ИНСТИТУТ ЗА ФИЗИКУ

ПРИМЉЕНО: 01. 07. 2021

Рад.јед.	број	Арх.шифра	Прилог
0901	488/1		

Научном већу Института за физику у Београду

Предмет: Покретање поступка за избор у звање научни сарадник

Молим Научно веће Института за физику у Београду да покрене поступак за мој избор у звање научни сарадник.

У прилогу достављам:

1. мишљење руководиоца лабораторије са предлогом комисије за избор у звање;
2. биографске и стручне податке;
3. преглед научне активности;
4. елементе за квалитативну оцену научног доприноса;
5. елементе за квантитативну оцену научног доприноса;
6. списак објављених радова;
7. податке о цитираности;
8. копије објављених радова;
9. копију докторске дисертације;
10. доказ о признању дипломе докторских студија;
11. докази о преосталим елементима оцена научног доприноса.

С поштовањем,

Марија Јанковић

Марија Јанковић

ИНСТИТУТ ЗА ФИЗИКУ

ПРИМЉЕНО: 01. 07. 2021

Рад.јед.	број	Арх.шифра	Прилог
0801	486/1		

Научном већу Института за физику у Београду

Предмет: Мишљење руководиоца лабораторије о избору др Марије Јанковић у звање научна сарадница

Др Марија Јанковић је заинтересована за сарадњу са Лабораторијом за примену рачунара у науци, у оквиру Националног центра изузетних вредности за изучавање комплексних система Института за физику у Београду. У истраживачком раду бави се темама везаним за проучавање теорије акреционих дискова око младих звезда и теорије формирања планета помоћу нумеричких симулација. С обзиром да испуњава све предвиђене услове у складу са Правилником о стицању истраживачких и научних звања МПНТР, сагласан сам са покретањем поступка за избор др Марије Јанковић у звање научна сарадница.

За састав комисије за избор др Марије Јанковић у звање научна сарадница предлажем:

(1) др Вељко Дмитрашиновић, научни саветник, Институт за физику у Београду
(2) др Марко Сталевски, виши научни сарадник, Астрономска опсерваторија у Београду
(3) др Антун Балаж, научни саветник, Институт за физику у Београду

др Антун Балаж
научни саветник
Руководилац Лабораторије за примену рачунара у науци

2. БИОГРАФСКИ И СТРУЧНИ ПОДАЦИ

Марија Јанковић је рођена 30.7.1991. у Београду, где је завршила основну школу „Раде Драинац“ и Рачунарску гимназију. Основне академске студије на Физичком факултету Универзитета у Београду, смер Теоријска и експериментална физика, завршила је 2015. године са просечном оценом 9,95. Мастер академске студије на истом факултету и смеру завршила је 2016. године са просечном оценом 10,00. Мастер рад на тему „*New periodic solutions of the three-body problem*“ („Нова периодична решења проблема три тела“) урадила је под менторством др Вељка Дмитрашиновића, научног саветника Института за физику у Београду. По завршетку мастер академских студија, уписала је докторске студије из области физике на Империјал Колеџ Лондон. Докторску дисертацију на тему „*Accretion discs and planet formation around young stars*“ („Акрециони дискови и формирање планета око младих звезда“) урадила је под менторством др Субханџој Мохантија, а одбранила у августу 2020. године. Диплома докторских студија призната је од стране Агенције за квалификације Републике Србије 21.4.2021. године, решење број 612-01-03-3-281/2021.

Током основних студија, током лета 2013. и 2014. године радила је стручну праксу на Бабрахам институту у Кембриџу у УК, односно на Институту за чисту и примењену математику у Лос Анђелесу у САД. Током докторских студија, током 2017. године радила је као асистент-демонстратор на курсу Рачунарска физика на основним академским студијама Империјал Колеџа Лондон. Од априла до септембра 2020. године била је запослена као научни асистент на Империјал Колеџ Лондон у оквиру *European Research Council (ERC)* пројекта *PEVAP* под руководством др Џејмса Овена. По завршетку основних студија, од октобра 2020. године запослена је као научни сарадник на постдокторском усавршавању на Институту за астрономију Универзитета у Кембриџу под руководством др Марка Вајата, у оквиру консолидованог пројекта на овој институцији, бр. *ST/S000623/1*, додељеног од стране британске агенције *Science and Technology research Council (STFC)*. На овој позицији истражује динамику прашине у егзо-Кајперовим појасевима.

Током основних студија била је добитник стипендије „Проф др Ђорђе Живановић“ за изузетне резултате, додељене од стране Физичког факултета и Института за физику у Београду. По завршетку мастер студија била је добитник награде „Проф др Љубомир Ћирковић“ за најбољу мастер тезу 2015/2016 школске године, додељене од стране Физичког факултета. Током докторских студија била је стипендиста Председникове стипендије Империјал Колеџа Лондон, као и Фонда за младе таленте "Доситеја" Републике Србије за најбоље студенте у иностранству.

Марија је до сада објавила осам научних радова. Радови су до сада цитирани 42 пута, не рачунајући аутоцитате.

3. ПРЕГЛЕД НАУЧНЕ АКТИВНОСТИ

У досадашњем научном раду кандидаткиња се примарно бавила истраживачким темама из области гравитације и астрофизике. Додатно, током основних и мастер академских студија, кандидаткиња је проучавала теме из физике, односно ужих области класичне физике и небеске механике и у једној студији учествовала у проучавању теме из физике плазме. Методолошки приступ кандидаткиње су нумеричке симулације.

Током докторских студија кандидаткиња се бавила проучавањем теорије акреционих дискова око младих звезда и теорије формирања планета. У протеклој деценији откривено је неколико хиљада планета ван Сунчевој система. Међу њима, најбројније су такозване супер-Земље, планете веће од Земље, а мање од Нептуна, које су од својих матичних звезда удаљене мање него Меркур од Сунца. Упркос бројности супер-Земљи, за сада не постоји прихваћено објашњење тога како су ове планете могле настати на тако малим удаљеностима од својих звезда. Планете примарно настају унутар протопланетарних дискова, дискова гаса и прашине који окружују младе, тек формиране звезде. Кандидаткиња се бавила нумеричким моделовањем делова протопланетарних дискова најближих централној звезди и истраживањем постојања услова за формирање супер-Земљи у тим деловима. Резултати истраживања објављени су у три студије:

- Subhanjoy Mohanty, Marija R. Jankovic, Jonathan C. Tan, and James E. Owen, Inside-out Planet Formation. V. Structure of the Inner Disk as Implied by the MRI, Astrophys. J. **861**, 144 (2018) (врхунски међународни часопис - категорија M21)
- Marija R. Jankovic, James E. Owen, and Subhanjoy Mohanty, Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc, Mon. Not. R. Astron. Soc. **484**, 2296 (2019) (врхунски међународни часопис - категорија M21)
- Marija R. Jankovic, James E. Owen, Subhanjoy Mohanty, Jonathan C. Tan, MRI-active inner regions of protoplanetary discs. I. A detailed model of disc structure, Mon. Not. R. Astron. Soc. **504**, 280 (2021) (врхунски међународни часопис - категорија M21)

У првој студији представљен је модел унутрашњих делова протопланетарних дискова који је само-доследно повезао структуру гаса у диску са акрецијом (акумулацијом) гаса из диска на звезду насталом услед магнетно-ротационе нестабилности (енг. *MRI*). Добијени модел структуре гаса коришћен је у другој студији, која се фокусирала на питање постојања услова за формирање планета. Прво, показано је да је настала структура гаса повољна за локалну акумулацију прашине, што је неопходан услов за настанак чврстих језгара планета. Друго, показано је да је разматрани модел гаса повољнији од конвенционалних модела у погледу помирења теорије настанка планетних атмосфера и маса атмосфера одређених посматрањем код детектованих супер-Земљи. Разлике у предвиђеним и одређеним величинама нису у потпуности објашњене, те ова студија указује на неопходност унапређења модела настанка планетних атмосфера. Коначно, трећа наведена студија је значајно унапредила многе аспекте модела из прве студије, разматрајући многе физичке и хемијске процесе који су у првој студији занемарени. Студија је показала да је температура у диску примарно одређена топлотом ослобођеном у процесу акреције диска, као и да је вертикална структура диска подложна развоју конвекције. Додатно, у складу са претходним предлозима у литератури, унутрашњи делови диска су примарно јонизовани термионичком и емисијом јона са честица прашине. Резултати студије показали су да је структура гаса у унутрашњим деловима диска одређена критичном температуром изнад које термионичка и емисија јона постају ефикасне. Овом студијом постављен је темељ за будућа истраживања међу-зависности структуре гаса и структуре прашине и акреције у унутрашњим деловима протопланетарних дискова.

У првој студији, кандидаткиња је унапредила нумеричке методе коришћених симулација и произвела и припремила резултате симулација за објављивање. У другој студији, кандидаткиња је имплементирала нумерички прорачун везан за настанак и еволуцију планетних атмосфера, произвела све резултате студије, значајно допринела анализи резултата и водила писање објављеног рада. У трећој студији, кандидаткиња је водила одабир метода студије, израдила нумеричке симулације коришћене у студији, произвела резултате и водила анализу резултата и писање објављеног рада.

У оквиру исте научне области, кандидаткиња се такође бавила темом моделовања акреционих дискова око масивних младих звезда, у сврхе предвиђања резултата интерферометарских детекција ових система, као и техникама обраде тих резултата. Главни резултат кандидаткиње унутар ове истраживачке теме тиче се могућности детекције спиралних структура које под одређеним условима могу настати у овим системима, односно тога како могућност детекције зависи од разних параметара система. Кандидаткиња је показала да АЛМА интерферометар може детектовати спиралне структуре само у најмасивнијим дисковима, код којих ефекти само-гравитације изазивају максималне пертурбације гаса и прашине, и то само уз одговарајуће технике филтрирања просторних и кинематичких података. Значај модела и алата развијених у овој студији огледа се и у њиховој примени при састављању предлога за пројекте посматрања АЛМА интерферометром. Кандидаткиња је повезала постојеће моделе акреционих дискова са алатима за тзв. синтетичка посматрања и алатима за анализу интерферометарских резултата, произвела већину резултата студије и водила писање публикације:

- Marija R. Jankovic, Thomas J. Haworth, John D. Ilee, Duncan H. Forgan, Claudia J. Cyganowski, Catherine Walsh, Crystal L. Brogan, Todd R. Hunter, Subhanjoy Mohanty, Observing substructure in circumstellar discs around massive young stellar objects, Mon. Not. R. Astron. Soc. **482**, 4673 (2019) (врхунски међународни часопис - категорија M21)

Током мастер студија, студенткиња се бавила проучавањем проблема три тела у класичној механици. Овај проблем тиче се три тачкасте масе које се међусобно гравитационо привлаче. За разлику од класичног проблема два тела, периодична решења проблема три тела су многобројна и тополошки разноврсна. Огроман број ових решења пронађен је нумеричким симулацијама у последњих неколико година, што је довело до открића повезаности кинематичких и тополошких особина ових решења. Унутар једне студије у којој је кандидаткиња учествовала у развоју нумеричких симулација и водила припрему резултата за публикацију, кандидаткиња је показала да је таква повезаност особина периодичних решења само приближна у специјалном случају тзв. БХХ фамилије решења:

- Marija R. Jankovic, V. Dmitrasinovic, Angular momentum and topological dependence of Kepler's third law in the Broucke-Hadjidemetriou-Henon family of periodic three-body orbits, Phys. Rev. Lett. **116**, 064301 (2016) (међународни часопис изузетних вредности - категорија M21a),
- Marija R. Jankovic, V. Dmitrasinovic, and Milovan Suvakov, A guide to hunting periodic three-body orbits with non-vanishing angular momentum, Comput. Phys. Commun. **250**, 107052 (2020) (међународни часопис изузетних вредности - категорија M21a).

Током основних студија, кандидаткиња је учествовала у једној студији из области физике плазме, у којој је као део већег тима допринела аналитичком поједностављењу, односно спектралном развоју квантне Ленард-Балеску једначине у кинетичкој теорији плазме. Студија је демонстрирала предности овог приступа за нумеричко решавање кинетичких једначина. Резултати студије објављени су у публикацији:

- Christian R. Scullard, Andrew P. Belt, Susan C. Fennell, Marija R. Janković, Nathan Ng, Susana Serna, and Frank R. Graziani, Numerical solution of the quantum Lenard-Balescu equation for a non-degenerate one-component plasma, Phys. Plasmas **23**, 092119 (2016) (истакнути међународни часопис - категорија M22)

4. ЕЛЕМЕНТИ ЗА КВАЛИТАТИВНУ ОЦЕНУ НАУЧНОГ ДОПРИНОСА КАНДИДАТА

4.1. Квалитет научних резултата

4.1.1. Научни ниво и значај резултата, утицај научних радова

Кандидаткиња је у досадашњој каријери објавила 8 научних радова, од чега 2 рада категорије М21а, 5 радова категорије М21 и 1 рад категорије М22. Своја истраживања је представила и на 4 конференције, тј. остварила 4 доприноса категорије М34.

До сада најутицајнији рад кандидаткиње из теме доктората је:

- Marija R. Jankovic, James E. Owen, and Subhanjoy Mohanty, Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc, Mon. Not. R. Astron. Soc. **484**, 2296 (2019)
  DOI: https://doi.org/10.1093/mnras/stz004

Тема рада је тестирање хипотезе да је најбројнија класа до сада откривених планета, тзв. супер-Земље, настала акумулацијом прашине и гаса у унутрашњим деловима протопланетарних дискова, те да кључну улогу у том процесу игра магнетно-ротациона нестабилност. Кандидаткиња је произвела све резултате објављене у овом раду и водила интерпретацију и дискусију резултата и писање рада.

У студији која је претходила наведеној, кандидаткиња је произвела модел структуре гаса у унутрашњим деловима протопланетарних дискова. У наведеној студији, користећи нумеричке симулације развијене од стране сарадника, кандидаткиња је истраживала како та структура гаса утиче на еволуцију честица прашине. Резултати симулација предвиђају да количина прашине у унутрашњим деловима протопланетарних дискова расте током времена, што је предуслов за акумулацију прашине у чврста језгра планета.

Под претпоставком да овај процес успешно формира чврста језгра планета, кандидаткиња је имплементирала модел акумулације гаса на језгро планете, односно формирање атмосфере планете. Теоријска предвиђања овог модела упоређена су са посматрачким мерењима масе и радијуса једне групе детектованих планета. Како акумулација гаса на језгро планете зависи од структуре гаса у протопланетарном диску, ово је дозволило индиректну проверу коришћеног модела диска. У овом смислу, за коришћени модел показано је да је повољнији од конвенционалних модела који не узимају у обзир утицај магнетно-ротационе нестабилности. Међутим, неслагања између теорије и мерења нису у потпуности уклоњена, те ови резултати такође указују на неопходност бољег разумевања процеса акумулације и еволуције атмосфера планета.

4.1.2. Цитираност научних радова кандидата

Према подацима о цитираности аутора изведених из базе *Web of Science* 28.6.2021., радови чији је кандидаткиња ко-аутор цитирани су 46 пута, од чега 42 пута без аутоцитата, а Хиршов фактор је 5.

4.1.3. Параметри квалитета радова и часописа

Кандидаткиња је објавила 8 радова у часописима:

- 4 рада у часопису *Monthly Notices of the Royal Astronomical Society* (ISSN: 0035-8711), категорија М21, IF (2019) = 5.357, SNIP (2020) = 1.09, SNIP (2019) = 1.13;

- 1 рад у часопису *Computer Physics Communications* (ISSN: 0010-4655), категорија M21a, IF (2020) = 4.390, SNIP (2020) = 1.68;
- 1 рад у часопису *Astrophysical Journal* (ISSN: 0004-637X), категорија M21, IF (2018) = 5.580, SNIP (2018) = 1.22;
- 1 рад у часопису *Physical Review Letters* (ISSN: 0031-9007), категорија M21a, IF (2016) = 8.462, SNIP (2016) = 2.62;
- 1 рад у часопису *Physics of Plasmas* (ISSN: 1070-664X), категорија M22, IF (2016) = 2.115, SNIP (2016) = 1.16.

Додатни библиометријски показатељи квалитета часописа у којима је кандидаткиња објављивала радове приказани су у табели:

	ИФ	М	СНИП
Укупно	41.975	65	11.12
Усредњено по чланку	5.246	8.125	1.39
Усредњено по аутору	12.183	18.20	3.33

4.1.4. Степен самосталности и степен учешћа у реализацији радова у научним центрима у земљи и иностранству

Кандидаткиња је водећи аутор на 5 објављених радова, којима је дала кључан допринос у погледу израде резултата нумеричких симулација и њиховој интерпретацији, као и допринос писању радова. Од преосталих објављених радова, у 1 раду кандидаткиња је дала кључан допринос у погледу израде резултата нумеричких симулација, у 1 раду допринос почетној теоретској анализи проблема, а у 1 раду допринос интерпретацији резултата.

Кандидаткиња је већину досадашњих научних активности обављала на Империјал Колеџу Лондон у Великој Британији. Кроз наведене доприносе остварила је сарадњу, између осталог, са истраживачима са Института за физику у Београду, Чалмерс Универзитета у Шведској, Харвард Универзитета у САД и Универзитета у Кембриџу у Великој Британији.

4.2. Нормирање броја коауторских радова, патената и техничких решења

Од 8 радова, 6 радова има пет или мање аутора, док 1 рад има 7, а један 9 аутора. Сви радови су из области нумеричких симулација, тако да се 6 радова рачуна са пуним бројем бодова, а за преостала 2 рада број бодова се нормира. Укупан број М бодова је 73, а нормирани број је 68.02.

4.3. Утицај научних резултата

Утицај научних резултата огледа се у подацима о цитираности, наведеним у секцији 4.1.2.

Кандидаткиња је одржала и 2 семинара по позиву на којима је представила своје научне резултате, на Чалмерс Универзитету у Шведској 2018. године и на Институту за астрономију Универзитета у Кембриџу у Великој Британији 2019. године.

4.4. Конкретан допринос кандидата у реализацији радова у научним центрима у земљи и иностранству

У 5 објављених радова, кандидаткиња је дала кључан допринос у погледу израде резултата нумеричких симулација и њиховој интерпретацији, као и допринос писању радова.

Кандидаткиња је већину досадашњих научних активности обављала на Империјал Колеџу Лондон у Великој Британији, као студент докторских студија. Преостале научне активности кандидаткиња је обављала као мастер студент на Институту за физику у Београду, као студент на летњој пракси на Институту за чисту и примењену физику у САД и као научни сарадник на Универзитету у Кембриџу у Великој Британији.

5. ЕЛЕМЕНТИ ЗА КВАНТИТАТИВНУ ОЦЕНУ НАУЧНОГ ДОПРИНОСА КАНДИДАТА

Остварени резултати у периоду након одлуке Научног већа о предлогу за стицање претходног научног звања:

Категорија	**М бодова по раду**	**Број радова**	**Укупно М бодова**	**Нормирани број М бодова**
M21a	10	2	20	20
M21	8	5	40	36.44
M22	5	1	5	3.57
M34	0.5	4	2	2
M70	6	1	6	6

Поређење са минималним квантитативним условима за избор у звање научни сарадник:

Минимални број М бодова	Неопходно	Остварено, број М бодова без нормирања	**Остварено, нормирани број М бодова**
Укупно	16	73	**68.02**
M10+M20+M31+M32+M33+M41+M42	10	65	**60.02**
M11+M12+M21+M22+M23	6	65	**60.02**

6. СПИСАК ОБЈАВЉЕНИХ РАДОВА

6.1 Радови у међународним часописима изузетних вредности (M21a):

- **Marija R. Jankovic**, V. Dmitrasinovic, Angular momentum and topological dependence of Kepler's third law in the Broucke-Hadjidemetriou-Henon family of periodic three-body orbits, Phys. Rev. Lett. 116, 064301 (2016)
  DOI: https://doi.org/10.1103/PhysRevLett.116.064301
  M21a, IF (2016) = 8.462, SNIP (2016) = 2.62

- **Marija R. Jankovic,** V. Dmitrasinovic, and Milovan Suvakov, A guide to hunting periodic three-body orbits with non-vanishing angular momentum, Comput. Phys. Commun. 250, 107052 (2020)
  DOI: https://doi.org/10.1016/j.cpc.2019.107052
  M21a, IF (2020) = 4.390, SNIP (2020) = 1.68

6.2 Радови у врхунским међународним часописима (M21):

- **Marija R. Jankovic**, James E. Owen, Subhanjoy Mohanty, Jonathan C. Tan, MRI-active inner regions of protoplanetary discs. I. A detailed model of disc structure, Mon. Not. R. Astron. Soc. 504, 280 (2021)
  DOI: https://doi.org/10.1093/mnras/stab920
  M21, IF (2019) = 5.357, SNIP (2020) = 1.09

- Richard Teague, **Marija R. Jankovic**, Thomas J. Haworth, Chunhua Qi, John D. Ilee, A three-dimensional view of Gomez’s hamburger, Mon. Not. R. Astron. Soc. 495, 451 (2020)
  DOI: https://doi.org/10.1093/mnras/staa1167
  M21, IF (2019) = 5.357, SNIP (2020) = 1.09

- **Marija R. Jankovic**, James E. Owen, and Subhanjoy Mohanty, Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc, Mon. Not. R. Astron. Soc. 484, 2296 (2019)
  DOI: https://doi.org/10.1093/mnras/stz004
  M21, IF (2019) = 5.357, SNIP (2019) = 1.13

- **Marija R. Jankovic**, Thomas J. Haworth, John D. Ilee, Duncan H. Forgan, Claudia J. Cyganowski, Catherine Walsh, Crystal L. Brogan, Todd R. Hunter, Subhanjoy Mohanty, Observing substructure in circumstellar discs around massive young stellar objects, Mon. Not. R. Astron. Soc. 482, 4, 4673–4686 (2019)
  DOI: https://doi.org/10.1093/mnras/sty3038
  M21, IF (2019) = 5.357, SNIP (2019) = 1.13

- Subhanjoy Mohanty, **Marija R. Jankovic**, Jonathan C. Tan, and James E. Owen, Inside-out Planet Formation. V. Structure of the Inner Disk as Implied by the MRI, Astrophys. J. 861, 144 (2018)
  DOI: https://doi.org/10.3847/1538-4357/aabcd0
  M21, IF (2018) = 5.580, SNIP (2018) = 1.22

6.3 Радови у истакнутим међународним часописима (M22):

- Christian R. Scullard, Andrew P. Belt, Susan C. Fennell, **Marija R. Janković**, Nathan Ng, Susana Serna, and Frank R. Graziani, Numerical solution of the quantum Lenard-Balescu equation for a non-degenerate one-component plasma, Phys. Plasmas 23, 092119 (2016)
  DOI: https://doi.org/10.1063/1.4963254
  M22, IF (2016) = 2.115, SNIP (2016) = 1.16

6.4 Саопштења са међународног скупа штампана у изводу (M34):

- M. Jankovic, S. Mohanty, J. Owen, J. Tan
  In situ formation of close-in super-Earths due to the MRI
  From Stars to Planets II, Chalmers, Sweden, 17-21 June 2019
- M. Jankovic, J. Owen, S. Mohanty
  Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc
  UK Exoplanet Community Meeting, London, UK, 15-17 April 2019
- M. Jankovic, J. Owen, S. Mohanty
  Formation of close-in super-Earths: dust enrichment of the inner disk due to the MRI
  ESO workshop “Take a Closer Look”, Garching, Germany, 15-19 October 2018
- M. Jankovic, J. Owen, S. Mohanty
  Dust enhancement in the inner disk due to the MRI
  UK Exoplanet Community Meeting, Oxford, UK, 21-23 March, 2018

Web of Science InCites Journal Citation Reports Essential Science Indicators EndNote Publons Kopernio Master Journal List Sign In Help English

Web of Science

Clarivate Analytics

Search Tools Searches and alerts Search History Marked List

Citation report for 8 results from Web of Science Core Collection between 1900 and 2021 Go

You searched for: **Article Group for:** Jankovic, Marija R. ...More

This report reflects citations to source items indexed within Web of Science Core Collection. Perform a Cited Reference Search to include citations to items not indexed within Web of Science Core Collection.

Export Data: Save to Excel File

Total Publications

8 Analyze

h-index

5

Average citations per item

5.75

Sum of Times Cited

46

Without self citations

42

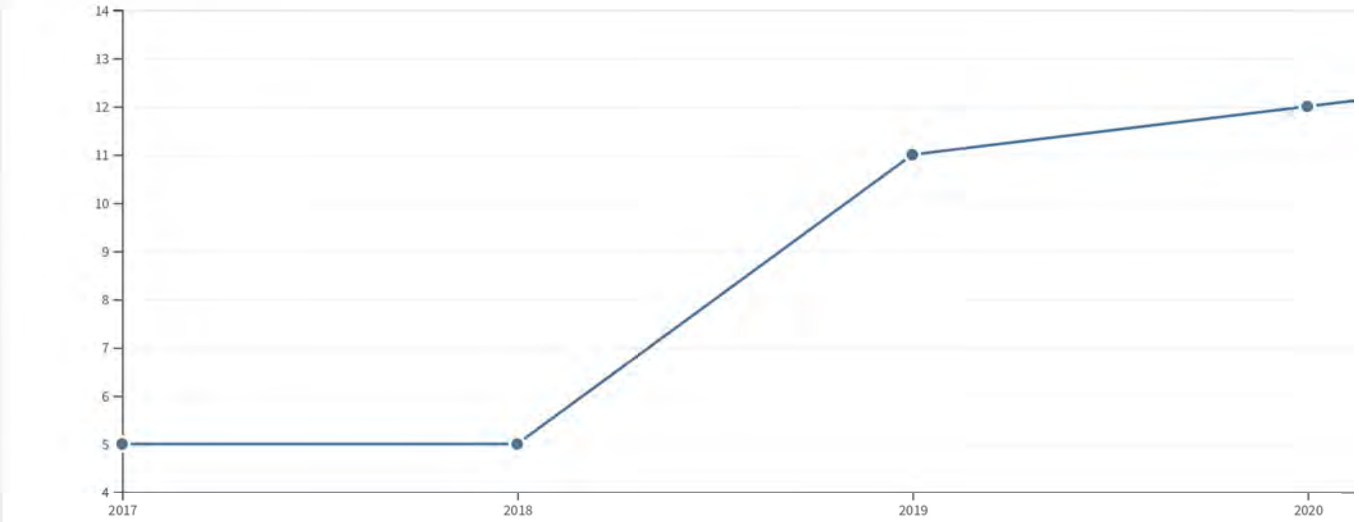
Citing articles

44 Analyze

Without self citations

41 Analyze

Sum of Times Cited per Year



Sort by: Times Cited Date More 1 of 1

How are these totals calculated?

2017 ◀	2018	2019	2020	2021 ▶	Total	Average Citations per Year

Use the checkboxes to remove individual items from this Citation Report or restrict to items published between 1900 and 2021 Go	5	5	11	12	13	46	9.20
☐ 1. Observing substructure in circumstellar discs around massive young stellar objects By: Jankovic, M. R.; Haworth, T. J.; Ilee, J. D.; et al. MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 482 Issue: 4 Pages: 4673-4686 Published: FEB 2019	0	0	6	4	2	12	4.00
☐ 2. Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc By: Jankovic, Marija R.; Owen, James E.; Mohanty, Subhanjoy MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 484 Issue: 2 Pages: 2296-2308 Published: APR 2019	0	0	2	3	3	8	2.67
☐ 3. Inside-out Planet Formation. V. Structure of the Inner Disk as Implied by the MRI By: Mohanty, Subhanjoy; Jankovic, Marija R.; Tan, Jonathan C.; et al. ASTROPHYSICAL JOURNAL Volume: 861 Issue: 2 Article Number: 144 Published: JUL 10 2018	0	2	3	1	2	8	2.00
☐ 4. Angular Momentum and Topological Dependence of Kepler's Third Law in the Broucke-Hadjidemetriou-Henon Family of Periodic Three-Body Orbits By: Jankovic, Marija R.; Dmitrasinovic, V. PHYSICAL REVIEW LETTERS Volume: 116 Issue: 6 Article Number: 064301 Published: FEB 10 2016	3	2	0	1	2	8	1.33
☐ 5. Numerical solution of the quantum Lenard-Balescu equation for a non-degenerate one-component plasma By: Scullard, Christian R.; Belt, Andrew P.; Fennell, Susan C.; et al. PHYSICS OF PLASMAS Volume: 23 Issue: 9 Article Number: 092119 Published: SEP 2016	2	1	0	1	1	5	0.83
☐ 6. A three-dimensional view of Gomez's hamburger By: Teague, Richard; Jankovic, Marija R.; Haworth, Thomas J.; et al. MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 495 Issue: 1 Pages: 451-459 Published: JUN 2020	0	0	0	1	2	3	1.50
☐ 7. A guide to hunting periodic three-body orbits with non-vanishing angular momentum By: Jankovic, Marija R.; Dmitrasinovic, V; Suvakov, Milovan COMPUTER PHYSICS COMMUNICATIONS Volume: 250 Article Number: 107052 Published: MAY 2020	0	0	0	1	1	2	1.00
☐ 8. MRI-active inner regions of protoplanetary discs. I. A detailed model of disc structure By: Jankovic, Marija R.; Owen, James E.; Mohanty, Subhanjoy; et al. MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 504 Issue: 1 Pages: 280-299 Published: JUN 2021	0	0	0	0	0	0	0.00

☐ Select Page Save to Excel File

Sort by: Times Cited Date More 1 of 1

8 records matched your query of the 80,234,500 in the data limits you selected.
Key: = Structure available.

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Web of Science | InCites | Journal Citation Reports | Essential Science Indicators | EndNote | Publons | Kopernio | Master Journal List | Sign In | Help | English

Web of Science

Clarivate Analytics

Search | Tools | Searches and alerts | Search History | Marked List

Citing Articles: 12

(from Web of Science Core Collection)

For: Observing substructure in circumstellar discs around massive young stellar objects ...Less

Times Cited Counts

- 12 in All Databases
- 12 in Web of Science Core Collection
- 0 in Arabic Citation Index
- 0 in BIOSIS Citation Index
- 0 in Chinese Science Citation Database
- 0 data sets in Data Citation Index
- 0 publication in Data Citation Index
- 0 in Russian Science Citation Index
- 0 in SciELO Citation Index

View Additional Times Cited Counts

Refine Results

Search within results for...

Filter results by:

☐ Open Access (6)

Refine

Publication Years

☐ 2021 (2)
☐ 2020 (4)
☐ 2019 (6)

more options / values...

Refine

Web of Science Categories

☐ ASTRONOMY ASTROPHYSICS (12)
☐ COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS (1)

more options / values...

Refine

Document Types

☐ ARTICLE (12)

Refine

Organizations-Enhanced

☐ MAX PLANCK SOCIETY (4)
☐ EBERHARD KARLS UNIVERSITY OF TUBINGEN (3)
☐ UNIVERSITY OF LEICESTER (3)

Refine

Sort by: **Date** | Times Cited | Usage Count | More

1 of 1

☐ Select Page | Print | More | Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **Searching for wide-orbit gravitational instability protoplanets with ALMA in the dust continuum**

By: Humphries, J.; Hall, C.; Haworth, T. J.; et al.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 502 Issue: 1 Pages: 953-968 Published: MAR 2021

ejournals@cambridge - find full tex | View Abstract

Times Cited: 1 *(from Web of Science Core Collection)*

Usage Count

☐ 2. **Parameter study for the burst mode of accretion in massive star formation**

By: Meyer, D. M-A; Vorobyov, E., I; Elbakyan, V. G.; et al.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 500 Issue: 4 Pages: 4448-4468 Published: FEB 2021

ejournals@cambridge - find full tex | Full Text from Publisher | View Abstract

Times Cited: 1 *(from Web of Science Core Collection)*

Usage Count

☐ 3. **Analytical solutions of radiative transfer equations in accretion discs with finite optical depth**

By: Samadi, M.; Habibi, F.; Abbassi, S.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 496 Issue: 2 Pages: 1655-1666 Published: AUG 2020

ejournals@cambridge - find full tex | Full Text from Publisher | View Abstract

Times Cited: 0 *(from Web of Science Core Collection)*

Usage Count

☐ 4. **New maser species tracing spiral-arm accretion flows in a high-mass young stellar object**

By: Chen, Xi; Sobolev, Andrej M.; Ren, Zhi-Yuan; et al.

NATURE ASTRONOMY Volume: 4 Issue: 12 Published: DEC 2020

Early Access: JUL 2020

ejournals@cambridge - find full tex | Full Text from Publisher | View Abstract

Times Cited: 5 *(from Web of Science Core Collection)*

Usage Count

☐ 5. **Massive discs around low-mass stars**

By: Haworth, Thomas J.; Cadman, James; Meru, Farzana; et al.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 494 Issue: 3 Pages: 4130-4148 Published: MAY 2020

ejournals@cambridge - find full tex | Free Published Article From Repository

View Abstract

Times Cited: 9 *(from Web of Science Core Collection)*

Usage Count

☐ 6. **Constraining planet formation around 6-8M(circle dot) stars**

By: Veras, Dimitri; Tremblay, Pier-Emmanuel; Hermes, J. J.; et al.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 493 Issue: 1 Pages: 765-775 Published: MAR 2020

ejournals@cambridge - find full tex | Free Accepted Article From Repository

View Abstract

Times Cited: 8 *(from Web of Science Core Collection)*

Usage Count

☐ CHINESE ACADEMY OF SCIENCES (2)
☐ IMPERIAL COLLEGE LONDON (2)

more options / values...

Funding Agencies

Authors

Source Titles

View all options

For advanced refine options, use

Analyze Results

☐ 7. **Disc kinematics and stability in high-mass star formation Linking simulations and observations**

By: Ahmadi, A.; Kuiper, R.; Beuther, H.

ASTRONOMY & ASTROPHYSICS Volume: 632 Article Number: A50 Published: NOV 26 2019

ejournals@cambridge - find full tex Free Full Text from Publisher

View Abstract

Times Cited: 7
(from Web of Science Core Collection)

Usage Count

☐ 8. **Fragmentation, rotation, and outflows in the high-mass star-forming region IRAS 23033+5951 A case study of the IRAM NOEMA large program CORE**

By: Bosco, F.; Beuther, H.; Ahmadi, A.; et al.

Group Author(s): CORE Team

ASTRONOMY & ASTROPHYSICS Volume: 629 Article Number: A10 Published: AUG 23 2019

ejournals@cambridge - find full tex Free Full Text from Publisher

View Abstract

Times Cited: 5
(from Web of Science Core Collection)

Usage Count

☐ 9. **On the ALMA observability of nascent massive multiple systems formed by gravitational instability**

By: Meyer, D. M-A; Kreplin, A.; Kraus, S.; et al.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 487 Issue: 4 Pages: 4473-4491 Published: AUG 2019

ejournals@cambridge - find full tex Full Text from Publisher View Abstract

Times Cited: 4
(from Web of Science Core Collection)

Usage Count

☐ 10. **Substructures in the Keplerian disc around the O-type (proto-)star G17.64+0.16**

By: Maud, L. T.; Cesaroni, R.; Kumar, M. S. N.; et al.

ASTRONOMY & ASTROPHYSICS Volume: 627 Article Number: L6 Published: JUL 9 2019

ejournals@cambridge - find full tex Free Full Text from Publisher

View Abstract

Times Cited: 12
(from Web of Science Core Collection)

Usage Count

☐ 11. **Massive star formation via torus accretion: the effect of photoionization feedback**

By: Sartorio, N. S.; Vandenbroucke, B.; Falceta-Goncalves, D.; et al.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 486 Issue: 4 Pages: 5171-5183 Published: JUL 2019

ejournals@cambridge - find full tex Full Text from Publisher View Abstract

Times Cited: 5
(from Web of Science Core Collection)

Usage Count

☐ 12. **The TORUS radiation transfer code**

By: Harries, T. J.; Haworth, T. J.; Acreman, D.; et al.

ASTRONOMY AND COMPUTING Volume: 27 Pages: 63-95 Published: APR 2019

ejournals@cambridge - find full tex Free Full Text from Publisher

View Abstract

Times Cited: 16
(from Web of Science Core Collection)

Usage Count

☐ Select Page Print More Add to Marked List

Sort by: Date Times Cited Usage Count More

1 of 1

Show: 50 per page

12 records matched your query of the 80,257,729 in the data limits you selected.

Key: = Structure available.

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Citing Articles: 8
(from Web of Science Core Collection)

For: Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc **...More**

Times Cited Counts

8 in All Databases
8 in Web of Science Core Collection
0 in Arabic Citation Index
0 in BIOSIS Citation Index
0 in Chinese Science Citation Database
0 data sets in Data Citation Index
0 publication in Data Citation Index
0 in Russian Science Citation Index
0 in SciELO Citation Index

View Additional Times Cited Counts

Refine Results

Search within results for...

Filter results by:

☐ Open Access (4)

Refine

Publication Years

☐ 2021 (3)
☐ 2020 (3)
☐ 2019 (2)

more options / values...

Refine

Web of Science Categories

☐ ASTRONOMY ASTROPHYSICS (7)
☐ GEOCHEMISTRY GEOPHYSICS (1)

more options / values...

Refine

Document Types

☐ ARTICLE (6)
☐ REVIEW (2)

more options / values...

Refine

Organizations-Enhanced

☐ IMPERIAL COLLEGE LONDON (8)
☐ AUSTRIAN ACADEMY OF SCIENCES (1)

Refine

Sort by: Date Times Cited Usage Count More 1 of 1

☐ Select Page Print More Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **MRI-active inner regions of protoplanetary discs. I. A detailed model of disc structure**
By: Jankovic, Marija R.; Owen, James E.; Mohanty, Subhanjoy; et al.
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY **Volume:** 504 **Issue:** 1 **Pages:** 280-299 **Published:** JUN 2021
ejournals@cambridge - find full tex Free Published Article From Repository
View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

☐ 2. **Unveiling the planet population at birth**
By: Rogers, James G.; Owen, James E.
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY **Volume:** 503 **Issue:** 1 **Pages:** 1526-1542 **Published:** MAY 2021
ejournals@cambridge - find full tex Free Published Article From Repository
View Abstract
Times Cited: 8 *(from Web of Science Core Collection)*
Usage Count

☐ 3. **The Nature and Origins of Sub-Neptune Size Planets**
By: Bean, Jacob L.; Raymond, Sean N.; Owen, James E.
JOURNAL OF GEOPHYSICAL RESEARCH-PLANETS **Volume:** 126 **Issue:** 1 **Article Number:** e2020JE006639 **Published:** JAN 2021
ejournals@cambridge - find full tex Free Full Text from Publisher
View Abstract
Times Cited: 6 *(from Web of Science Core Collection)*
Usage Count

☐ 4. **Hydrogen Dominated Atmospheres on Terrestrial Mass Planets: Evidence, Origin and Evolution**
By: Owen, J. E.; Shaikhislamov, I. F.; Lammer, H.; et al.
SPACE SCIENCE REVIEWS **Volume:** 216 **Issue:** 8 **Article Number:** 129 **Published:** NOV 4 2020
ejournals@cambridge - find full tex Free Full Text from Publisher
View Abstract
Times Cited: 3 *(from Web of Science Core Collection)*
Usage Count

☐ 5. **Constraining the entropy of formation from young transiting planet**
By: Owen, James E.
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY **Volume:** 498 **Issue:** 4 **Pages:** 5030-5040 **Published:** NOV 2020
ejournals@cambridge - find full tex View Abstract
Times Cited: 1 *(from Web of Science Core Collection)*
Usage Count

☐ 6. **Testing exoplanet evaporation with multitransiting systems**
By: Owen, James E.; Estrada, Beatriz Campos
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY **Volume:** 491 **Issue:** 4 **Pages:** 5287-5297 **Published:** FEB 2020
ejournals@cambridge - find full tex View Abstract
Times Cited: 17 *(from Web of Science Core Collection)*
Usage Count

☐ CARNEGIE INSTITUTION FOR SCIENCE (1)

☐ CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS (1)

☐ CHALMERS UNIVERSITY OF TECHNOLOGY (1)

more options / values...

Funding Agencies

Authors

Source Titles

View all options

For advanced refine options, use

Analyze Results

☐ 7. **Effects of magnetic fields on the location of the evaporation valley for low-mass exoplanets**

By: Owen, James E.; Adams, Fred C.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY **Volume:** 490 **Issue:** 1 **Pages:** 15-20 **Published:** NOV 2019

ejournals@cambridge - find full tex   View Abstract

Times Cited: 9
(from Web of Science Core Collection)

Usage Count

☐ 8. **Radiation pressure clear-out of dusty photoevaporating discs**

By: Owen, James E.; Kollmeier, Juna A.

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY **Volume:** 487 **Issue:** 3 **Pages:** 3702-3714 **Published:** AUG 2019

ejournals@cambridge - find full tex   Full Text from Publisher   View Abstract

Times Cited: 12
(from Web of Science Core Collection)

Usage Count

☐ Select Page   Print   More   Add to Marked List

Sort by: Date   Times Cited   Usage Count   More   1 of 1

Show: 50 per page

8 records matched your query of the 80,257,729 in the data limits you selected.
Key: = Structure available.

Clarivate
Accelerating innovation

© 2021 Clarivate   Copyright notice   Terms of use   Privacy statement   Cookie policy

Sign up for the Web of Science newsletter   Follow us

Citing Articles: 8

(from Web of Science Core Collection)

For: Inside-out Planet Formation. V. Structure of the Inner Disk as Implied by the MRI **...More**

Times Cited Counts

8 in All Databases
8 in Web of Science Core Collection
0 in Arabic Citation Index
0 in BIOSIS Citation Index
0 in Chinese Science Citation Database
0 data sets in Data Citation Index
0 publication in Data Citation Index
0 in Russian Science Citation Index
0 in SciELO Citation Index

View Additional Times Cited Counts

Refine Results

Search within results for...

Filter results by:

☐ Open Access (5)

Refine

Publication Years

☐ 2021 (2)
☐ 2020 (1)
☐ 2019 (3)
☐ 2018 (2)

more options / values...

Refine

Web of Science Categories

☐ ASTRONOMY ASTROPHYSICS (8)

Refine

Document Types

☐ ARTICLE (8)

Refine

Organizations-Enhanced

☐ IMPERIAL COLLEGE LONDON (3)
☐ MAX PLANCK SOCIETY (2)
☐ SPACE TELESCOPE SCIENCE INSTITUTE (2)
☐ AARHUS UNIVERSITY (1)
☐ ARC CTR EXCELLENCE ALL SKY ASTROPHYS ASTRO 3D (1)

more options / values...

Refine

Funding Agencies

Authors

Sort by: **Date** Times Cited Usage Count More 1 of 1

☐ Select Page Print More Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **MRI-active inner regions of protoplanetary discs. I. A detailed model of disc structure**
By: Jankovic, Marija R.; Owen, James E.; Mohanty, Subhanjoy; et al.
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 504 Issue: 1 Pages: 280-299 Published: JUN 2021
ejournals@cambridge - find full tex Free Published Article From Repository View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

☐ 2. **Orbital misalignment of the super-Earth pi Men c with the spin of its star**
By: Hodzic, Vedad Kunovac; Triaud, Amaury H. M. J.; Cegla, Heather M.; et al.
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 502 Issue: 2 Pages: 2893-2911 Published: APR 2021
ejournals@cambridge - find full tex Free Published Article From Repository View Abstract
Times Cited: 2 *(from Web of Science Core Collection)*
Usage Count

☐ 3. **Dynamics of small grains in transitional discs**
By: Krumholz, Mark R.; Ireland, Michael J.; Kratter, Kaitlin M.
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 498 Issue: 2 Pages: 3023-3042 Published: OCT 2020
ejournals@cambridge - find full tex Full Text from Publisher View Abstract
Times Cited: 1 *(from Web of Science Core Collection)*
Usage Count

☐ 4. **The GRAVITY Young Stellar Object survey I. Probing the disks of Herbig Ae/Be stars in terrestrial orbits**
By: Perraut, K.; Labadie, L.; Lazareff, B.; et al.
Group Author(s): GRAVITY Collaboration
ASTRONOMY & ASTROPHYSICS Volume: 632 Article Number: A53 Published: NOV 27 2019
ejournals@cambridge - find full tex Free Full Text from Publisher View Abstract
Times Cited: 4 *(from Web of Science Core Collection)*
Usage Count

☐ 5. **Complex Magnetospheric Accretion Flows in the Low Accretor CVSO 1335**
By: Thanathibodee, Thanawuth; Calvet, Nuria; Muzerolle, James; et al.
ASTROPHYSICAL JOURNAL Volume: 884 Issue: 1 Article Number: 86 Published: OCT 10 2019
ejournals@cambridge - find full tex View Abstract
Times Cited: 5 *(from Web of Science Core Collection)*
Usage Count

☐ 6. **Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc**
By: Jankovic, Marija R.; Owen, James E.; Mohanty, Subhanjoy
MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 484 Issue: 2 Pages: 2296-2308 Published: APR 2019
ejournals@cambridge - find full tex Full Text from Publisher Free Published Article From Repository View Abstract
Times Cited: 8 *(from Web of Science Core Collection)*
Usage Count

☐ 7. **Evolution of Circumbinary Protoplanetary Disks with Photoevaporative Winds Driven by External Far-ultraviolet Radiation**
By: Shadmehri, M.; Ghoreyshi, S. M.; Alipour, N.
ASTROPHYSICAL JOURNAL Volume: 867 Issue: 1 Article Number: 41 Published: NOV 1 2018
ejournals@cambridge - find full tex Full Text from Publisher View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

Source Titles ▼

View all options

For advanced refine options, use

Analyze Results

☐ 8. Photoevaporation and high-eccentricity migration created the sub-Jovian desert

By: Owen, James E.; Lai, Dong

MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY Volume: 479 Issue: 4 Pages: 5012-5021
Published: OCT 2018

ejournals@cambridge - find full tex Free Full Text from Publisher View Abstract ▼

Times Cited: 52
(from Web of Science Core Collection)

Usage Count ⌄

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Web of Science InCites Journal Citation Reports Essential Science Indicators EndNote Publons Kopernio Master Journal List Sign In Help English

Web of Science

Clarivate Analytics

Search Tools Searches and alerts Search History Marked List

Citing Articles: 8

(from Web of Science Core Collection)

For: Angular Momentum and Topological Dependence of Kepler's Third Law in the Broucke-Hadjidemetriou-Heno
...More

Times Cited Counts

9 in All Databases
8 in Web of Science Core Collection
0 in Arabic Citation Index
0 in BIOSIS Citation Index
2 in Chinese Science Citation Database
0 data sets in Data Citation Index
0 publication in Data Citation Index
0 in Russian Science Citation Index
0 in SciELO Citation Index

View Additional Times Cited Counts

Refine Results

Search within results for...

Filter results by:

☐ Open Access (1)

Refine

Publication Years

☐ 2021 (2)
☐ 2020 (1)
☐ 2018 (2)
☐ 2017 (3)

more options / values...

Refine

Web of Science Categories

☐ PHYSICS MATHEMATICAL (4)
☐ PHYSICS MULTIDISCIPLINARY (4)
☐ COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS (1)
☐ MECHANICS (1)
☐ OPTICS (1)

more options / values...

Refine

Document Types

Refine

Sort by: Date Times Cited Usage Count More 1 of 1

☐ Select Page Print More Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **On "generalized Kepler's third law" and mass dependence of periods of three-body orbits**

By: Raonic, Bogdan; Dmitrasinovic, V.

MECCANICA Volume: 56 Issue: 5 Pages: 1011-1024 Published: MAY 2021

Early Access: MAR 2021

ejournals@cambridge - find full tex View Abstract

Times Cited: 0
(from Web of Science Core Collection)

Usage Count

☐ 2. **One family of 13315 stable periodic orbits of non-hierarchical unequal-mass triple systems**

By: Li, XiaoMing; Li, XiaoChen; Liao, ShiJun

SCIENCE CHINA-PHYSICS MECHANICS & ASTRONOMY Volume: 64 Issue: 1 Article Number: 219511 Published: JAN 2021

ejournals@cambridge - find full tex View Abstract

Times Cited: 0
(from Web of Science Core Collection)

Usage Count

☐ 3. **A guide to hunting periodic three-body orbits with non-vanishing angular momentum**

By: Jankovic, Marija R.; Dmitrasinovic, V; Suvakov, Milovan

COMPUTER PHYSICS COMMUNICATIONS Volume: 250 Article Number: 107052 Published: MAY 2020

ejournals@cambridge - find full tex Full Text from Publisher

View Abstract

Times Cited: 2
(from Web of Science Core Collection)

Usage Count

☐ 4. **Periodic three-body orbits in the Coulomb potential**

By: Sindik, Marija; Sugita, Ayumu; Suvakov, Milovan; et al.

PHYSICAL REVIEW E Volume: 98 Issue: 6 Article Number: 060101 Published: DEC 7 2018

ejournals@cambridge - find full tex View Abstract

Times Cited: 5
(from Web of Science Core Collection)

Usage Count

☐ 5. **Linear stability of periodic three-body orbits with zero angular momentum and topological dependence of Kepler's third law: a numerical test**

By: Dmitrasinovic, V; Hudomal, Ana; Shibayama, Mitsuru; et al.

JOURNAL OF PHYSICS A-MATHEMATICAL AND THEORETICAL Volume: 51 Issue: 31 Article Number: 315101 Published: AUG 3 2018

ejournals@cambridge - find full tex Full Text from Publisher

View Abstract

Times Cited: 4
(from Web of Science Core Collection)

Usage Count

☐ ARTICLE (8)

Organizations-Enhanced

☐ UNIVERSITY OF BELGRADE (5)
☐ OSAKA CITY UNIVERSITY (2)
☐ SHANGHAI JIAO TONG UNIVERSITY (2)
☐ ETH ZURICH (1)
☐ IMPERIAL COLLEGE LONDON (1)

more options / values...

Refine

Funding Agencies

Authors

Source Titles

View all options

For advanced refine options, use

Analyze Results

☐ 6. **More than six hundred new families of Newtonian periodic planar collisionless three-body orbits**

By: Li, XiaoMing; Liao, ShiJun

SCIENCE CHINA-PHYSICS MECHANICS & ASTRONOMY **Volume:** 60 **Issue:** 12 **Article Number:** 129511 **Published:** DEC 2017

ejournals@cambridge - find full tex | Full Text from Publisher

View Abstract

Times Cited: 23
(from Web of Science Core Collection)

Usage Count

☐ 7. **Periodic three-body orbits with vanishing angular momentum in the Jacobi-Poincare 'strong' potential**

By: Dmitrasinovic, V.; Petrovic, Luka V.; Suvakov, Milovan

JOURNAL OF PHYSICS A-MATHEMATICAL AND THEORETICAL **Volume:** 50 **Issue:** 43 **Article Number:** 435102 **Published:** OCT 27 2017

ejournals@cambridge - find full tex | Full Text from Publisher

View Abstract

Times Cited: 4
(from Web of Science Core Collection)

Usage Count

☐ 8. **Variational principle for the determination of unstable periodic orbits and instanton trajectories at saddle points**

By: Junginger, Andrej; Main, Joerg; Wunner, Guenter; et al.

PHYSICAL REVIEW A **Volume:** 95 **Issue:** 3 **Article Number:** 032130 **Published:** MAR 29 2017

ejournals@cambridge - find full tex | Free Full Text from Publisher

View Abstract

Times Cited: 4
(from Web of Science Core Collection)

Usage Count

☐ Select Page | Print | More | Add to Marked List

Sort by: Date | Times Cited | Usage Count | More

1 of 1

Show: 50 per page

8 records matched your query of the 80,257,729 in the data limits you selected.
Key: = Structure available.

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Web of Science

Clarivate Analytics

Citing Articles: 5

(from Web of Science Core Collection)

For: Numerical solution of the quantum Lenard-Balescu equation for a non-degenerate one-component plasma
...More

Times Cited Counts

5 in All Databases
5 in Web of Science Core Collection
0 in Arabic Citation Index
0 in BIOSIS Citation Index
0 in Chinese Science Citation Database
0 data sets in Data Citation Index
0 publication in Data Citation Index
0 in Russian Science Citation Index
0 in SciELO Citation Index

View Additional Times Cited Counts

Sort by: **Date** Times Cited Usage Count More 1 of 1

☐ Select Page Print More Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **Quantum Fokker-Planck modeling of degenerate electrons**
By: Le, Hai P.
JOURNAL OF COMPUTATIONAL PHYSICS Volume: 434 Article Number: 110230 Published: JUN 1 2021
ejournals@cambridge - find full tex View Abstract
Times Cited: 0
(from Web of Science Core Collection)
Usage Count

☐ 2. **Adaptive spectral solution method for the Landau and Lenard-Balescu equations**
By: Scullard, Christian R.; Hickok, Abigail; Sotiris, Justyna O.; et al.
JOURNAL OF COMPUTATIONAL PHYSICS Volume: 402 Article Number: 109110 Published: FEB 1 2020
ejournals@cambridge - find full tex Full Text from Publisher
View Abstract
Times Cited: 1
(from Web of Science Core Collection)
Usage Count

☐ 3. **Analytic expressions for electron-ion temperature equilibration rates from the Lenard-Balescu equation**
By: Scullard, Christian R.; Serna, Susana; Benedict, Lorin X.; et al.
PHYSICAL REVIEW E Volume: 97 Issue: 1 Article Number: 013205 Published: JAN 31 2018
ejournals@cambridge - find full tex Free Full Text from Publisher
View Abstract
Times Cited: 4
(from Web of Science Core Collection)
Usage Count

☐ 4. **A Conservative, Entropic Multispecies BGK Model**
By: Haack, Jeffrey R.; Hauck, Cory D.; Murillo, Michael S.
JOURNAL OF STATISTICAL PHYSICS Volume: 168 Issue: 4 Pages: 826-856 Published: AUG 2017
ejournals@cambridge - find full tex Full Text from Publisher
View Abstract
Times Cited: 15
(from Web of Science Core Collection)
Usage Count

☐ 5. **Molecular dynamics studies of electron-ion temperature equilibration in hydrogen plasmas within the coupled-mode regime**
By: Benedict, Lorin X.; Surh, Michael P.; Stanton, Liam G.; et al.
PHYSICAL REVIEW E Volume: 95 Issue: 4 Article Number: 043202 Published: APR 10 2017
ejournals@cambridge - find full tex Free Full Text from Publisher
View Abstract
Times Cited: 18
(from Web of Science Core Collection)
Usage Count

☐ Select Page Print More Add to Marked List

Refine Results

Search within results for...

Filter results by:

☐ Open Access (2)

Refine

Publication Years

☐ 2021 (1)
☐ 2020 (1)
☐ 2018 (1)
☐ 2017 (2)

more options / values...

Refine

Web of Science Categories

☐ PHYSICS MATHEMATICAL (5)
☐ COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS (2)
☐ PHYSICS FLUIDS PLASMAS (2)

more options / values...

Refine

Document Types

Refine

☐ ARTICLE (5)

Organizations-Enhanced

☐ UNITED STATES DEPARTMENT OF ENERGY DOE (5)
☐ LAWRENCE LIVERMORE NATIONAL LABORATORY (4)
☐ LOS ALAMOS NATIONAL LABORATORY (2)
☐ MICHIGAN STATE UNIVERSITY (2)
☐ AUTONOMOUS UNIVERSITY OF BARCELONA (1)

more options / values...

Refine

Funding Agencies

Authors

Source Titles

View all options

For advanced refine options, use

Analyze Results

Sort by: Date Times Cited Usage Count More

1 of 1

Show: 50 per page

5 records matched your query of the 80,257,729 in the data limits you selected.
Key: = Structure available.

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Web of Science InCites Journal Citation Reports Essential Science Indicators EndNote Publons Kopernio Master Journal List Sign In Help English

Web of Science

Clarivate Analytics

Search Tools Searches and alerts Search History Marked List

Citing Articles: 3

(from Web of Science Core Collection)

For: A three-dimensional view of Gomez's hamburger **...More**

Times Cited Counts

- 3 in All Databases
- 3 in Web of Science Core Collection
- 0 in Arabic Citation Index
- 0 in BIOSIS Citation Index
- 0 in Chinese Science Citation Database
- 0 data sets in Data Citation Index
- 0 publication in Data Citation Index
- 0 in Russian Science Citation Index
- 0 in SciELO Citation Index

View Additional Times Cited Counts

Refine Results

Search within results for...

Filter results by:

☐ Open Access (1)

Refine

Publication Years

☐ 2021 (2)
☐ 2020 (1)

more options / values...

Refine

Web of Science Categories

☐ ASTRONOMY ASTROPHYSICS (3)

Refine

Document Types

☐ ARTICLE (3)

Refine

Organizations-Enhanced

☐ CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS (3)
☐ COMMUNAUTE UNIVERSITE GRENOBLE ALPES (3)

Refine

Sort by: **Date** Times Cited Usage Count More 1 of 1

☐ Select Page Print More Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **The Anatomy of an Unusual Edge-on Protoplanetary Disk. II. Gas Temperature and a Warm Outer Region**

By: Flores, C.; Duchene, G.; Wolff, S.; et al.

ASTRONOMICAL JOURNAL **Volume:** 161 **Issue:** 5 **Article Number:** 239 **Published:** MAY 2021

ejournals@cambridge - find full tex View Abstract

Times Cited: 1
(from Web of Science Core Collection)

Usage Count

☐ 2. **ALMA chemical survey of disk-outflow sources in Taurus (ALMA-DOT): V. Sample, overview, and demography of disk molecular emission**

By: Garufi, A.; Podio, L.; Codella, C.; et al.

ASTRONOMY & ASTROPHYSICS **Volume:** 645 **Article Number:** A145 **Published:** FEB 1 2021

ejournals@cambridge - find full tex View Abstract

Times Cited: 0
(from Web of Science Core Collection)

Usage Count

☐ 3. **ALMA chemical survey of disk-outflow sources in Taurus (ALMA-DOT) II. Vertical stratification of CO, CS, CN, H2CO, and CH3OH in a Class I disk**

By: Podio, L.; Garufi, A.; Codella, C.; et al.

ASTRONOMY & ASTROPHYSICS **Volume:** 642 **Article Number:** L7 **Published:** OCT 9 2020

ejournals@cambridge - find full tex Free Full Text from Publisher

View Abstract

Times Cited: 7
(from Web of Science Core Collection)

Usage Count

☐ Select Page Print More Add to Marked List

Sort by: **Date** Times Cited Usage Count More 1 of 1

Show: 50 per page

3 records matched your query of the 80,257,729 in the data limits you selected.
Key: = Structure available.

☐ INSTITUT DE PLANETOLOGIE ET D ASTROPHYSIQUE DE GRENOBLE IPAG (3)

☐ UNIVERSITE GRENOBLE ALPES UGA (3)

☐ EUROPEAN SOUTHERN OBSERVATORY (2)

more options / values...

Funding Agencies ▼

Authors ▼

Source Titles ▼

View all options

For advanced refine options, use

Analyze Results

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Web of Science InCites Journal Citation Reports Essential Science Indicators EndNote Publons Kopernio Master Journal List Sign In Help English

Web of Science

Clarivate Analytics

Search Tools Searches and alerts Search History Marked List

Citing Articles: 2

(from Web of Science Core Collection)

For: A guide to hunting periodic three-body orbits with non-vanishing angular momentum **...More**

Times Cited Counts

3 in All Databases
2 in Web of Science Core Collection
0 in Arabic Citation Index
0 in BIOSIS Citation Index
1 in Chinese Science Citation Database
0 data sets in Data Citation Index
0 publication in Data Citation Index
0 in Russian Science Citation Index
0 in SciELO Citation Index

View Additional Times Cited Counts

Sort by: **Date** Times Cited Usage Count More 1 of 1

☐ Select Page Print More Add to Marked List

Analyze Results
Create Citation Report

☐ 1. **One family of 13315 stable periodic orbits of non-hierarchical unequal-mass triple systems**

By: Li, XiaoMing; Li, XiaoChen; Liao, ShiJun

SCIENCE CHINA-PHYSICS MECHANICS & ASTRONOMY **Volume:** 64 **Issue:** 1 **Article Number:** 219511 **Published:** JAN 2021

ejournals@cambridge - find full tex View Abstract

Times Cited: 0
(from Web of Science Core Collection)

Usage Count

☐ 2. **Approximate Analytical Periodic Solutions to the Restricted Three-Body Problem with Perturbation, Oblateness, Radiation and Varying Mass**

By: Gao, Fabao; Wang, Yongqing

UNIVERSE **Volume:** 6 **Issue:** 8 **Article Number:** 110 **Published:** AUG 2020

ejournals@cambridge - find full tex Free Full Text from Publisher

View Abstract

Times Cited: 1
(from Web of Science Core Collection)

Usage Count

☐ Select Page Print More Add to Marked List

Sort by: **Date** Times Cited Usage Count More 1 of 1

Show: 50 per page

2 records matched your query of the 80,257,729 in the data limits you selected.
Key: = Structure available.

Refine Results

Search within results for...

Filter results by:

☐ Open Access (1)

Refine

Publication Years

☐ 2021 (1)
☐ 2020 (1)

more options / values...

Refine

Web of Science Categories

☐ ASTRONOMY ASTROPHYSICS (1)
☐ PHYSICS MULTIDISCIPLINARY (1)
☐ PHYSICS PARTICLES FIELDS (1)

more options / values...

Refine

Document Types

☐ ARTICLE (2)

Refine

Organizations-Enhanced

Refine

☐ AUTONOMOUS UNIVERSITY OF BARCELONA (1)
☐ JINAN UNIVERSITY (1)
☐ MASSACHUSETTS INSTITUTE OF TECHNOLOGY MIT (1)
☐ SHANGHAI JIAO TONG UNIVERSITY (1)
☐ SOUTH CHINA UNIVERSITY OF TECHNOLOGY (1)

more options / values...

Funding Agencies ▼

Authors ▼

Source Titles ▼

View all options

For advanced refine options, use

Analyze Results

Clarivate
Accelerating innovation

© 2021 Clarivate Copyright notice Terms of use Privacy statement Cookie policy

Sign up for the Web of Science newsletter Follow us

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **504,** 280–299 (2021)
Advance Access publication 2021 April 1

doi:10.1093/mnras/stab920

MRI-active inner regions of protoplanetary discs. I. A detailed model of disc structure

Marija R. Jankovic[1]★ James E. Owen[2] Subhanjoy Mohanty[2] and Jonathan C. Tan[3,4]

[1] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[2] *Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*
[3] *Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA*
[4] *Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg, Sweden*

Accepted 2021 March 29. Received 2021 March 29; in original form 2020 September 11

ABSTRACT

Short-period super-Earth-sized planets are common. Explaining how they form near their present orbits requires understanding the structure of the inner regions of protoplanetary discs. Previous studies have argued that the hot inner protoplanetary disc is unstable to the magnetorotational instability (MRI) due to thermal ionization of potassium, and that a local gas pressure maximum forms at the outer edge of this MRI-active zone. Here we present a steady-state model for inner discs accreting viscously, primarily due to the MRI. The structure and MRI-viscosity of the inner disc are fully coupled in our model; moreover, we account for many processes omitted in previous such models, including disc heating by both accretion and stellar irradiation, vertical energy transport, realistic dust opacities, dust effects on disc ionization, and non-thermal sources of ionization. For a disc around a solar-mass star with a standard gas accretion rate ($\dot{M} \sim 10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$) and small dust grains, we find that the inner disc is optically thick, and the accretion heat is primarily released near the mid-plane. As a result, both the disc mid-plane temperature and the location of the pressure maximum are only marginally affected by stellar irradiation, and the inner disc is also convectively unstable. As previously suggested, the inner disc is primarily ionized through thermionic and potassium ion emission from dust grains, which, at high temperatures, counteract adsorption of free charges on to grains. Our results show that the location of the pressure maximum is determined by the threshold temperature above which thermionic and ion emission become efficient.

Key words: planets and satellites: formation – protoplanetary discs .

1 INTRODUCTION

Close-in super-Earths – planets with radii 1–4 $\mathrm{R}_\oplus$ and orbital periods shorter than ∼100 d – are common around solar-type and lower mass stars (Dressing & Charbonneau 2013, 2015; Fressin et al. 2013; Mulders et al. 2018; Hsu et al. 2019; Zink, Christiansen & Hansen 2019). Yet how these planets form is still an open problem. One theory posits that they are born further away from the star, and subsequently migrate through the protoplanetary disc to their present orbits (Terquem & Papaloizou 2007; Ogihara & Ida 2009; McNeil & Nelson 2010; Cossou et al. 2014; Izidoro et al. 2017, 2019; Bitsch, Raymond & Izidoro 2019). However, this hypothesis predicts that the planets should be water-rich, in contrast to the water-poor composition inferred from their observed radius distribution combined with atmospheric evolution models (Owen & Wu 2017; Van Eylen et al. 2018; Wu 2019).

An alternative proposal is that these super-Earths form at or near their present orbits, out of solid material which migrates to the inner disc prior to planet formation. These solids are expected to arrive in the form of pebbles, which radially drift inwards from the outer disc due to gas drag (Boley & Ford 2013; Hansen & Murray 2013; Chatterjee & Tan 2014; Hu et al. 2018; Jankovic, Owen & Mohanty 2019). However, these pebbles need to be trapped, i.e. their radial drift must be halted, in order for them to coalesce into planets instead of drifting into the star. Such trapping may occur as follows.

In the inner protoplanetary disc, at the short orbital periods at which close-in super-Earths are observed, the gaseous disc is thought to accrete via the magnetorotational instability (MRI; Balbus & Hawley 1991). The MRI leads to turbulence in the disc, which drives viscous accretion, but its magnitude is sensitive to the disc's ionization state. The MRI is expected to be efficient in the hot, innermost parts of the disc (where it can be activated by thermal ionization of trace alkali elements, at temperatures $\gtrsim$1000 K), and largely suppressed in cold regions further away from the star (the latter being known as the MRI dead zone; Gammie 1996). At the transition between the two regions, a local gas pressure maximum forms when the disc is in steady state, and this pressure maximum may trap the pebbles that drift in from the outer disc (Chatterjee & Tan 2014). Concurrently, the MRI-driven turbulence in the inner disc (interior to the pressure maximum) limits the grain size thereby inducing high collisional velocities between grains, causing them to fragment. A decrease in the size of grains reduces their radial drift, providing an alternative way to accumulate the grains arriving from the outer disc (Jankovic et al. 2019). In general, therefore, the structure of the inner disc, undergoing MRI-induced accretion, is likely to play a key role in the formation of close-in super-Earths.

★ E-mail: mj577@cam.ac.uk

© 2021 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

In previous work (Mohanty et al. 2018), we presented a model of the inner disc in which the disc structure and accretion due to the MRI are treated self-consistently. We obtained an inner disc structure in line with the expectations sketched above, and inferred a location for the pressure maximum consistent with the orbital distances of close-in super-Earths. However, this model includes a number of simplifying assumptions about the disc's physical and chemical structure. First, the disc is considered to be vertically isothermal, and secondly, heating by stellar irradiation is neglected. In reality, the disc temperature will vary vertically, with a profile that is particularly non-trivial when both accretion heating and stellar irradiation are accounted for (e.g. D'Alessio et al. 1998). Thirdly, the model assumes a constant dust opacity, when in fact the latter depends on the temperature and on grain properties and abundance.

Fourthly, Mohanty et al. (2018) assumed that the only source of ionization in the disc is thermal (collisional) ionization of potassium, and only considered gas-phase interactions. Potassium is indeed a good representative of thermally ionized species in the inner disc due to its low ionization potential and high abundance (Desch & Turner 2015). However, ions and free electrons are also adsorbed on to the surfaces of dust grains in the disc, where they quickly recombine. Dust can thus reduce the disc ionization level and suppress the MRI (Sano et al. 2000; Ilgner & Nelson 2006; Wardle 2007; Salmeron & Wardle 2008; Bai & Goodman 2009; Mohanty, Ercolano & Turner 2013). Conversely, hot grains ($\gtrsim$500 K) can also emit electrons and ions into the gas phase (for electrons, this process is known as thermionic emission). Such temperatures are easily attainable in the inner disc, and thermionic and ion emission have been hypothesized to be important sources of ionization there (Desch & Turner 2015). Neither of these grain effects are treated by Mohanty et al. (2018). Finally, while Mohanty et al. (2018) showed a posteriori that, in the MRI-accreting inner disc, X-ray ionization of molecular hydrogen may be competitive with thermal ionization of potassium in supplying free electrons (due to the low gas surface density in this region), they did not actually include X-ray ionization in their model.

In this paper, we present a new model of an MRI-accreting, steady-state inner disc that addresses each of the shortcomings of the Mohanty et al. (2018) model listed above. The vertical structure in our model is calculated self-consistently from viscous dissipation (due to the MRI-induced viscosity), stellar irradiation, and radiative and convective cooling, with realistic opacities due to dust grains. Ionization in the disc is determined by thermal ionization of potassium, thermionic and ion emission from dust grains, and ionization of molecular hydrogen by stellar X-rays, cosmic rays, and radionuclides. In Section 2, we detail all the components of our model and the methods used to find self-consistent steady-state solutions for the disc structure. We present our results in Section 3, discuss the relative importance of the various physical and chemical processes in Section 4, and summarize our findings in Section 5.

The aim of this paper is two-fold: to present the methodology of our calculations, and to discuss the detailed physics of the inner disc in the context of fiducial disc and stellar parameters. In a companion paper (Paper II: Jankovic et al., in preparation), we investigate how the inner disc structure varies as a function of these parameters, and discuss the implications for the formation of super-Earths in the inner disc.

2 METHODS

We consider a disc that is viscously accreting and in steady-state, i.e. has a constant mass accretion rate. Our model of the disc structure is described in Section 2.1. The disc structure depends on the disc's radiative properties, i.e. opacities, and on the viscosity. Our calculation of opacities is summarized in Section 2.2, and the prescription for MRI-driven viscosity given in Section 2.3. The latter viscosity is a function of the disc's ionization state, calculated using a chemical network described in Section 2.4. The disc structure, opacities, ionization, and viscosity are calculated self-consistently at every point in the disc, using numerical methods supplied in Section 2.5.

The key parameters of our model are the steady-state gas accretion rate $\dot{M}$ through the disc, stellar mass M_*, stellar radius R_*, stellar effective temperature T_*, and value of the viscosity in the absence of the MRI (the dead-zone viscosity). Additionally, the disc opacities and ionization state, and thus the disc structure, depend on the properties of the dust; most importantly, on the dust-to-gas ratio $f_{\rm dg}$ and the maximum dust grain size $a_{\rm max}$.

We assume that the disc only accretes viscously due to the MRI. Disc accretion may additionally be partially driven by magnetic winds, which could also affect the inner disc structure (e.g. Suzuki et al. 2016). Therefore, inclusion of wind-driven accretion is an important issue; however, it is one which we do not tackle here.

2.1 Disc model

Our disc model largely follows the work of D'Alessio et al. (1998, 1999). We consider a thin, axisymmetric, Keplerian, steady-state disc that is viscously accreting. We assume that the disc is in vertical hydrostatic equilibrium, heated by viscous dissipation and stellar irradiation, and that energy is transported by radiation and convection. Since the disc is vertically thin, we neglect energy transport in the radial direction. Furthermore, at a given disc radius, our viscosity depends only on local conditions and the vertical mass column (see Section 2.3). As such, the disc structures at different radii are only coupled by the stellar irradiation, as it penetrates the disc along the line of sight to the central star.

2.1.1 *Hydrostatic equilibrium*

In a thin Keplerian disc in vertical hydrostatic equilibrium, the gas pressure profile at any given radius follows from

$$\frac{{\rm d}P}{{\rm d}z} = -\rho\Omega^2 z, \tag{1}$$

where P is the gas pressure, ρ the gas volume density, Ω the Keplerian angular velocity, and z the height above the disc mid-plane. We adopt the ideal gas law.

2.1.2 *Viscous heating and stellar irradiation*

The disc is heated by the viscosity that drives the accretion. The viscosity ν is parametrized by $\nu = \alpha c_{\rm s}^2/\Omega$, where α is the viscosity parameter and $c_{\rm s}$ the isothermal sound speed (Shakura & Sunyaev 1973). The local viscous dissipation rate at any location in the disc is given by

$$\Gamma_{\rm acc} = \frac{9}{4}\alpha P\Omega. \tag{2}$$

The flux generated by viscous dissipation that is radiated through one side of the disc at any disc radius r is

$$F_{\rm acc} = \frac{3}{8\pi}\dot{M} f_r \Omega^2, \tag{3}$$

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

where $f_r = 1 - \sqrt{R_{\rm in}/r}$ comes from the thin boundary-layer condition at the inner edge of the disc, $R_{\rm in} = R_*$ is the radius of the inner edge of the disc, and we are assuming a zero-torque inner boundary condition (e.g. Frank, King & Raine 2002).

We also consider heating due to stellar irradiation. Stellar flux propagates spherically outwards from the star, and the resultant heating at any disc location depends on the attenuation of this flux along the line of sight to the star. Accounting for this, however, while simultaneously neglecting scattering of starlight and radial energy transport within the disc (as we do in our 1 + 1D model here), leads to multiple equilibrium solutions or none at all (e.g. Chiang et al. 2001). Such behaviour does not appear in 2D disc models (e.g. Dullemond 2002), which are nevertheless too complex for our purposes here. Instead, we treat the irradiation heating at each disc radius in isolation, by considering heating only due to the stellar flux that impinges on the disc surface at that radius (at some grazing angle ϕ calculated self-consistently; see further below), and then propagates vertically towards the mid-plane (Calvet et al. 1992; Chiang & Goldreich 1997). In this framework, the attenuation, i.e. the optical depth to the stellar flux, is approximated as $\tau_{\rm irr} \approx \tau_{\rm irr,\,z}/\mu$, where $\tau_{\rm irr,\,z}$ is the optical depth in the vertical direction, and $\mu \equiv \sin\phi$. Local heating due to stellar irradiation is then given by

$$\Gamma_{\rm irr} = \kappa_{\rm P}^* \rho \frac{F_{\rm irr}}{\mu} e^{-\tau_{\rm irr,z}/\mu}, \tag{4}$$

where $\kappa_{\rm P}^*$ is the disc Planck opacity to stellar irradiation (see Section 2.2) and $F_{\rm irr}$ is the total incident stellar flux at the specified disc radius. In an optically thick disc, the latter is given by

$$F_{\rm irr} = \sigma_{SB} T_*^4 \left(\frac{R_*}{s}\right)^2 \mu, \tag{5}$$

where s is the distance to the star (spherical radius) from the disc surface at the specified radius.[1]

Finally, the grazing angle ϕ is given by

$$\phi = \sin^{-1}\frac{4}{3\pi}\frac{R_*}{r} + \tan^{-1}\frac{{\rm dlog}z_{\rm irr}}{{\rm dlog}r}\frac{z_{\rm irr}}{r} - \tan^{-1}\frac{z_{\rm irr}}{r}. \tag{6}$$

Here the first term is the value of ϕ for a flat disc and comes from the finite size of the stellar disc, and the other two terms are due to disc flaring. $z_{\rm irr}(r)$ is the height above the disc mid-plane at which the stellar flux is absorbed; we take $z_{\rm irr}(r)$ to be the height where the optical depth to the stellar irradiation is $\tau_{\rm irr} = 2/3$. Specifically, in calculating the height $z_{\rm irr}$, the optical depth $\tau_{\rm irr}$ is obtained by integrating along the spherical radius s to the star, as opposed to the approximation used in the local heating term. We determine $z_{\rm irr}$ and ϕ self-consistently with the disc structure following the procedure by D'Alessio et al. (1999), as outlined in Section 2.5.

2.1.3 *Radiative energy transport*

In a thin disc, the optical depth to the disc's own radiation is much smaller in the vertical direction than in the radial. Thus we expect radiative energy transport to be primarily vertical, and that is the only direction we consider.

The frequency-integrated moments of the radiative transfer equation in the Eddington approximation (i.e. assuming that the radiation is isotropic, as is valid in the optically thick regime) and the equation for energy balance are then

$$\frac{{\rm d}F}{{\rm d}z} = \Gamma_{\rm acc} + \Gamma_{\rm irr}, \tag{7}$$

$$\frac{{\rm d}J}{{\rm d}z} = -\tfrac{3\rho\kappa_{\rm R}}{4\pi} F, \tag{8}$$

$$4\rho\kappa_{\rm P}(\sigma_{\rm SB}T^4 - \pi J) = \Gamma_{\rm acc} + \Gamma_{\rm irr}, \tag{9}$$

where F and J are the radiative flux and mean intensity, respectively. We have also assumed here that the J and F weighted opacities can be approximated by the Planck mean opacity $\kappa_{\rm P}$ and the Rosseland mean opacity $\kappa_{\rm R,}$ respectively (following e.g. Hubeny 1990, see Section 2.2).

Together with the ideal gas equation of state, equations (1), (7)–(9) form a closed set in P, F, J, and T. Together with appropriate boundary conditions, they determine the disc vertical structure. One boundary condition is imposed at the disc mid-plane, where the flux $F(0) = 0$ by symmetry. The remaining boundary conditions are supplied at the disc surface, at height $z_{\rm surf}$ above the mid-plane. The boundary condition for the flux F that is obtained by integrating equation (7) from $z = 0$ to $z_{\rm surf}$, from which it follows that $F(z_{\rm surf}) = F_{\rm acc} + F_{\rm irr}$. The boundary condition for the mean intensity J is given by $J(z_{\rm surf}) = \frac{1}{2\pi}F(z_{\rm surf})$. Finally, we assume that the gas pressure at the top of the disc has a small constant value, $P(z_{\rm surf}) = 10^{-10}$ dyn cm^{-2}, which is another boundary condition. The precise value of $P(z_{\rm surf})$ is arbitrary, but does not affect our results as long as it is sufficiently small. Overall, then, we have four boundary conditions on three differential equations. Note that the temperature $T(z_{\rm surf})$ at the disc surface follows from the algebraic equation (9), the boundary conditions on J and P, and the ideal gas law.

2.1.4 *Energy transport by convection*

If radiative energy transport yields a thermal structure such that the gradient $\nabla = \frac{{\rm dln}T}{{\rm dln}P}$ is greater than the adiabatic gradient $\nabla_{\rm ad} = (\gamma - 1)/\gamma$, then the gas is unstable to convection. In disc regions where this is the case, we assume that energy transport by convection is efficient and the gas is vertically isentropic, so that $\nabla = \nabla_{\rm ad}$ at such locations (e.g. Shu 1992; Garaud & Lin 2007; Rafikov 2007). We adopt $\gamma = 1.4$, valid for an H_2 dominated disc. Equations (8) and (9) are then replaced by

$$\frac{{\rm d}T}{{\rm d}z} = -\nabla_{\rm ad}\tfrac{T}{P}\rho(T,\,P)\Omega^2 z = -\nabla_{\rm ad}\tfrac{\mu m_{\rm H}}{k_{\rm B}}\Omega^2 z. \tag{10}$$

2.2 Opacities

Radiative transport is controlled by the Rosseland-mean opacity $\kappa_{\rm R}$ in optically thick regions (equation 8), and by the Planck-mean opacity $\kappa_{\rm P}$ in optically thin regions (equation 9). Additionally, the absorption coefficient for the stellar flux is a Planck-mean opacity $\kappa_{\rm P}^*$ at the stellar effective temperature. We assume that the only source of these opacities are dust grains. Gas opacities are important in the very innermost regions of discs, where most dust species have sublimated. However, these regions are not of particular significance for the early stages of planet formation that we are interested in, as dust is required to form solid planet cores. Beyond the silicate

---

[1] When calculating the factor f_r in equation (3) we have assumed that the inner disc edge is at the stellar radius, which would imply that the disc surface can only see a half of the stellar disc. If the disc surface can only see a half of the stellar disc, the expression for the total absorbed stellar flux, equation (5), should include an additional factor of 1/2 (e.g. Chiang & Goldreich 1997). However, in a more realistic case of the disc being truncated at several stellar radii, the entire stellar disc should be visible (see e.g. Estrada, Cuzzi & Morgan 2016), and so we do not include this additional factor.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

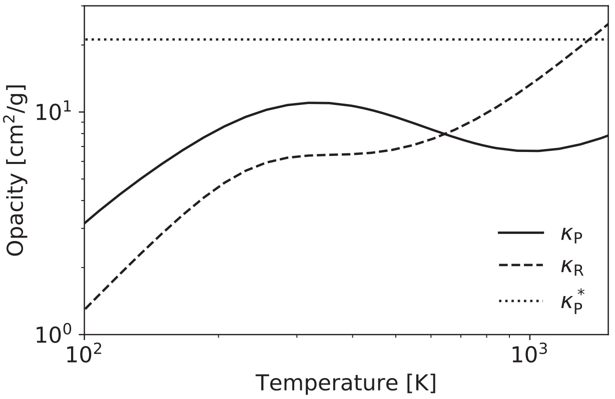


Figure 1. Planck-mean opacity $\kappa_{\rm P}$, Rosseland-mean opacity $\kappa_{\rm R}$, and Planck-mean opacity at the stellar effective temperature $\kappa_{\rm P}^*$, as functions of disc temperature, assuming dust-to-gas ratio $f_{\rm dg} = 10^{-2}$ and a maximum dust grain size $a_{\rm max} = 1$ µm.

sublimation line, gas opacities may still be important in hot, optically thin regions (Malygin et al. 2014). However, as we find, the inner disc is significantly optically thick and including gas opacities would only alter the structure of the hot disc atmosphere. Therefore, we completely ignore the contribution of gas to the opacities.

Furthermore, we assume that the only dust species present are silicate grains. Other species that can be comparable in abundance to silicates are water ice and carbonaceous grains (e.g. organics; Pollack et al. 1994). However, due to their low sublimation temperatures, neither water ice nor carbonaceous grains are expected in the hot MRI-active regions, and we generally limit our calculations to the inner 1 AU of the disc.

To calculate the opacities, we adopt the optical constants of 'astronomical silicates' from Draine (2003). We use the MIESCAT module of the python wrapper RADMC3DPY for RADMC3D (Dullemond et al. 2012) to obtain dust absorption and scattering coefficients as functions of radiation wavelength and grain size. For each grain size a, the coefficients are averaged within a size bin of width $\Delta \ln a = 0.02$. Next, we assume that the grain bulk density is $\rho_{\rm gr} = 3.3\,{\rm g\,cm^{-3}}$, and that the grain sizes are described by the MRN distribution, wherein the number density of grains in the size range $[a, a + {\rm d}a]$ is given by $n(a){\rm d}a \propto a^{-q}{\rm d}a$ (Mathis, Rumpl & Nordsieck 1977), with $q = 3.5$. We adopt a minimum grain size of $a_{\rm min} = 0.1$ µm, and a variable maximum grain size $a_{\rm max}$. The size-dependent absorption and scattering coefficients are then weighted by grain mass and averaged over the grain size distribution. Finally, the absorption coefficient is integrated over frequency to obtain the Planck-mean opacity $\kappa_{\rm P}(T)$, and the total extinction coefficient yields the Rosseland-mean opacity $\kappa_{\rm R}(T)$. Following D'Alessio et al. (1998), we calculate the mean absorption coefficient for the stellar flux as a frequency-integrated absorption coefficient weighted by the Planck function at the stellar effective temperature: $\kappa_{\rm P}^* \equiv \kappa_{\rm P}(T_*)$.

Fig. 1 shows the opacities per unit mass of gas, assuming a dust-to-gas ratio of $f_{\rm dg} = 10^{-2}$, a maximum grain size of $a_{\rm max} = 1$ µm, and $T_* = 4400$ K. The Planck-mean opacity at the stellar effective temperature $\kappa_{\rm P}^*$ is a constant, since the wavelength-dependent dust absorption coefficient does not depend on the local temperature. The Planck-mean opacity $\kappa_{\rm P}$ is in general expected to increase with increasing temperature. This is because the wavelength at which the Planck function peaks is inversely proportional to the temperature, and for grains smaller than the wavelength of peak emission (and small grains contribute to the opacities most) absorption is expected to increase with decreasing wavelength. However, due to the silicate absorption feature at 10 µm, $\kappa_{\rm P}$ decreases with temperature in the range ~ 500–1000 K.

2.3 Viscosity

Our model of the MRI-driven viscosity closely follows that of Mohanty et al. (2018) (see also Bai 2011). Here we only summarize the main points. A well-ionized circumstellar disc which follows the laws of ideal magnetohydrodynamics (MHD) is susceptible to the MRI (Balbus & Hawley 1991). The MRI leads to turbulence, producing an accretion stress and acting as a source of viscosity. The resulting Shakura–Sunyaev viscosity parameter α is

$$\alpha_{\rm AZ} = \frac{1}{3\beta}, \tag{11}$$

where the subscript 'AZ' denotes an MRI-active zone, $\beta \equiv P/P_B$ is the plasma parameter, and $P_B = B^2/8\pi$ is the magnetic field pressure (Sano et al. 2004).[2]

However, even in the inner regions of protoplanetary discs, non-ideal MHD effects can quench the MRI, leading to the so-called dead zones. The non-ideal effects considered here are Ohmic and ambipolar diffusion. Ohmic diffusion will not suppress the MRI if (Sano & Stone 2002)

$$\Lambda = \frac{v_{\mathcal{A}z}^2}{\eta_{\rm O}\Omega} > 1, \tag{12}$$

where Λ is the Ohmic Elsasser number, $v_{\mathcal{A}z} \equiv B_z/\sqrt{4\pi\rho}$ is the vertical component of the local Alfven velocity, and $\eta_{\rm O}$ is the Ohmic resistivity. Here we utilize the relationship between the strength of the vertical component of the magnetic field, B_z, and the strength of the r.m.s. field, B: $B_z^2 \sim B^2/25$ (Sano et al. 2004). We also assume that B is vertically constant. Our method of determining the value of B is described in Section 2.5.

Similarly, the ambipolar Elsasser number is defined by

$$Am = \frac{v_{\mathcal{A}}^2}{\eta_{\rm A}\Omega}, \tag{13}$$

where $\eta_{\rm A}$ is the ambipolar magnetic resistivity. However, in the strong-coupling limit, valid in protoplanetary discs, the MRI can be active even if $Am < 1$, as long as the magnetic field is sufficiently weak (Bai & Stone 2011). The criterion for active MRI in the presence of ambipolar diffusion is

$$\beta/\beta_{\rm min} > 1, \tag{14}$$

where the minimum value of β neccessary to sustain the MRI is a function of the ambipolar Elsasser number:

$$\beta_{\rm min}(Am) = \left[\left(\frac{50}{Am^{1.2}}\right)^2 + \left(\frac{8}{Am^{0.3}} + 1\right)^2\right]^{1/2}. \tag{15}$$

Thus, whether the MRI is active or not depends on the magnetic resistivities, $\eta_{\rm O}$ and $\eta_{\rm A}$, as well as on the magnetic field strength B. The magnetic resistivities, calculated following Wardle (1999), express the coupling between the gas and the magnetic field, which is principally determined by the degree of ionization of the gas.

If either of the two criteria given by equations (12) and (14) is not fulfilled, the MRI is not active. In such MRI-dead zones, we assume there is a small residual viscosity $\alpha_{\rm DZ}$, driven either

---

[2] The numerical factor is indeed 1/3 for the Shakura–Sunyaev α parameter in equation (11); see discussion in Appendix B of Mohanty et al. (2018).

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

by propagation of turbulence from the MRI-active zone, or by hydrodynamic instabilities (for more discussion, see Mohanty et al. 2018). In this work, we also impose a smooth transition between the active and the dead zones, necessary to ensure numerical stability in the integration of the equations of disc structure (such smoothing was not employed by Mohanty et al. 2018). To cover both non-ideal effects that lead to dead zones, we define $C \equiv \min(\Lambda, \beta/\beta_{\rm min})$. Then, at any location in the disc where $|C - 1| < 0.5$ (i.e. in the vicinity of the switch from active to dead), we adopt a smoothed α given by

$$\alpha = \alpha_{\rm DZ} + \frac{\alpha_{\rm AZ} - \alpha_{\rm DZ}}{1 + \exp\left(-\frac{C-1}{\Delta}\right)}, \tag{16}$$

where $\Delta = 10^{-2}$.

2.4 Ionization

2.4.1 Chemical network

We implement the same simple chemical network adopted by Desch & Turner (2015). This network tracks the number densities of five species: free electrons ($n_{\rm e}$), potassium ions ($n_{\rm K^+}$), neutral potassium atoms ($n_{\rm K^0}$), potassium atoms adsorbed (condensed) on to dust grains ($n_{\rm K,\,cond}$), and atomic ions ($n_{\rm i}$; i.e. ions of atomic species other than potassium; see below).

In the gas-phase, potassium atoms can be thermally ionized: collisions of neutral potassium atoms with H_2 molecules produce potassium ions and free electrons. The potassium ions and free electrons can also recombine in the gas phase, either radiatively or via three-body collisions with H_2 molecules (the latter process dominates at the high densities prevalent in the inner disc).

Furthermore, non-thermal sources (we consider stellar X-rays, cosmic rays and radionuclides; see Section 2.4.3) can ionize H_2 (Glassgold, Najita & Igea 1997; Ercolano & Glassgold 2013). The charge is quickly transferred from the ionized hydrogen to other abundant gas species through collisions, producing molecular and atomic ions (e.g. HCO^+, Mg^+). Notably, in application to the MRI, the exact composition of the gas that this leads to is unimportant in the presence of dust, and simple chemical networks reproduce the gas ionization levels well (Ilgner & Nelson 2006). Thus, it is assumed that the ionization of molecular hydrogen at a rate ζ directly produces atomic ions and free electrons at a volumetric rate $\zeta n_{\rm H_2}$. The atomic ion species in this chemical network may thus be understood as a representative of the various chemical species abundant in the gas-phase, whose mass is taken to be that of magnesium. It is assumed that the number density of molecular hydrogen is constant, which is valid for low ionization rates. Just as potassium, the atomic ions also recombine radiatively and in three-body recombinations.

Importantly, all gas-phase species (electrons, ions, and neutral atoms) collide with and are adsorbed on to dust grains at a rate

$$\mathcal{R}_{j,\,\rm coll} = n_j n_{\rm gr} \pi a_{\rm gr}^2 \left(\frac{8 k_{\rm B} T}{\pi m_j}\right)^{1/2} \tilde{J}_j S_j, \tag{17}$$

where n_j is the number density of the gas-phase species, $n_{\rm gr}$ is the number density of the grains, $a_{\rm gr}$ is the grain size, m_j is the gas-phase species mass, $\tilde{J}_j$ is the modification of the collisional cross-sections for charged species due to dust grain charge (Draine & Sutin 1987), and S_j is the sticking coefficient. It is assumed that all grains have the same charge; this is valid since the dispersion in the distribution of charge states is generally found to be small (Draine & Sutin 1987). It is further assumed that potassium ions rapidly recombine on the grain surface to form condensed potassium atoms. Thus the ions are effectively destroyed upon adsorption.

At high temperatures electrons on the dust grains have a finite probability of leaving the grain, producing the so-called thermionic emission. The emission depends on the energy required for the electron to escape the grain. For a neutral grain this is the work function W, a property of the material out of which the grains are made. The rate at which free electrons are produced through thermionic emission is

$$\mathcal{R}_{\rm therm} = n_{\rm gr} 4\pi a_{\rm gr}^2 \lambda_{\rm R} \frac{4\pi m_{\rm e} (k_{\rm B} T)^2}{h^3} \exp\left(-\frac{W_{\rm eff}}{k_{\rm B} T}\right), \tag{18}$$

where

$$W_{\rm eff} = W + \frac{Ze^2}{a_{\rm gr}} \tag{19}$$

is the effective work function due to grain charge Ze.

Potassium atoms will also evaporate from the grains only at high temperatures. The vaporization rate of condensed potassium atoms is given by

$$\mathcal{R}_{\rm K,evap} = n_{\rm K,cond} \nu \exp\left(-\frac{E_{\rm a}}{k_{\rm B} T}\right), \tag{20}$$

where ν is the vibration frequency of potassium atoms on the dust grain surface lattice, and $E_{\rm a} = 3.26$ eV is the binding energy, whose value is chosen to reproduce the condensation temperature of potassium (1006 K; Lodders 2003). These potassium atoms may be emitted into the gas phase as both neutral atoms and ions, contributing further to the gas' ionization state. The ratio of ions to neutrals among the emitted particles is given by

$$\frac{n_{\rm K}^+}{n_{\rm K^0}} = \frac{g_+}{g_0} \exp\left(+\frac{W_{\rm eff} - {\rm IP}}{k_{\rm B} T}\right), \tag{21}$$

where $\frac{g_+}{g_0}$ is the ratio of statistical weights of the ionized and neutral state of potassium, and IP the ionization potential of potassium. The fraction of all emitted particles that leave the grain as ions is then given by

$$f_+ = \frac{n_{\rm K}^+/n_{\rm K^0}}{1 + n_{\rm K}^+/n_{\rm K^0}}, \tag{22}$$

so the rate at which potassium ions evaporate from the grains is given by $\mathcal{R}_{\rm K,evap} f_+$.

It is important to note here that there are two sources of the neutral potassium atoms condensed on grains, whose evaporation is described above: first, gas-phase potassium ions that are adsorbed on to grain surfaces and recombine there into neutral atoms; and secondly, gas-phase neutral potassium atoms that are adsorbed directly on to the grains. The reionization and subsequent ejection from grains of the former simply returns ions to the gas-phase that were originally adsorbed from it, and thus clearly cannot increase the ion fraction in the gas-phase beyond what it would be in the absence of grains. The ionization on the grain surface, and ejection as ions, of particles that were originally neutral in the gas-phase, however, can increase the gas-phase ion fraction beyond what it would be without grains; it is this channel that makes ion emission from grains such a crucial effect.

Clearly, the contribution of the dust grains to gas ionization levels depends on the work function W of the grain material. We adopt the fiducial value of Desch & Turner (2015), $W = 5$ eV, and refer the reader to their work for a discussion of the experimental results supporting this choice. This value is close to the ionization potential of potassium, IP = 4.34 eV, indicating that thermionic emission is important for the production of free electrons in the same temperature range as thermal ionization of potassium. Importantly, for this given

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

value of the work function, grains become negatively charged at high temperatures as a large fraction of potassium evaporating from the grains is in ionized state. This results in a reduction of the effective work function $W_{\rm eff}$, since, for negatively charged grains thermionic emission is higher (equation 19).

In this work, we assume that the abundances of hydrogen and potassium atoms are $x_{\rm H} = 9.21 \times 10^{-1}$ and $x_{\rm K} = 9.87 \times 10^{-8}$, respectively, relative to the total number density of all atomic particles (Keith & Wardle 2014), and that the mean molecular weight is $\mu = 2.34 m_{\rm H}$.[3] The grain material density is $\rho_{\rm gr} = 3.3\,{\rm g\,cm^{-3}}$, the same as in our calculation of the dust opacities. The input for the chemical network are temperature T, pressure P, hydrogen ionization rate ζ, dust-to-gas ratio $f_{\rm dg}$, and dust grain size $a_{\rm gr}$. All other kinetic rates, parameters, and coefficients are the same as in Desch & Turner (2015).[4] We note that the chosen value of the sticking coefficient for electrons ($S_e = 0.6$) is compatible with the detailed calculation by Bai (2011), who consider work function values of 1 and 3 eV. His results suggest that for a work function of 5 eV, at 1000 K, S_e is indeed a few times 0.1 for neutral grains, and increases further for negatively charged grains. For the sticking coefficients for the ions we adopt $S_j = 1$, and similarly for the neutral atoms we adopt a sticking coefficient of $S_0 = 1$.

For a given dust-to-gas ratio $f_{\rm dg}$ and grain size $a_{\rm gr}$, we pre-calculate and tabulate the equilibrium number densities of electrons and ions, and the average grain charge, as functions of temperature T, pressure P, and hydrogen ionization rate ζ. We find the equilibrium solution following the same method as Desch & Turner (2015). Time derivatives of all number densities are set to zero, and so rate equations yield an algebraic system of equations. For a given average grain charge Z, this system of equations is solved iteratively to find number densities of all five species. The grain charge is then found by solving the equation of charge neutrality.

2.4.2 *Grain size distribution / effective dust-to-gas ratio*

The described chemical network incorporates only one grain size population. Ideally, we would consider a number of grain size populations, with the same size distribution used in our calculation of dust opacities. However, this would greatly enhance the computational complexity of the problem. At the same time, it is clear that dust grains of different size contribute differently to the equilibrium ionization levels. To the lowest order of approximation, all dust-related reaction rates are regulated by the total grain surface area. Thus, it can be expected that the ionization levels are most sensitive to the smallest grains. Bai & Goodman (2009) considered the effects of dust on the ionization levels in the cold regions of protoplanetary discs, in application to the onset of the MRI due to non-thermal sources of ionization. They considered chemical networks with two grain size populations and found that the grain populations behave independently, as the charge transfer between the grains is negligible. They further found that the ionization levels are largely controlled by a quantity $f_{\rm dg}/a_{\rm gr}^p$, where the exponent p varies between $p = 1$ (i.e. the total grain surface area) and $p = 2$.

We repeat a similar exercise for the above chemical network, suited to the hot inner regions of protoplanetary discs. For a set of values of the exponent $p = 1$, 1.25, and 1.5, we vary the grain size $a_{\rm gr}$ and dust-to-gas mass ratio $f_{\rm dg}$ while keeping $f_{\rm dg}/a_{\rm gr}^p$ constant. We calculate the ionization levels as a function of temperature, and for different sets of pressure and hydrogen ionization rates, so as to probe different conditions in different regions of the inner disc. The results are shown in Fig. 2. The grain surface area ($p = 1$) controls the equilibrium ionization levels when the hydrogen ionization dominates, but does not determine the temperature at which the ionization levels rise due to thermionic and ion emission. On the other hand, for $p = 1.5$, this temperature depends very weakly on the grain size. That is, regardless of the actual dust grain size, a quantity $f_{\rm dg}/a_{\rm gr}^{1.5}$ regulates the temperature at which dense interior of the disc becomes ionized.

Therefore, for the chemical network calculations, we use a single grain size of $a_{\rm gr} = 10^{-5}$ cm, but employ an effective dust-to-gas ratio $f_{\rm dg,\,eff}$ that satisfies $f_{\rm dg,eff} a_{\rm gr}^{-p} = \int_{a_{\rm min}}^{a_{\rm max}} {\rm d}n(a) m(a)/\rho_{\rm g} a^{-p}$, where $n(a)$ is the same grain size distribution used to calculate dust opacities. Since we find that thermionic and ion emission are more important than the non-thermal ionization of hydrogen in the inner disc, we use $p = 1.5$. From right-hand panel of Fig. 2, it appears that this choice will lead to a large error in the gas ionization state we derive in the low-density non-thermally ionized disc regions, due to the large variation in such regions in the ionization state produced by different grain sizes . However, in reality the error will be much smaller than implied by Fig. 2, since the majority of grains are skewed towards small sizes in a realistic grain size distribution.

2.4.3 *Hydrogen ionization rate*

In our calculation of the MRI-driven viscosity, we consider molecular hydrogen ionization rate due to radionuclides, cosmic rays, and stellar X-rays. The ionization rate of molecular hydrogen due to short-lived and long-lived radionuclides is

$$\zeta_{\rm R} = 7.6 \times 10^{-19} \quad {\rm s^{-1}}, \tag{23}$$

which predominantly comes from decay of ^{26}Al (Umebayashi & Nakano 2009). The ionization rate of molecular hydrogen due to interstellar cosmic rays is (Umebayashi & Nakano 2009)

$$\zeta_{\rm CR}(z) = \frac{\zeta_{\rm CR,ISM}}{2} e^{-\frac{\Sigma(z)}{\lambda_{\rm CR}}} \left(1 + \left(\frac{\Sigma(z)}{\lambda_{\rm CR}}\right)^{\frac{3}{4}}\right)^{-\frac{4}{3}}, \tag{24}$$

where $\zeta_{\rm CR,\,ISM} = 10^{-17}\,{\rm s^{-1}}$ is the interstellar cosmic ray ionization rate, $\Sigma(z)$ is the integrated density column from the top of the disc to the height z above disc mid-plane, and $\lambda_{\rm CR} = 96\,{\rm g\,cm^{-2}}$ is the attenuation length for cosmic rays (Umebayashi & Nakano 1981).

For the ionization rate of molecular hydrogen due to stellar X-rays, we use Bai & Goodman (2009) fits to the Igea & Glassgold (1999) Monte Carlo simulations,

$$\zeta_{\rm X}(z) = \frac{L_{\rm X}}{10^{29}\ \ {\rm erg}\ \ {\rm s^{-1}}} \left(\frac{r}{1\ \ {\rm AU}}\right)^{-2.2} \left(\zeta_1 e^{-(\Sigma(z)/\lambda_1)^{c_1}} + \zeta_2 e^{-(\Sigma(z)/\lambda_2)^{c_2}}\right), \tag{25}$$

where $L_{\rm X}$ is stellar X-ray luminosity, $\zeta_1 = 6 \times 10^{-12}\,{\rm s^{-1}}$, $\lambda_1 = 3.4 \times 10^{-3}\,{\rm g\,cm^{-2}}$, and $c_1 = 0.4$ characterize absorption of X-rays, and $\zeta_2 = 10^{-15}\,{\rm s^{-1}}$, $\lambda_2 = 1.6\,{\rm g\,cm^{-2}}$, and $c_2 = 0.65$ characterize the contribution from scattered X-rays. Here we have recalculated the attenuation lengths given by Bai & Goodman in terms of column densities of hydrogen nucleus into the surface density lengths using the hydrogen abundance given above. We adopt the saturated

---

[3] The total number densities of molecular hydrogen and potassium are then related to the gas density as $n_{\rm H_2} = x_{\rm H}/(2 - x_{\rm H})\rho/\mu$ and $n_{\rm K} = 2x_{\rm K}/(2 - x_{\rm H})\rho/\mu$, respectively.

[4] Note that the fiducial value of the hydrogen ionization rate used by Desch & Turner (2015) is in fact $\zeta = 1.4 \times 10^{-22}\,{\rm s^{-1}}$ (S. Desch, private communication).

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

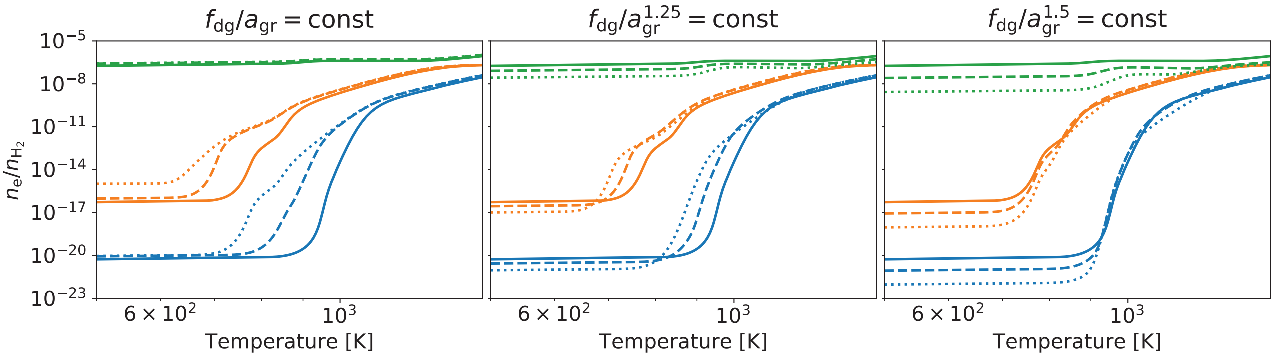


Figure 2. Ionization fraction (n_e/n_{H_2}) as a function of temperature. Different colours correspond to different combinations of pressure and hydrogen ionization rates: $P = 10^2$ dyn cm^{-2} and $\zeta = 10^{-19}$ s^{-1} (blue), $P = 10^{-2}$ dyn cm^{-2} and $\zeta = 10^{-19}$ s^{-1} (orange), $P = 10^{-2}$ dyn cm^{-2} and $\zeta = 10^{-11}$ s^{-1} (green). Different linestyles correspond to different grain sizes, $a_{gr} = 10^{-5}$ cm (solid), $a_{gr} = 10^{-3}$ cm (dashed), $a_{gr} = 10^{-1}$ cm (dotted), each with a different dust-to-gas ratio such that the ratio f_{dg}/a_{gr}^p remains constant and as evaluated for $a_{gr} = 10^{-5}$ cm, $f_{dg} = 0.01$. Different panels show the calculations for different values of the exponent p, as indicated in each panel title. Exponent $p = 1$ is equivalent to keeping the total grain surface area constant. Adopting exponent $p = 1.5$ yields approximately the same threshold temperature at which ionization fraction sharply increases irrespective of grain size a_{gr}. See Section 2.4.2.

relationship $L_X = 10^{-3.5} L_{bol}$ (e.g. Wright et al. 2011). For both cosmic rays and X-rays, we ignore the contribution coming through the other side of the disc. This is valid since we find that the gas surface densities are mostly larger than the attenuation lengths of the ionizing particles and, even at low gas surface densities, this can only increase the ionization rates by at most a factor of 2.

2.5 Numerical methods

2.5.1 *Equilibrium vertical disc structure*

At a given orbital radius, magnetic field strength and grazing angle ϕ, the disc's vertical structure is determined as a solution to the boundary value problem given by equations (1), (7)–(9) and the ideal gas law, and, where convectively unstable, by equations (1), (7), and (10). This boundary value problem is solved using the shooting method (Press et al. 2002). The equations are integrated from the top of the disc ($z = z_{surf}$) to the disc mid-plane ($z = 0$). The height of the disc z_{surf} is then found, such that $F(0) = 0$. This root-finding problem is solved using the Ridders' method, with an exit criterion that $|F(0)| < 10^{-5} F(z_{surf})$.

Equations of the vertical disc structure are integrated on a fixed vertical grid (i.e. in the z-direction), with points uniform in the polar angle. Since equation (9) is an algebraic equation, and for numerical stability, we use a fully implicit integration method. For the radiative-transfer problem we use a Runge–Kutta method of the second order, i.e. the trapezoidal method. In the trapezoidal method the equations (1), (7)–(9) are discretized as a system of non-linear equations to be solved in every integration step (a system of non-linear equations in P_{n+1}, F_{n+1}, J_{n+1}, T_{n+1} to be solved by a root-finding algorithm),

$$\begin{aligned}
P_{n+1} &= P_n + \frac{h}{2}\Omega^2(-\rho_n z_n - \rho_{n+1} z_{n+1}),\\
F_{n+1} &= F_n + \frac{h}{2}(\Gamma_n + \Gamma_{n+1}),\\
J_{n+1} &= J_n + \frac{h}{2}(-\frac{3}{4\pi})(\rho_n \kappa_R(T_n) F_n + \rho_{n+1}\kappa_R(T_{n+1}) F_{n+1}),\\
0 &= 4\rho_{n+1}\kappa_P(T_{n+1})(\sigma_{SB} T_{n+1}^4 - \pi J_{n+1}) - \Gamma_{n+1},
\end{aligned}$$

where h is the integration step, $\rho = \rho(T, P)$ is given by the ideal gas law, and $\Gamma = \Gamma_{acc} + \Gamma_{irr}$. Here, $\Gamma_{irr,\,n} = \Gamma_{irr,\,n}(T_n, P_n, \tau_{irr,\,z,\,n})$, with the optical depth to stellar irradiation obtained using $\tau_{irr,z,n+1} = \tau_{irr,z,n} + \frac{h}{2}\kappa_P^*(\rho_n + \rho_{n+1})$. Furthermore, viscous dissipation is a function of the MRI-driven viscosity α and thus depends on the local ionization levels. The latter are a function of the local temperature, pressure, and the hydrogen ionization rate which depends on the column density from the top of the disc. Thus, $\Gamma_{acc,\,n} = \Gamma_{acc,\,n}(T_n, P_n, N_n)$, with the column density given by $N_{n+1} = N_n + \frac{h}{2}{k_B}^{-1}(P_n/T_n + P_{n+1}/T_{n+1})$.

The equation for pressure P_{n+1} can be rearranged into an explicit form

$$P_{n+1} = \frac{P_n - \frac{h}{2}\Omega^2 \rho_n z_n}{1 + \frac{h}{2}\Omega^2 \frac{\mu m_H}{k_B}\frac{z_{n+1}}{T_{n+1}}}.$$

Then, the above system of equations is equivalent to a single non-linear equation in T_{n+1}, greatly simplifying the problem. In every integration step we use the Ridders' method to solve this equation for the temperature T_{n+1} (down to a relative precision of 10^{-7}) and consequently for all other quantities. This includes the MRI-driven viscosity α, which is thus calculated self-consistently at each step of integration.[5] At every step and in every iteration of the root-solver opacities are interpolated from pre-calculated tables using cubic splines, and the ionization levels (e.g. free electron number density) using tri-linear interpolation.

Additionally, at each integration step we check if the resulting temperature gradient is unstable to convection, and if so, the temperature T_{n+1} is obtained analytically using

$$T_{n+1} = T_n + \frac{h}{2}\nabla_{ad}\frac{\mu m_H}{k_B}\Omega^2(-z_n - z_{n+1}).$$

With T_{n+1} known, all other quantities follow same as above. A disc column can, in principle, become convectively stable again at some height above disc mid-plane. To calculate the mean intensity J at a boundary between a convective and a radiative zone, we use the energy balance equation (9).

For some model parameters, there can be a range of orbital radii and values of the magnetic field strength for which there are multiple solutions for the disc height z_{surf} (i.e. multiple solutions for the

---

[5] This is indeed necessary. An iterative method in which the disc thermal structure is decoupled from the density structure and the heating terms (e.g. Dullemond et al. 2002) does not converge to a solution in the case of MRI-driven viscosity.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

equilibrium vertical disc structure). This happens when a complex ionization structure leads to strong variations in the viscosity α as a function of height above disc mid-plane, making the total produced viscous dissipation a non-monotonous function of z_{surf}. When there are multiple solutions, we choose the solution with smallest z_{surf}. It is likely that at least some of the additional solutions are thermally unstable and/or unphysical, as the strong variations in both the levels of turbulence and the levels of ionization should be removed by turbulent mixing (see Section 3.2.3). Note that the multiple solutions in our model, when they exist, correspond to the same, radially constant input accretion rate. Steady accretion is thus assured regardless of the choice of the solution.

2.5.2 *Magnetic field strength*

At a given orbital radius, for a fixed grazing angle, ϕ, and magnetic field strength, B, the above procedure yields an equilibrium vertical disc structure characterized by a vertically averaged viscosity

$$\bar{\alpha} = \frac{\int_0^{z_{\mathrm{surf}}} \alpha P \mathrm{d}z}{\int_0^{z_{\mathrm{surf}}} P \mathrm{d}z}.$$

As in the vertically isothermal model (Mohanty et al. 2018), for a sufficiently small and a sufficiently large magnetic field strength B the MRI is suppressed in the entire disc column and $\bar{\alpha} = \alpha_{\mathrm{DZ}}$. There can be an intermediate range of magnetic fields strengths for which the MRI is active, and the vertically averaged viscosity $\bar{\alpha}$ peaks at some value of B. At every orbital radius, we choose B such that $\bar{\alpha}$ is maximized. The underlying assumption here is that the magnetic fields are strengthened by the MRI-driven turbulence. It is assumed that the initial magnetic field configuration was sufficient to start the instability. The induced magnetohydrodynamic turbulence then amplifies the field strength. We assume the equilibrium configuration is one in which the turbulence is maximized, where the turbulence is parametrized by $\bar{\alpha}$. Similar arguments were employed by Bai (2011) and Mohanty et al. (2013).

To maximize $\bar{\alpha}(B)$, we use the Brent method with a target absolute precision of 10^{-3} in log$B$. There can be a range of orbital radii where there are multiple local maxima in $\bar{\alpha}$ as a function of B. This is essentially for the same reasons that cause multiple solutions in disc height z_{surf} at a fixed value of B. In general, we choose B corresponding to the global maximum in $\bar{\alpha}$. However, in some cases, this is a function of the grazing angle ϕ at a fixed orbital radius, and the procedure to determine the grazing angle (described below) does not converge. There we choose a local maximum with a largest magnetic field strength.

2.5.3 *Grazing angle*

At any given orbital radius, the angle between the incident stellar radiation and the disc surface is determined self-consistently with the disc structure following D'Alessio et al. (1999). A self-consistent disc structure is found by iteratively updating the grazing angle and re-calculating the entire disc structure. We use a logarithmic grid for orbital radius. First, we calculate ϕ using equation (6) by assuming that $z_{\mathrm{irr}} = 0$ and solve for the vertical disc structure and the magnetic field strength at all radii. We integrate through the obtained disc structure along lines of sight to the star to calculate $\tau_{\mathrm{irr}}(r, z)$, which yields an updated z_{irr} at each radius.

Critically, to calculate the updated value of the grazing angle ϕ at each radius, the derivative dlogz_{irr}/dlog$r$ is approximated by assuming that z_{irr} is a power law, $z_{\mathrm{irr}} \propto r^b$, within a radius bin centred at the given radius. So, at each radius we fit for the slope b using z_{irr} at that radius and at a number of radial grid points interior and exterior to it. Then, the value of the grazing angle is updated and the vertical disc structure recalculated at all radii.

This procedure is repeated until the grazing angle has converged at every radius, i.e. until the relative difference in ϕ between two consecutive iterations is less than 0.5 per cent at all radii. For the first radial point, we always assume the flat-disc approximation ($z_{\mathrm{irr}} = 0$) and do not include it in the fitting routine. In this work, we use a total of 100 radial points between 0.1 and 1 AU, and a total of 10 radial points in fits for dlogz_{irr}/dlog$r$.

3 RESULTS

Using our self-consistent model, we can now investigate the structure of the inner regions of a protoplanetary disc that is viscously accreting with a constant accretion rate.

In Section 3.1, we consider a model in which the vertical structure of the disc is calculated self-consistently from viscous heating, heating by stellar irradiation, radiative, and convective energy transport, and self-consistent radiative properties of dust, while the disc's ionization state is calculated by only considering the thermal (collisional) ionization of potassium, using the Saha equation.

In Section 3.2, we present results for our full model that includes additional chemical species, including dust grains, in the chemical network, as well as non-thermal sources of ionization.

Unless otherwise stated, throughout this paper we assume a solar-mass star, $M_* = 1\,\mathrm{M}_{\odot}$, with a stellar radius $R_* = 3\,\mathrm{R}_{\odot}$, effective stellar temperature $T_* = 4400\,\mathrm{K}$, gas accretion rate $\dot{M} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, and viscosity in the MRI-dead zone $\alpha_{\mathrm{DZ}} = 10^{-4}$. Our adopted gas accretion rate is the median from observations of solar-mass classical T Tauri stars (e.g. Hartmann et al. 1998; Manara et al. 2016, 2017). In reality, there is more than an order of magnitude scatter in the data around this median; we explore the effects of varying $\dot{M}$, as well as other parameters, in a companion paper, as noted below. The stellar parameters are from the evolutionary models of Baraffe et al. (2015), for a solar-mass star at an age of 5×10^5 yr. At later times, the stellar luminosity decreases, as the star contracts towards the zero-age main-sequence. Our choice of a relatively young classical T Tauri star thus maximizes the stellar luminosity, allowing us to examine the largest possible effect of stellar irradiation on the inner disc. Lastly, we adopt a standard ISM dust-to-gas ratio of $f_{\mathrm{dg}} = 10^{-2}$, and a maximum dust grain size of $a_{\mathrm{max}} = 1\ \mu\mathrm{m}$.

The above set of parameters represents our fiducial model. In this work, we investigate in detail the main physical and chemical processes that shape the structure of the inner disc for this model. In a companion paper (Jankovic et al., in preparation), we examine how the inner disc structure changes as a function of the model parameters, and build a picture of how this structure may affect planet formation.

3.1 Disc thermal structure and the MRI

In this section, we explore how thermal processes shape the inner disc structure. The ionization state of the disc is set here only by the thermal ionization and recombination of potassium, which is assumed to be entirely in the gas phase. Note that there is no other chemistry in this model: in the absence of grain-related reactions (grains here only contribute to the opacity), and without non-thermal sources to ionize H_2, gas-phase thermal ionization and recombination of potassium are the only reaction pathways available.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

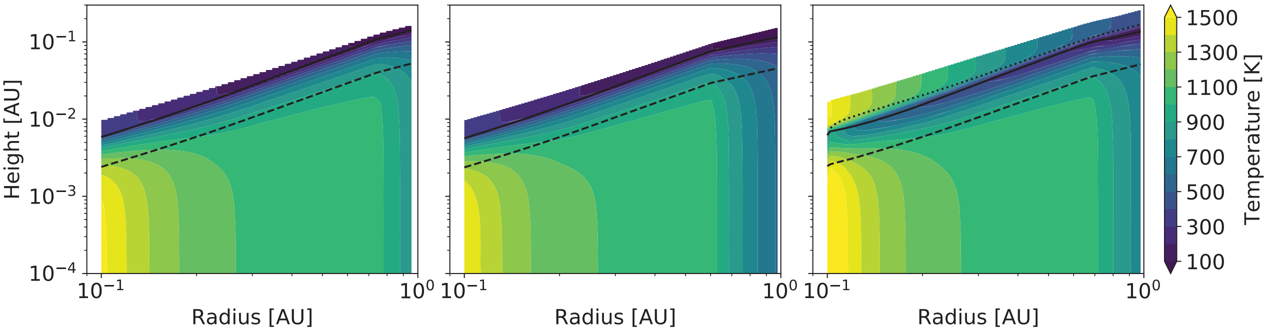


Figure 3. Temperature as a function of location in the disc for a viscously heated constant-opacity model (left-hand panel), viscously heated model with realistic opacities (middle), and a viscously and irradiation-heated model with realistic opacities (right-hand panel). In each panel, the solid line shows the disc photosphere ($\tau_R = 2/3$) and the dashed line shows the pressure scale height ($P = e^{-1/2} P_{\rm mid}$). The dotted line in the right-hand-side panel shows the surface at which $\tau_{\rm irr} = 2/3$. Note that the inclusion of heating by stellar irradiation does not strongly affect the disc mid-plane temperature. See Section 3.1.1.

Mohanty et al. (2018) presented a model in which the disc is vertically isothermal, with a constant opacity ($= 10\,{\rm cm}^2\,{\rm g}^{-1}$). Throughout this section, we shall compare this simple model with three models of increasing complexity within our framework: (i) a model with the same constant opacity ($\kappa_R = \kappa_P = 10\,{\rm cm}^2\,{\rm g}^{-1}$), but with the disc vertical structure calculated self-consistently from viscous heating and cooling by radiation and convection; (ii) a model with the same heating and cooling processes as (i), but now with opacities also determined self-consistently; and (iii) similar to (ii), but now also including heating by stellar irradiation.

To make it easier to interpret the effects of these thermal processes, we first focus on how the disc's temperature changes when the increasingly complex models are used. We then discuss how this leads to differences in the disc's ionization state, the locations where the MRI is active, and the radial surface density and mid-plane pressure profiles. Finally we provide a short summary of the key findings at the end of this section.

3.1.1 *Thermal structure of the inner disc*

Fig. 3 shows the temperature as a function of disc radius and height above the mid-plane for our three models, increasing in complexity from the left-hand to right-hand panels.

Noticeably, the temperature profiles deviate from being vertically isothermal. The dashed lines here show the disc pressure scale height, and the solid lines show the disc photosphere for outgoing radiation (i.e. where the Rosseland-mean optical depth is $\tau_R = 2/3$). In all three models, the temperature increases towards the mid-plane below the photosphere, as the disc becomes more and more optically thick to its own radiation.

The resulting temperature gradient is sufficiently high to make the disc convectively unstable, as shown in Fig. 4. From the mid-plane up to a couple of pressure scale heights, energy is thus transported by convection, and the temperature gradient here is essentially isentropic. Importantly, the strong temperature gradient that yields this convective instability is not specific to MRI-driven accretion: it is a general feature of any active, optically thick disc where the accretion heat is released near the mid-plane (see also Garaud & Lin 2007). We show this analytically in Section 4.

The model that includes heating by stellar irradiation, shown in the right-hand panels of Figs 3 and 4, features a temperature inversion in the disc upper layers. This inversion has been discussed in detail by D'Alessio et al. (1998), and is a simple consequence of the fact that the photosphere for the outgoing radiation (solid line) lies below the photosphere for the incoming stellar radiation (dotted line), where the latter corresponds to the irradiation surface $z_{\rm irr}$ at which $\tau_{\rm irr} = 2/3$. Above this surface, the disc upper layers are heated by stellar irradiation. Below, the disc becomes optically thick to stellar radiation (i.e. stellar photons do not penetrate here) and the temperature drops. Going deeper still, below the disc photosphere to the outgoing radiation, the disc becomes optically thick to its own radiation and the temperature rises again.

In the irradiated disc, close to the star the disc mid-plane is as hot as the disc upper layers, and further away the mid-plane is significantly hotter than the upper layers. It would thus appear that accretion heating dominates in the inner disc. However, at a given radius, the total flux of stellar radiation absorbed by a vertical column in the disc is in fact at least an order of magnitude greater than the total flux generated by viscous dissipation ($F_{\rm irr} \gtrsim 10 F_{\rm acc}$), throughout the inner disc (see top panel of Fig. 5). In spite of this, the disc temperature near the mid-plane is only weakly affected by irradiation. We discuss this result in Section 4.2.

The ratio $F_{\rm irr}/F_{\rm acc}$ varies non-monotonically, following the grazing angle ϕ, shown in the bottom panel of Fig. 5. At the inner edge of our calculation domain, the first $\sim$10 points (shown in grey) are affected by boundary effects. This is a well-known problem in disc models that account for stellar irradiation using the grazing angle prescription (Chiang et al. 2001). Importantly, far from the inner edge, the value of the grazing angle and the disc structure do not depend on the disc structure at the inner edge.

3.1.2 *Ionization levels and non-ideal MHD effects*

The ionization structure in our models with a self-consistent vertical structure is qualitatively different from that in the vertically isothermal model of Mohanty et al. (2018). In the latter, the temperature is constant with height above the mid-plane while the density decreases, and thus the ionization fraction ($n_e/n_{\rm H_2}$) increases with height (by the Saha equation). In our models without stellar irradiation heating, the temperature declines monotonically with height, and hence the ionization fraction decreases as well (Fig. 6: middle row, solid curves). When we include irradiation heating, on the other hand, the fractional ionization does increase in the uppermost irradiated layers, due to the rising temperature there (dashed curves).

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

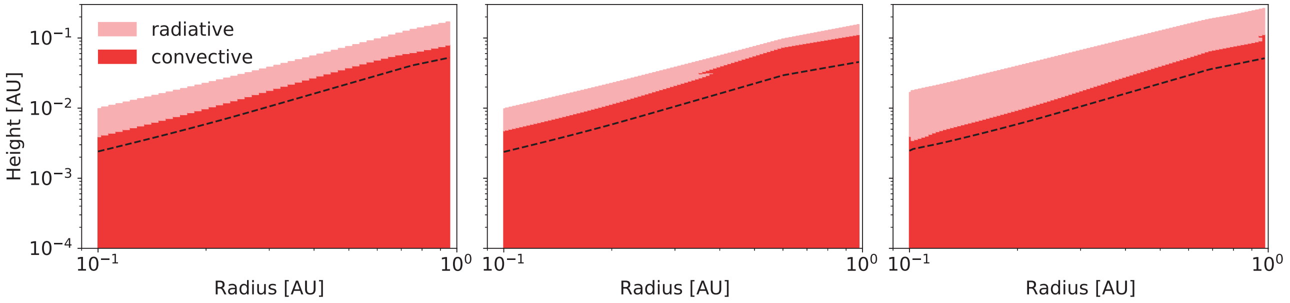


Figure 4. Radiative and convective zones (light and dark red, respectively) for a viscously heated constant-opacity model (left-hand panel), viscously heated model with realistic opacities (middle), and a viscously and irradiation-heated model with realistic opacities (right-hand panel). In each panel, the dashed line shows the pressure scale height ($P = e^{-1/2}P_{\rm mid}$). In all three models, the disc is convectively unstable within few scale heights. See Section 3.1.1.

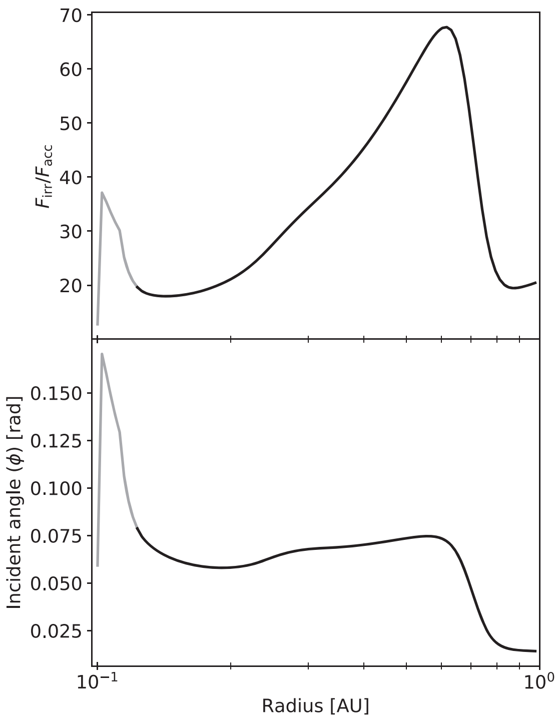


Figure 5. Top: Ratio of the total irradiation heating to the total viscous dissipation ($F_{\rm irr}/F_{\rm acc}$) as a function of radius. The total (vertically integrated) absorbed stellar flux $F_{\rm irr}$ is at least an order of magnitude larger than the total viscous dissipation $F_{\rm acc}$ at any given radius. Bottom: Grazing angle of stellar irradiation (ϕ) as a function of radius. Note that the $F_{\rm irr}/F_{\rm acc}$ curve follows the trend in ϕ. In both panels, the grey portion of the curve indicates the region affected by the inner boundary condition on ϕ. See Section 3.1.1.

As a consequence of the above, both ambipolar and Ohmic resistivities increase with height in our non-irradiated discs. In other words, in these models, both ambipolar and Ohmic diffusion quench the MRI from above. In the irradiated disc, the Ohmic resistivity falls in the uppermost layers; however, ambipolar diffusion still increases in these layers, again stifling the MRI there. Consequently, the vertical extent of the MRI-active region (i.e. where the local $\alpha > \alpha_{\rm DZ}$) at any given radius in our irradiated disc model is nearly identical to that in the non-irradiated discs (Fig. 6, top row).

3.1.3 MRI-active and dead zones

The differences in the ionization structure of our vertically self-consistent models and the vertically isothermal model lead to differences in where the MRI is active in the disc. A plot of the viscosity parameter α as a function of disc radius and height above mid-plane, Fig. 7, shows that the MRI-active zone (i.e. where $\alpha > \alpha_{\rm DZ}$) is confined to the vicinity of the mid-plane in our models. The same was found by Terquem (2008), who considered a similar vertically self-consistent disc model, for a simple viscosity prescription in which the MRI is activated above a fixed critical temperature. This is qualitatively different from the vertically isothermal case, where the MRI-active zone occurs around the mid-plane in the innermost region, but rises into the upper layers at larger radii (see Mohanty et al. 2018). The latter configuration, where the active zone is sandwiched between a dead zone around the mid-plane below and an ambipolar-dead zone above, emerges as a consequence of the ionization increasing with height above the disc mid-plane in the isothermal case, as discussed above. Moreover, comparing our non-irradiated discs (first two panels of Fig. 7) with the irradiated one (last panel), we see that heating by stellar irradiation makes little difference to the extent of the MRI-active zone, since it only weakly affects the mid-plane temperature, and the hot uppermost layers are dead due to ambipolar diffusion.

Finally, in Fig. 8 we compare radial profiles of the vertically averaged viscosity parameter $\bar{\alpha}$ and the MRI-generated magnetic field strength B from the four models (the vertically isothermal and constant opacity disc of Mohanty et al. 2018, plus our three non-isothermal discs).

In line with theoretical expectations and the results of Mohanty et al. (2018), the vertically averaged viscosity parameter $\bar{\alpha}$ decreases as a function of orbital radius. At some distance from the star the MRI is completely quenched and the average viscosity parameter reaches the minimum value, $\bar{\alpha} = \alpha_{\rm DZ}$. Interestingly, the $\bar{\alpha}$ radial profile is both qualitatively and quantitatively similar in all four models. However, the radial profile of the magnetic field strength B reveals qualitative differences. In the vertically isothermal model, $B(r)$ features a sharp drop at $\sim$0.35 AU, which corresponds to the appearance of a dead zone in the disc mid-plane, as described above. From that point, until the MRI is completely smothered at $\sim$0.7 AU, the MRI is active in a thin layer high above the mid-plane, between a dead zone below and an ambipolar-dead zone above (see Mohanty et al. 2018). In the self-consistent models,

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

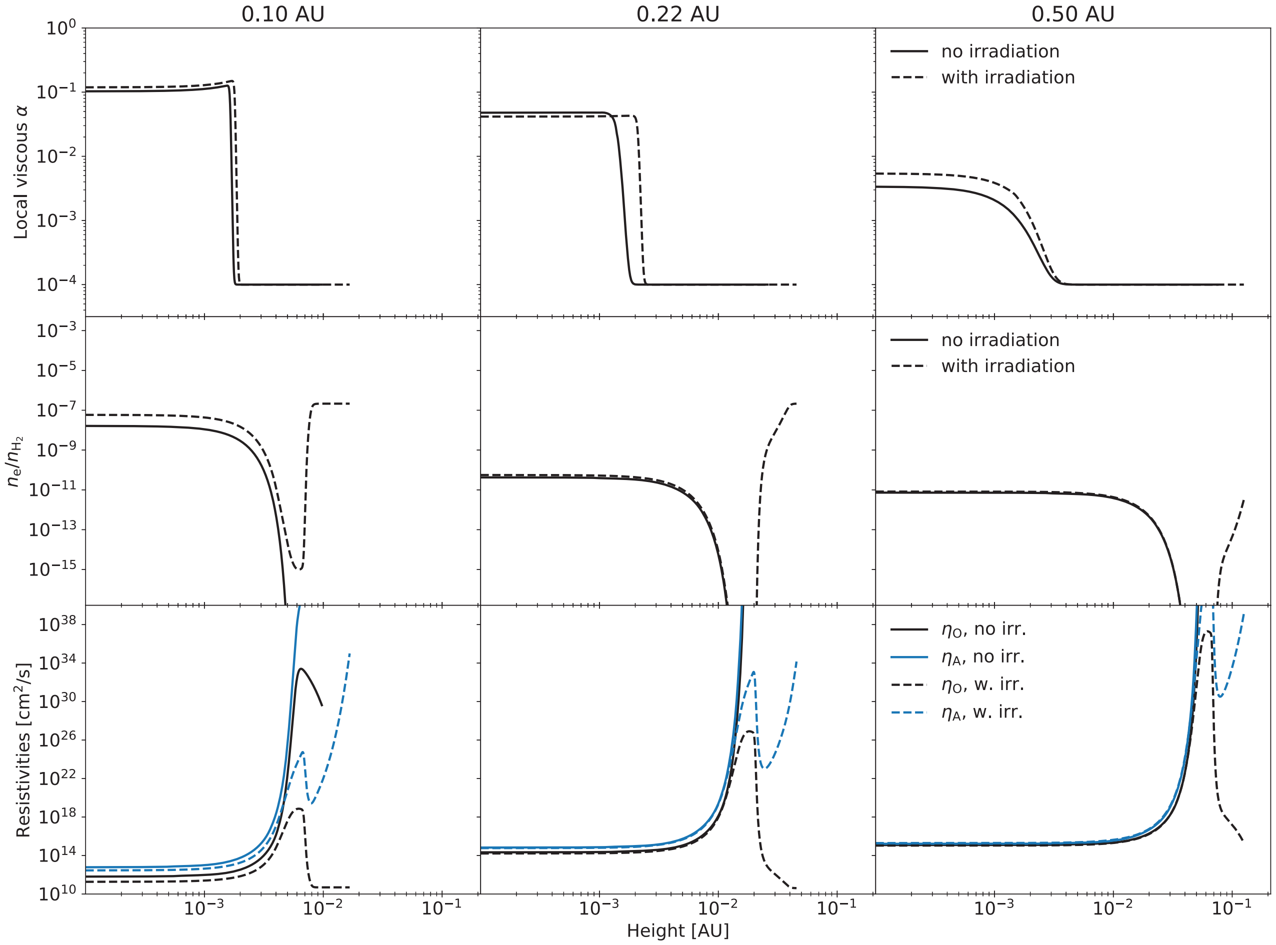


Figure 6. Local viscous parameter α (top), ionization fraction n_e/n_{H_2} (middle), and magnetic resistivities (bottom) at three different disc radii (as indicated in the panel titles), in two different models: one with realistic opacities but no irradiation (solid lines), and one that also includes irradiation (dashed lines). Despite the high ionization fraction in the irradiation-heated disc upper layers, η_A is still very high in these layers, i.e. ambipolar diffusion quenches the MRI there. See Section 3.1.2.

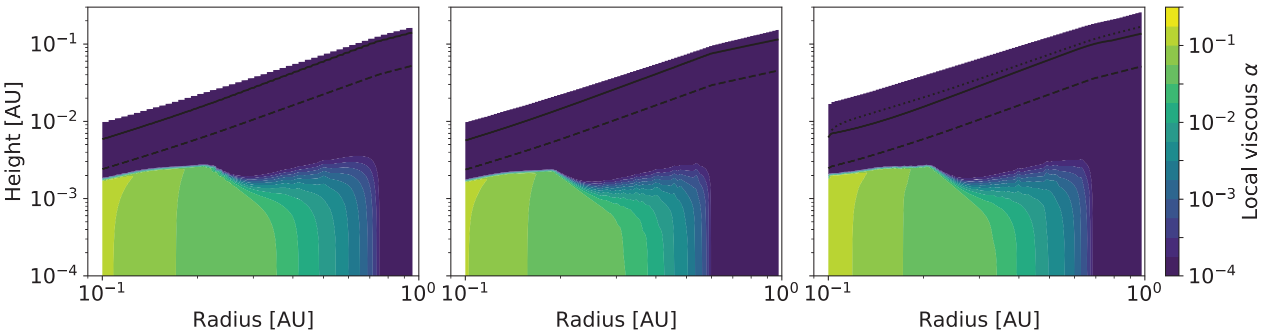


Figure 7. Local viscous parameter α as a function of location in the disc for a viscously heated constant-opacity model (left-hand panel), viscously heated model with realistic opacities (middle), and a viscously and irradiation-heated model with realistic opacities (right-hand panel). In each panel the solid line shows the disc photosphere ($\tau_R = 2/3$) and the dashed line shows the pressure scale height ($P = e^{-1/2}P_{mid}$). The dotted line in the right-hand-side panel shows the surface at which $\tau_{irr} = 2/3$. Note that the heating by stellar irradiation has a very weak effect on the extent of the MRI-active region. See Section 3.1.3.

however, this configuration does not appear: the magnetic field strength remains strong until the MRI is completely quenched, and the MRI-active zone in the inner disc is always restricted to the mid-plane regions.

3.1.4 Surface density and pressure maximum

All four steady-state models discussed above feature a maximum in both the local gas pressure and the surface density (Fig. 9), at the radial location where the vertically averaged viscosity parameter $\bar{\alpha}$

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

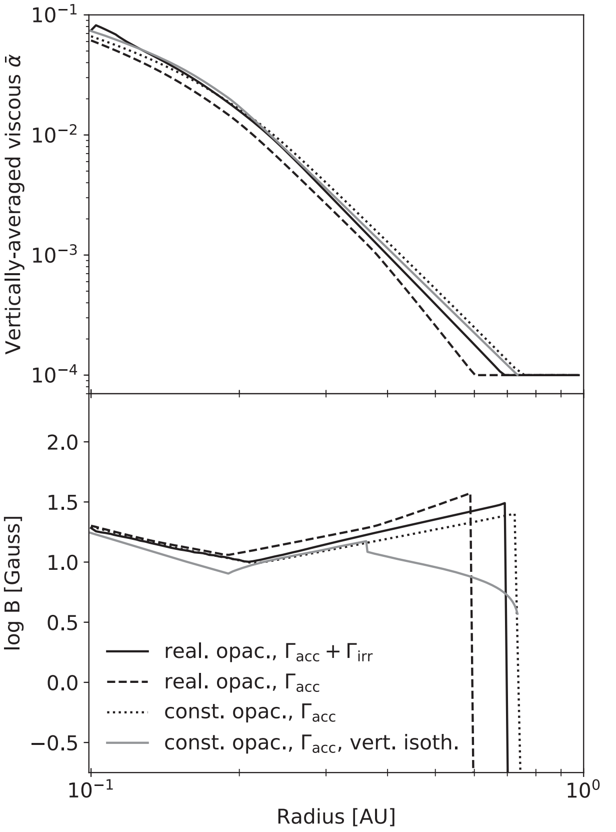


Figure 8. Vertically averaged viscosity parameter ($\bar{\alpha}$) and the MRI-generated magnetic field strength (B) as functions of radius, in a vertically isothermal model as well as in self-consistent models of varying complexity (viscously heated constant-opacity model, viscously heated model with realistic opacities, and viscously and irradiation-heated model with realistic opacities), as indicated in plot legend. The vertically averaged viscosity parameter profile is similar in all four models, but the magnetic field strength profile is qualitatively different in the vertically isothermal model (in which the field strength decreases as a function of radius outwards from ∼ 0.4 AU). See Section 3.1.3.

falls to its minimum value (see Fig. 8). Inwards of this location, the MRI-driven accretion efficiency (i.e. $\bar{\alpha}$) increases; for a radially constant (i.e. steady-state) gas accretion rate, this leads to a decrease in both the surface density and the mid-plane pressure. Since the radial profile of $\bar{\alpha}$ is similar in the four models, their surface density and mid-plane pressure profiles are similar too. Within our three self-consistent models, the addition of heating by stellar irradiation moves the pressure maximum outwards by ∼0.1 AU (compare the dashed and solid black lines in Fig. 9, bottom panel). At the same time, relative to the result of Mohanty et al. (2018) for an isothermal and constant opacity disc (grey solid line), the pressure maximum in our most complex model in this section (non-isothermal, with realistic opacities and irradiation heating; black solid line) has moved ∼0.1 AU inwards.

3.1.5 Summary

In this section, we have considered how the details of the disc thermal structure model affect the disc's ionization state and where in the disc the MRI is active. We have also examined the importance of the various physical effects on the location of the MRI-induced pressure maximum in the inner disc.

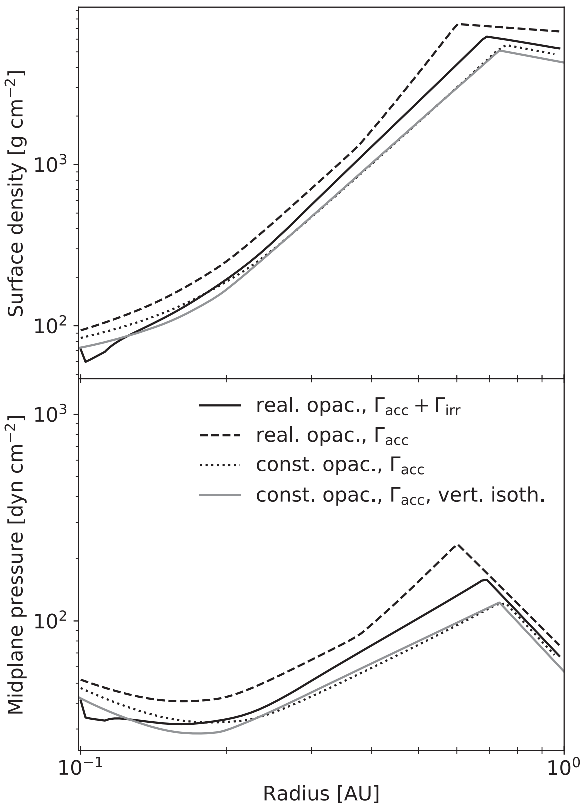


Figure 9. Surface density (top) and mid-plane pressure (bottom) as functions of radius in a vertically isothermal and self-consistent models of varying complexity (same as in Fig. 8). In each model the density and the mid-plane pressure maximum correspond to the point at which the viscous $\bar{\alpha}$ reaches a minimum value. See Section 3.1.3.

Our results show that when the disc vertical structure and the MRI-driven viscosity are considered self-consistently, the MRI is active around the disc mid-plane. In previous work using a simplified vertically isothermal model, a different configuration arises in the vicinity of the pressure maximum, where the MRI is active in a layer above (and below) the dead mid-plane. Though both approaches yield a similar location for the pressure maximum for our fiducial parameters, the difference in the physical behaviour of the disc cautions against the use of the simplified model.

Furthermore, for the fiducial parameters considered here, we find that the mid-plane temperature is primarily determined by viscous dissipation, and not stellar irradiation. As a result, the location of the pressure maximum, where the mid-plane temperature and the ionization levels fall below a critical value, is negligibly affected by stellar irradiation. Finally, we find that the inner disc is convectively unstable, even if the disc opacity is constant. Hence, even though the location of the pressure maximum is similar to that in previous work, the disc's structure is qualitatively different. These findings are further discussed in sections 4.2 and 4.3.

3.2 Disc chemical structure and the MRI

In this section, we build upon our models from the previous section. We first explore the effects of dust on the disc's ionization state and on the MRI. We then construct our complete model of the inner disc by also considering non-thermal sources of ionization.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

3.2.1 Effects of dust

Here we consider two models. The first is our most complex model from the previous section, which incorporates a self-consistent vertical structure and realistic opacities, and includes stellar irradiation, but accounts for only gas-phase thermal ionization and recombination of potassium (solid lines in Figs 10 and 11). The second is a model including all these processes, as well as dust effects on the ionization (dashed lines in Figs 10 and 11); the latter effects include adsorption of neutral atoms and free charges on to dust grains, recombinations on the grain surfaces, and thermionic and ion emission from the grains.

Fig. 10 depicts the radial profiles of the vertically averaged viscosity $\bar{\alpha}$, magnetic field strength, mid-plane temperature, and the mid-plane fractional ionization (n_e/n_{H_2}) for the two models. We see that the radial profile of $\bar{\alpha}$, and the radial extent of the MRI-active zone (i.e. where $\bar{\alpha} > \alpha_{DZ}$), are similar in the two models. Consequently, their surface density and mid-plane pressure profiles, plotted in Fig. 11, are also similar. Interestingly, Fig. 10 also shows that when the influence of dust on the ionization is included, both $\bar{\alpha}$ and the ionization fraction are somewhat higher at a given radius within the active region, while the mid-plane temperature is lower. We discuss these effects further in Section 4.1.

Fig. 12 illustrates the vertical structure of the model with dust at three different disc radii. The top row shows the local viscosity parameter α, the middle row shows the number densities of free electrons and ions, and the bottom row shows the various contributions to the free electron production rate per unit volume. As in the thermally ionized disc (Fig. 6), the MRI is active in the hot disc mid-plane. The ionization levels decrease with height above the mid-plane as the temperature decreases, and increase again in the disc atmosphere heated by stellar irradiation. In the hot upper layers ambipolar diffusion quenches the MRI.

The plots of free electron and ion production rates show that thermionic and ion emission are the dominant ionization sources. Thus, Fig. 10 is misleading in the sense that, while the differences in the global structure are very small when dust effects are added, it is not because dust effects are minor. Clearly, dust dominates the chemistry in the inner disc. Rather, for the parameters assumed here, the ionization levels as a function of temperature and density, when dust is included, are similar to the levels obtained from gas-phase thermal ionization only, due to the similarity between the ionization potential of potassium and the grain work function.

3.2.2 Non-thermal sources of ionization

In this section, we present our full model of the inner disc which, in addition to the above dust effects, also includes non-thermal sources of ionization: stellar X-rays, cosmic rays, and radionuclides all ionize H_2, ultimately producing metallic ions and free electrons. The resulting radial profiles of $\bar{\alpha}$, magnetic field strength, mid-plane temperature, density, and ionization levels are shown in Fig. 10 as dotted lines. The radial profile of $\bar{\alpha}$ seems to overlap with the model with only dust effects (dashed line). However, in the region where the MRI is dead in the dust-only model, $\bar{\alpha}$ is slightly higher than α_{DZ} in the new model including H_2 ionization. The magnetic field strength also does not drop to zero in this model, further revealing that the MRI remains active here.

Fig. 13 shows the vertical structure of this disc at three different radii. It reveals that, at large radii where the MRI was quenched in previous models, the MRI now remains active at large heights (see also Fig. 14). The situation is reminiscent of the appearance of

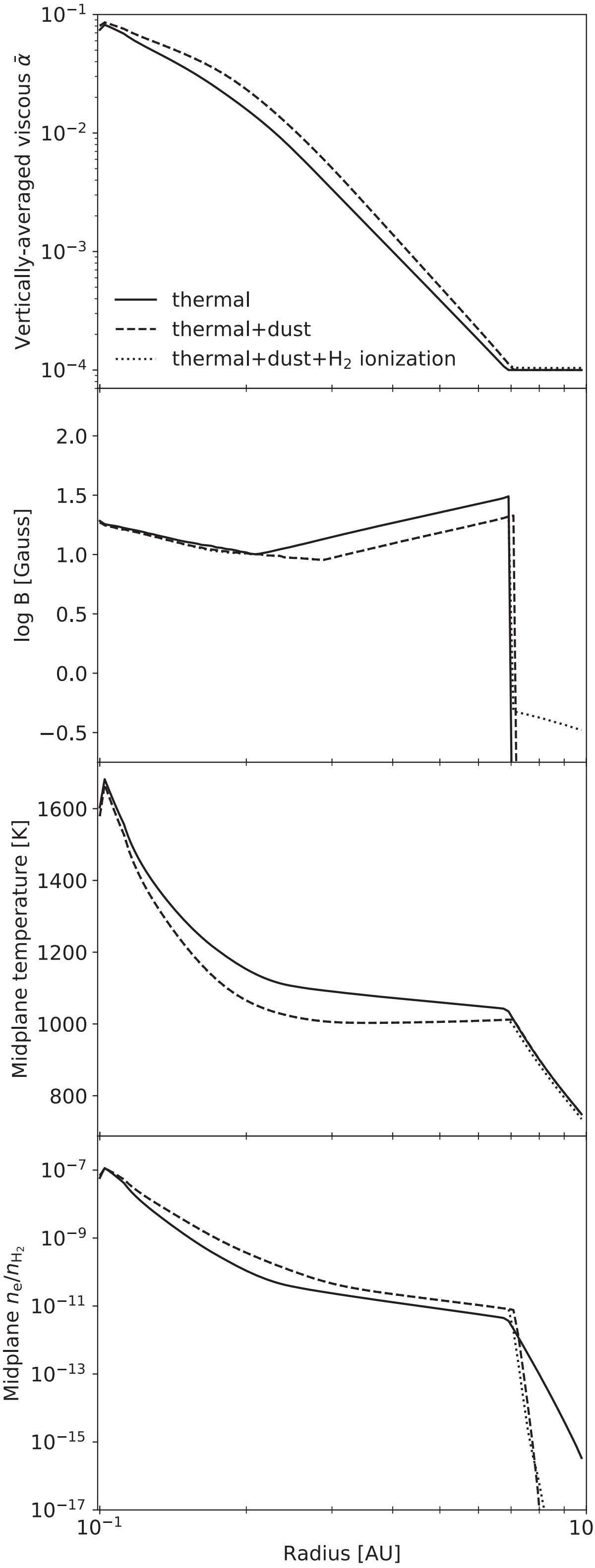


Figure 10. Vertically averaged viscous parameter ($\bar{\alpha}$), MRI-generated magnetic field strength (B), mid-plane temperature, and the mid-plane free electron fraction (n_e/n_{H_2}) as functions of radius, for a model with thermal ionization only, a model with thermal ionization and dust effects, and a model which also includes realistic non-thermal sources of ionization of H_2 ($\zeta = \zeta_R + \zeta_{CR} + \zeta_X$). The disc structure is quantitatively similar in all three models; however, the main sources of ionization in the models with dust are thermionic and ion emission. See Section 3.2.1.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

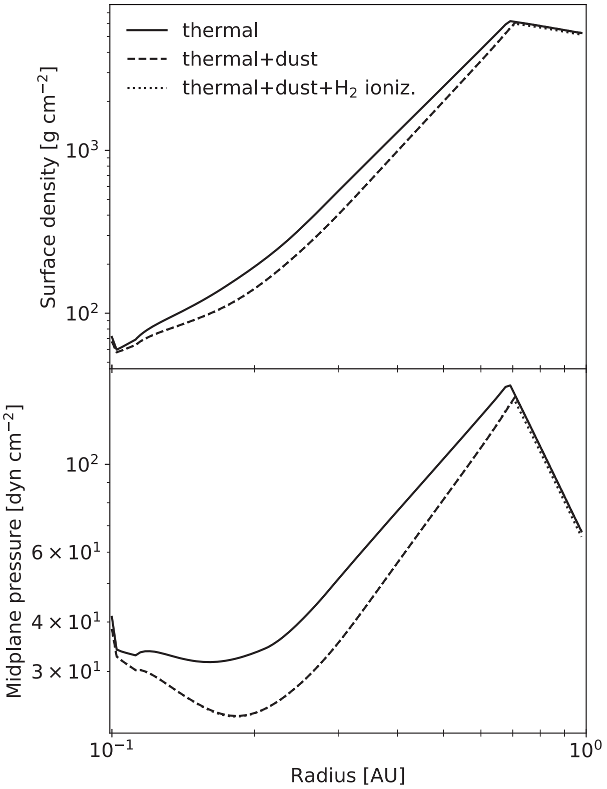


Figure 11. Surface density (top) and mid-plane pressure (bottom) as functions of radius for a model with thermal ionization only, a model with thermal ionization and dust effects, and a model which also includes realistic non-thermal sources of ionization of H_2 ($\zeta = \zeta_R + \zeta_{CR} + \zeta_X$). The disc structure is similar in all three. See Section 3.2.1.

an MRI-active layer near the disc surface in the thermally ionized, vertically isothermal model of Mohanty et al. (2018), but the physical reason is very different: in the latter model, it is due to the isothermal assumption, which leads to increasing fractional ionization with height; in our present non-isothermal model, it is due to additional, non-thermal sources of ionization (mainly X-rays; see below) which elevate the fractional ionization near the disc surface.

The plot of ionization levels in Fig. 13 shows that the fractional abundances of free electrons and metal ions increase strongly towards the disc surface, where non-thermal ionization dominates. As a result, potassium ions are depleted in the upper layers via recombination with the abundant electrons. Furthermore, the values of the metal ion abundance near the surface are themselves noteworthy: they greatly exceed the solar abundance of Mg, and even of C and O. The reason is as follows. In our chemical network, the ionization of an H_2 molecule by a non-thermal source effectively produces a free electron and a metal ion, the latter through (implicitly included) rapid charge exchange between a metal atom and the H_2 ion. This is valid as long as the number of H_2 ions remains lower than the total number of metal atoms; in this case, the precise total abundance of metals is unimportant, and is not accounted for in our calculations. When the H_2 ion abundance exceeds that of metal atoms, however, our network fails: it yields a spuriously high metal ion abundance, when in reality H_2 ions dominate (since there are no remaining metal atoms to transfer their charge to). Additionally, when the H_2 ion abundance becomes very high (e.g. when $n_{\mathrm{H}_2^+}/n_{\mathrm{H}_2}$ exceeds $\sim10^{-3}$), our assumption that H_2 is effectively a neutral species due to charge transfer also breaks down.

Clearly, above this level our chemical model is not applicable, as the ionized hydrogen would become an important species and our assumption of a constant hydrogen number density would be invalid. Nevertheless, this is above the MRI active zone at all orbital radii (see the black solid line in Fig. 14, and also the blue solid line in Fig. 13), and is thus not germane to our conclusions.

Fig. 15 compares the contributions from stellar X-rays and cosmic rays to the hydrogen ionization rate. As expected, the un-attenuated ionization rate due to X-rays is higher at the disc surface, but cosmic rays can penetrate deeper in the disc. Nevertheless, the MRI-active region in the upper disc layers is mostly ionized by X-rays, in agreement with previous work (e.g. Glassgold et al. 1997; Ercolano & Glassgold 2013).

Finally, note that in this outer region, where the MRI is active in the upper disc, the gas accretes primarily through the dead zone (driven by the dead-zone viscosity α_{DZ}), since the density at the dead disc mid-plane is much higher than in the X-ray-ionized MRI-active layer. This is why the vertically averaged viscosity parameter is close to the dead-zone value, $\bar{\alpha} \sim \alpha_{\mathrm{DZ}}$, outwards from the pressure maximum.

3.2.3 *Multiple solutions for the vertical disc structure*

In the results presented so far, for our fiducial choice of disc and dust parameters, solutions for the vertical disc structure (i.e. solutions in disc height z_{surf}) appear to be unique. In general, there is also a single peak in the vertically averaged viscosity $\bar{\alpha}$ as a function of magnetic field strength B (which determines our choice for B; see Section 2.5.2). The exception is the vicinity of the orbital radius at which the MRI is quenched at the disc mid-plane (see Fig. 14). There, $\bar{\alpha}(B)$ has two peaks, one corresponding to the solution where the MRI is active at disc mid-plane, and the other to the solution where the MRI is active in the upper disc layers, mostly ionized by stellar X-rays. As discussed in Section 2.5, we choose B such that $\bar{\alpha}$ is maximized in this case as well.

Note that there could be, in principle, an MRI-active layer high up in the disc at shorter orbital radii as well, in addition to the active layer at mid-plane. Here, this does not appear due to our assumption that the magnetic field strength B is vertically constant. At the high B necessary to drive efficient accretion at mid-plane, the MRI is quenched in the low-density disc atmosphere due to ambipolar diffusion (since the ambipolar criterion for active MRI also encapsulates the requirement that the magnetic pressure should be less than the thermal pressure). Therefore, it is only when the temperature drops and high-temperature ionization effects (thermal ionization of potassium, thermionic, and ion emission from grains) can no longer drive the accretion efficiently in the mid-plane that our model features the active layer (generated by X-ray ionization) high above the disc mid-plane.

For a different choice of parameters, e.g. if the maximum dust grain size is $a_{\max} = 100$ µm, there may exist an additional range of orbital radii where there are multiple peaks in $\bar{\alpha}(B)$ and also multiple solutions for the disc vertical structure (for z_{surf}) at a fixed value of the magnetic field strength B. Similarly to the case above, this issue arises due to the competing effects of high-temperature sources of ionization and X-rays. As these sources of free electrons have different dependencies on the disc structure (density, temperature, column density), their combination leads to non-monotonous variations in the electron number density as a function of height above the disc's mid-plane. Since the viscous dissipation due to the MRI is sensitive to the ionization fraction, the total dissipation can

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

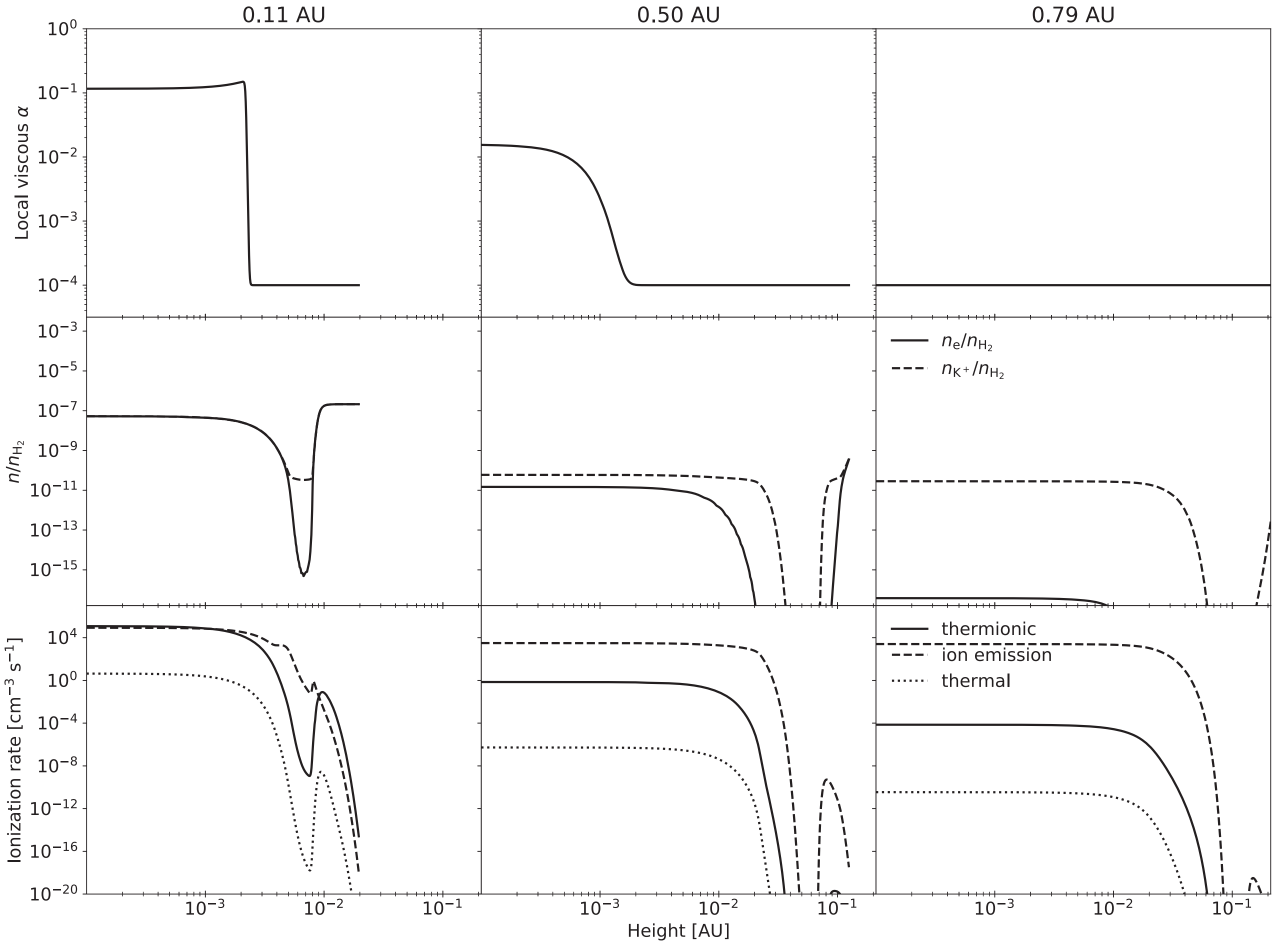


Figure 12. Local viscous parameter α (top), fractional abundance of charged species n/n_{H_2} (middle) and ionization rates (bottom; thermionic $\mathcal{R}_{\rm therm}$, ion emission $\mathcal{R}_{\rm K,evap} f_+$, thermal $k_2 n_{H_2} n_{K^0}$) at three different radii (as indicated in panel titles) for the model with thermal ionization and dust effects. Thermionic emission is the primary source of free electrons in the MRI-active regions. See Section 3.2.1.

be a non-monotonous function of $z_{\rm surf}$. Since the solution for the vertical disc structure is determined by an equilibrium between an input and an output total heat, this can lead to multiple solutions in $z_{\rm surf}$.

To illustrate this issue, we show in Fig. 16 an example of three thermally stable solutions for the vertical disc structure at a fixed value of magnetic field strength B that appears in our model for a maximum grain size $a_{\rm max} = 100\,\mu$m (we ignore thermally unstable solutions). Note that the dependence of the overall disc structure and the location of the pressure maximum on the dust grain size is presented and thoroughly discussed in our companion paper (Jankovic et al., in preparation). Here we only discuss how we deal with the multiple solutions. Fig. 16 shows the viscosity α as a function of height in the top panel and the ratio n_e/n_{H_2} in the bottom.

Evidently, small variations in the free electron number density correspond to large variations in the viscosity α, all at heights below one disc pressure scale height (indicated by grey lines). This implies that the difference between these solutions is likely unphysical for two reasons. First, the viscosity α is in reality driven by turbulence, and turbulent motions should not abruptly change over length scales much smaller than a single pressure scale height. Secondly, chemical species can also be expected to be spatially mixed by turbulence, and so such vertical variations in the ionization fraction as we obtain here might be smoothed over in reality. Since resolving these issues is beyond the scope of our models, we simply always choose a solution with minimum $z_{\rm surf}$, which also appears to always correspond to a maximum $\bar{\alpha}$ at the given magnetic field strength.

3.2.4 *Summary*

As previously suggested by Desch & Turner (2015), the inner disc is primarily ionized through thermionic and potassium ion emission from dust grains. These processes counteract adsorption of free charges on to dust grains at the high temperatures present in the inner disc. We show that, for our fiducial parameters, the introduction of dust effects on disc ionization has very little effect on the location of the pressure maximum. This is because thermionic and ion emission become efficient above roughly the same threshold temperature as thermal ionization of potassium. Additionally, we show that non-thermal sources of ionization are unimportant for the fiducial parameters considered here. Thus, similar to our results in Section 3.1, while the inclusion of new physics in our model has not fundamentally changed the position of the pressure maximum, the physics setting its

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

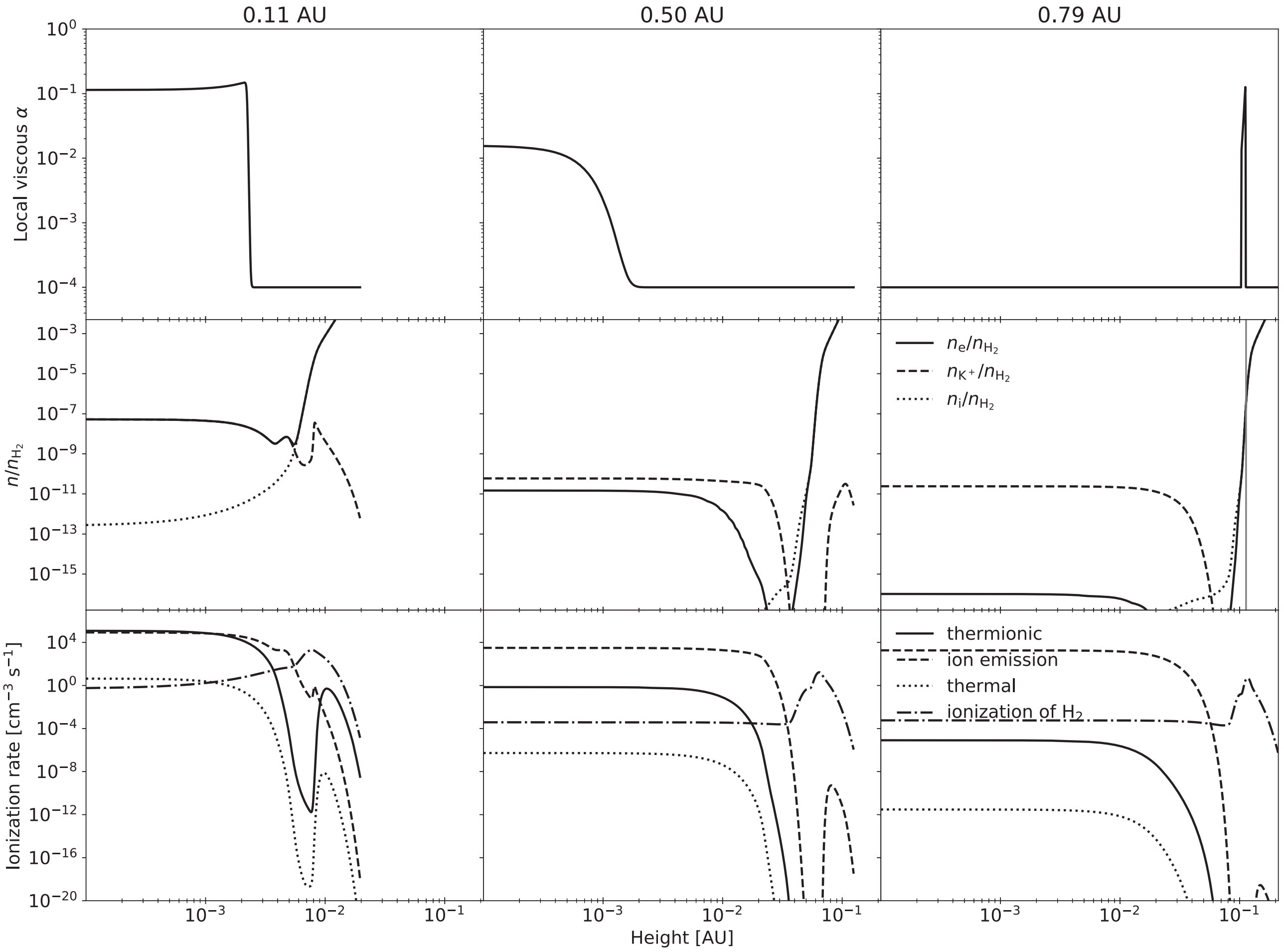


Figure 13. Local viscous parameter α (top), ionization fraction (middle; electrons: n_e/n_{H_2}, potassium ions: n_{K^+}/n_{H_2}, metal ions: n_i/n_{H_2}) and ionization rates (bottom; thermionic $\mathcal{R}_{\rm therm}$, ion emission $\mathcal{R}_{\rm K,evap}f_+$, thermal $k_2 n_{H_2} n_{K^0}$ and non-thermal ζn_{H_2}) at three different radii (as indicated in panel titles) for the model with all sources of thermal and non-thermal ionization. Non-thermal ionization produces an MRI-active region high above disc mid-plane at larger radii (see the top right-hand panel). The blue line in the right-hand panel on the second row indicates the upper boundary of this MRI-active region at the given radius, and the ionization fraction at which it occurs. Note that the metal ion fraction becomes unrealistically large only to the right of the blue line, i.e. only above the high-altitude MRI-active layer. See Section 3.2.2.

position is very different from that described in Mohanty et al. (2018).

4 DISCUSSION

4.1 Effects of dust

A key feature of the inner disc is that the MRI drives high viscosity in the innermost regions, close to the star, but becomes largely suppressed at larger orbital distances (in the so-called dead zone). This leads to the formation of a local gas pressure maximum that may play a key role in planet formation at short orbital distances (Chatterjee & Tan 2014). This decrease in viscosity is expected to arise because the innermost regions are hot enough (>1000 K) to thermally ionize potassium (coupling the gas to the magnetic field), but further out temperature and ionization levels decrease substantially (Gammie 1996). In a previous study, we showed that in a thermally ionized disc coupled self-consistently to an MRI viscosity, the inner edge of the dead zone lies at a few tenths of an AU (Mohanty et al. 2018).

One of the key differences between this work and that of Mohanty et al. (2018) is that here we also take into account the effects of dust on the disc's ionization state. Small dust grains present in the disc sweep up free electrons and ions from the gas, and these recombine quickly on the grain surfaces. In the bulk of the protoplanetary disc dust grains therefore efficiently lower the ionization fraction, decoupling the magnetic field from the gas and suppressing the MRI (Sano et al. 2000; Ilgner & Nelson 2006; Wardle 2007; Salmeron & Wardle 2008; Bai & Goodman 2009; Mohanty et al. 2013). However, in the inner regions of protoplanetary discs dust grains also act to increase the ionization levels, as at high temperatures they can also emit electrons and ions into the gas (Desch & Turner 2015). The balance between thermal ionization and these processes then determines how well ionized the inner disc is, and thus the extent of the high-viscosity region and the location of the dead zone inner edge.

The top panel of Fig. 10 shows that addition of dust only weakly affects the vertically averaged viscosity $\bar{\alpha}$ in the inner disc. At a given orbital radius, $\bar{\alpha}$ is even slightly higher than in the model with no dust, implying that thermionic and ion emission are important sources of ionization. In fact, as Desch & Turner (2015) showed,

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

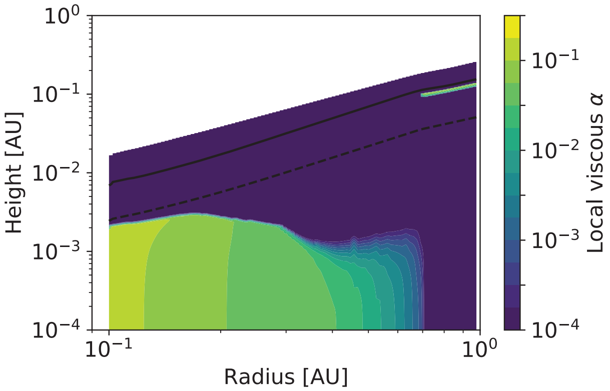


Figure 14. Local viscous parameter α as a function of location in the disc for the model with all sources of thermal and non-thermal ionization. In the innermost disc, thermionic and ion emission ionize the dense regions around the disc mid-plane, producing the high MRI-driven α there. At larger radii, the MRI is active in a low-density layer high above the disc mid-plane, dominated by non-thermal sources of ionization. The solid black line indicates the surface in the disc above which the ionization fraction $n_{\mathrm{e}}/n_{\mathrm{H_2}} > 10^{-4}$, i.e. above which the assumptions of our simple chemical network break; this surface is above the MRI-active region at all radii. See Section 3.2.2.

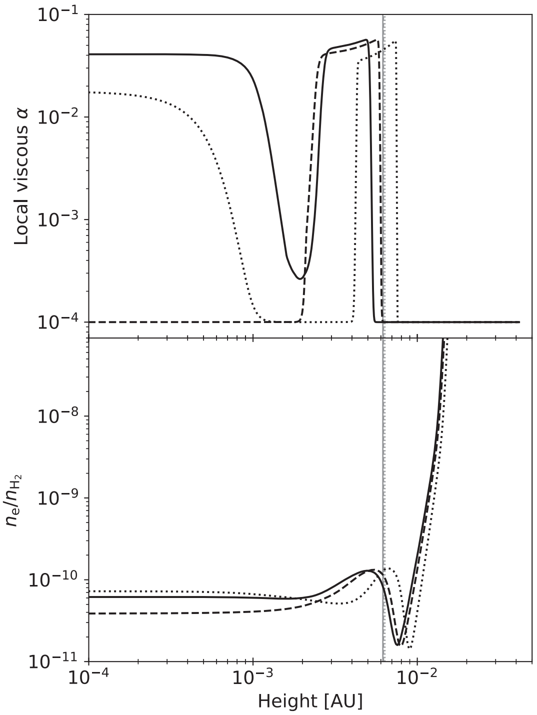


Figure 16. Degeneracy in the vertical disc structure at a fixed magnetic field strength in a model where the maximum dust grain size is $a_{\max} = 100$ μm: local viscous parameter α as a function of height (top), and the fractional electron number density ($n_{\mathrm{e}}/n_{\mathrm{H_2}}$; bottom) as functions of height in three different equilibrium solutions. The different solutions arise from vertical variations in the viscous α; the length scales of these variations are much smaller than the disc pressure scale height (shown for each solution by the vertical grey lines). The solutions shown here are at a radius of 0.22 AU, for field strength $\log B_z = 0.226$ and grazing angle $\phi = 0.06$. The vertically averaged viscous $\bar{\alpha}$ for the solid, dashed, and dotted lines are, respectively, 2×10^{-2}, 1.78×10^{-2}, and 1.28×10^{-2}. See Section 3.2.3.

thermionic and ion emission can become the main source of free electrons at high temperatures. For our disc model this can be seen in the bottom panel of Fig. 13, which shows that at the hot disc mid-plane thermionic and ion emission dominate over other sources of ionization. Clearly then, for the chosen parameters, the adsorption of charges on to grains is more than offset by the expulsion of charges from hot grain surfaces.

Similarity in the resulting disc structure in the models with and without dust grains can be explained by the similar dependence on temperature that thermal and thermionic/ion emission have (and which can also be deduced from the bottom panel of Fig. 13). As discussed by Desch & Turner (2015), thermal ionization of potassium becomes efficient at temperatures above ~1000 K in accordance with its ionization potential IP = 4.34 eV. The temperature at which thermionic and ion emission become important is determined by the work function W of the material out of which dust grains are made, and for silicates $W \sim 5$ eV. This alone would imply that thermionic emission becomes efficient at temperatures closer to 2000 K. However, above ~1000 K potassium-bearing minerals start evaporating from grain surfaces (Lodders 2003), and a fraction of potassium atoms leaves the grain surface as ions (since $W \sim$ IP; see equation 21). Dust grains then become negatively charged, which reduces the effective potential that electrons need to overcome for thermionic emission (the effective work function W_{eff}).

The above results could change significantly as dust grains grow or as they accumulate in the inner disc (as needed for the formation of solid planet cores). Dust adsorption of free charges, for example, becomes much less efficient for larger grains, since the total grain

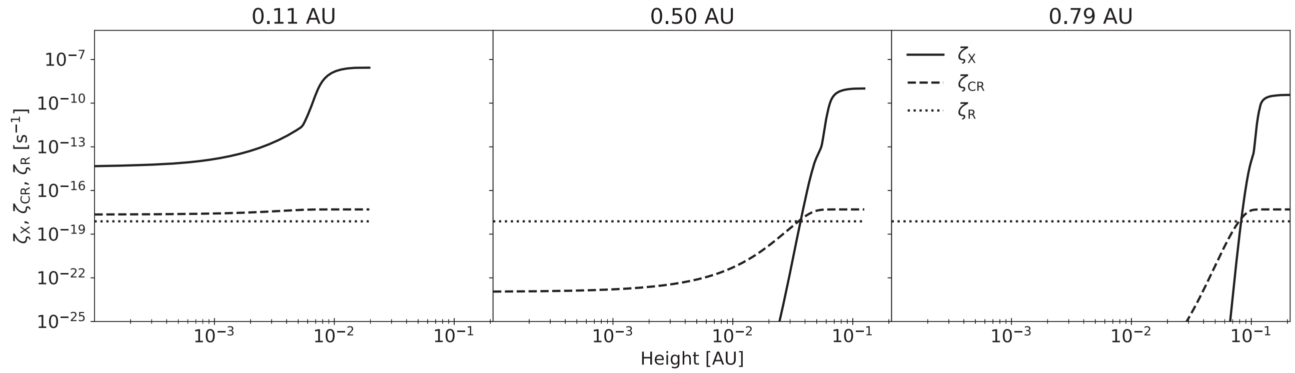


Figure 15. Ionization rates of molecular hydrogen due to stellar X-rays (ζ_{X}), cosmic rays (ζ_{CR}), and radionuclides (ζ_{R}) for the model with all sources of thermal and non-thermal ionization. Stellar X-rays are the dominant source of ionization in the disc upper layers. See Section 3.2.2.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

surface area decreases (Sano et al. 2000; Ilgner & Nelson 2006). We investigate how dust growth and varying dust-to-gas ratio affect the inner disc structure in our companion paper (Jankovic et al., in preparation).

4.2 Effect of stellar irradiation

We find that the absorbed flux of stellar irradiation is many times higher than the heat flux generated by accretion at any given radius, yet irradiation has a very small effect on the disc mid-plane temperature. Consequently, the ionization levels and the MRI-driven viscosity are similar in the models with and without stellar irradiation. Why does the irradiation have such a marginal effect?

Essentially, it is because the stellar irradiation heats the disc's optically thin regions, from which heat escapes easily. Accretion heat is generated deep in the disc, where the optical depth is much higher. In the absence of stellar irradiation, the mid-plane temperature in the optically thick disc is $\sigma_{\rm SB} T_{\rm mid}^4 \sim F_{\rm acc}\tau_{\rm mid}$ (Hubeny 1990, though in our work the mid-plane temperature is somewhat lower due to convection). If on top of a viscously heated layer of optical thickness $\tau_{\rm mid} \gg 1$ there is an irradiation-heated layer of optical thickness $\tau_{\rm upper}$, it follows from equations (8) and (9) that $\sigma_{\rm SB} T_{\rm mid}^4 \sim F_{\rm acc}\tau_{\rm mid} + F_{\rm irr}\tau_{\rm upper}$ (again, neglecting convection). Here $\tau_{\rm upper}$ is the optical depth of the disc to its own radiation down to a height at which the disc becomes optically thick to stellar irradiation. Then, if $\tau_{\rm mid}$ is sufficiently larger than $\tau_{\rm upper}$, the mid-plane temperature is determined by viscous dissipation.

Our results are consistent with those of D'Alessio et al. (1998), who also found that models with and without stellar irradiation yield roughly the same mid-plane temperatures in the optically thick inner disc. Similarly, Flock et al. (2019) considered 2D static radiation-hydrodynamics models of the inner disc heated by stellar irradiation only, but found that the mid-plane temperature (and the orbital radius of the dead zone inner edge) would increase appreciably if accretion heat were included, on the condition that the accretion heat is released near the optically thick mid-plane.

Note, additionally, that we have not considered the details of the inner disc edge or the dust sublimation line. If the inner rim of the disc is puffed-up, it would throw a shadow over a portion of the inner disc (Dullemond, Dominik & Natta 2001; Natta et al. 2001), further reducing the importance of stellar irradiation.

4.3 Convective instability in the inner disc

In Section 3.1, we showed that a large region of the inner disc is convectively unstable. We find this to be the case even when the opacities are constant, i.e. a superlinear growth of the opacity with temperature (Lin & Papaloizou 1980) is not needed. Here, the high temperature gradient is established because the heat is deposited deep within the optically thick disc. In the presented models most of the viscous dissipation happens near the mid-plane, where the MRI is active, but the same is also true for a vertically constant viscosity α. This result can also be confirmed analytically. We consider a simplified problem of radiative transfer in the optically thick limit, where the temperature is given by $\sigma T^4 = \frac{3}{4}\tau F(z_{\rm surf})$ (Hubeny 1990). Assuming a constant disc opacity, the equation of hydrostatic equilibrium can be re-written as $\frac{{\rm d}P}{{\rm d}\tau} = \frac{\Omega^2 z}{\kappa_{\rm R}}$. The temperature gradient is then given by

$$\nabla = \frac{{\rm d}\ln T}{{\rm d}\ln P} = \frac{\kappa_{\rm R} P}{4\Omega^2 z\tau}.$$

The appropriate upper boundary condition for this problem is the disc photosphere ($\tau = 2/3$), where the gas pressure is given by $P_{\rm surf} = \Omega^2 z_{\rm surf}\tau_{\rm surf}/\kappa_{\rm R}$ (assuming that the disc is vertically isothermal above the photosphere; Papaloizou & Terquem 1999). At the photosphere, given the chosen boundary condition, we have $\nabla = 1/4$. Near the mid-plane the optical depth is $\tau_{\rm MID} = \frac{1}{2}\kappa_{\rm R}\Sigma$, and we may estimate the mid-plane pressure as

$$P_{\rm MID} = \rho_{\rm MID} {c_{\rm s,MID}}^2 = \frac{\Sigma}{2H} {c_{\rm s,MID}}^2 = \frac{1}{2}\Omega^2 \Sigma H,$$

where the disc's scale height H is related to the mid-plane temperature through hydrostatic equilibrium. Substituting $\tau_{\rm MID}$ and $P_{\rm MID}$ into the expression for the temperature gradient, we have

$$\nabla_{\rm MID} = \frac{1}{4}\frac{H}{z},$$

implying that such a disc should become convectively unstable a bit below one scale height. However, we can further estimate the gradient ∇ near one scale height, by assuming that there $P_H \sim P_{\rm MID} e^{-1/2} \sim 0.6 P_{\rm MID}$, and $\tau_H \sim 0.3\tau_{\rm MID}$, as follows from vertically isothermal, Gaussian profiles of pressure and density. Thus, near $z = H$, we have

$$\nabla_H = \frac{1}{2}\frac{H}{z},$$

showing that the disc should become convectively unstable above one scale height.

Furthermore, in the convectively unstable regions we use a simple approximation that convection is efficient and the temperature gradient is isentropic. More detailed calculations would yield an answer in which the temperature gradient lies between the isentropic one and the gradient given by the radiative transport. As it turns out, the result would not differ much from what is obtained here. In the optically thick limit considered above, the difference in the temperature profile when the entire flux is transported by radiation and when the entire flux is transported by convection is a very weak function of optical depth, and remains small for rather large optical depths (Cassen 1993).

Limiting the temperature gradient is the only role of convection in our simple viscous model. In real discs, convection might interact with the MHD turbulence induced by the MRI. For example, in simulations of ideal MHD, Bodo et al. (2013) and Hirose et al. (2014) find that convection can increase the angular momentum transport driven by the MRI by increasing the magnetic field strength, although it appears that this is only the case when convection is particularly strong (Hirose 2015). Concurrently, it is found that the relationship between the induced stress and the magnetic field strength are not modified. The consequences for the non-ideal MHD regime, relevant in protoplanetary discs, are not clear. In our solutions the value of the vertically averaged MRI-driven viscosity would decrease with both a decrease and an increase in the magnetic field strength due to non-ideal effects, as discussed in Section 2.5.2. Convection itself is not expected to drive the angular momentum transport at a level comparable to the MRI (e.g. Lesur & Ogilvie 2010; Held & Latter 2018), and in any case it is not self-sustainable, i.e. it requires an additional source of heat near disc mid-plane to establish the high temperature gradient.

4.4 Energy transport by turbulent elements

In our model, we have not considered the possibility that the turbulent elements driving the angular momentum transport may also transport energy (Ruediger, Elstner & Tschaepe 1988). Such

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

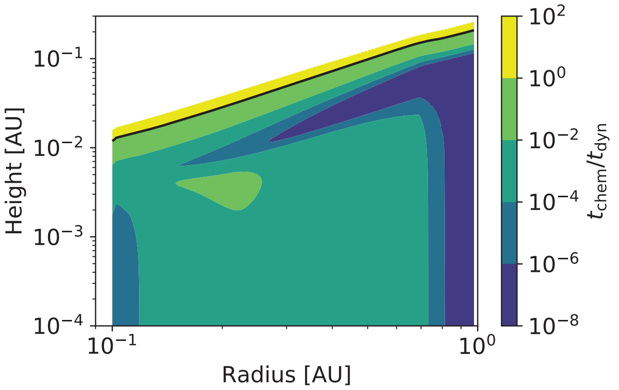


Figure 17. Ratio of the shortest recombination time-scale ($t_{\rm chem}$) to the dynamical time-scale ($t_{\rm dyn}$) as a function of the location in the disc. The solid line indicates where $t_{\rm chem}/t_{\rm dyn} = 1$. Everywhere below this line the chemical equilibrium time-scale is shorter than the dynamical time-scale, justifying our assumption of the strong-coupling regime. See Section 4.5.

turbulent energy transport flux would be analogous to convection, transporting energy down the entropy gradient, the difference being that turbulent elements may persist at sub-adiabatic temperature gradients (as they are driven by other instabilities) in which case they transport energy from cooler to hotter regions (Balbus 2000). We do not expect including this mode of energy transport in our calculations to appreciably change any of our results. D'Alessio et al. (1998) found that turbulent energy transport accounts for less than 20 per cent of the total energy flux at any given orbital radius. In our work, the convectively stable upper layers of the disc are MRI-dead, and thus the thermal diffusivity due to turbulence is likely very low. In the regions of the disc where the radiative flux alone would yield super-adiabatic temperature gradient, we already assume that convection efficiently establishes the adiabat. The proportion of turbulent energy flux could be higher than the convective energy flux in such regions, but the temperature would not change significantly. Hence, we have not included this effect in our analysis.

4.5 Ambipolar diffusion in the strong-coupling regime

The criterion for ambipolar diffusion to quench the MRI that we employ is valid in the strong-coupling regime (Bai & Stone 2011). Strong coupling requires that ionization equilibrium be achieved on a time-scale $t_{\rm chem}$ shorter than the dynamical time-scale $t_{\rm dyn} = 2\pi/\Omega$. Previously, Mohanty et al. (2018) reported that this condition is not fulfilled in most of the inner disc, as slow radiative recombinations make the chemical equilibrium time-scale long. However, even in the absence of dust, three-body recombination (recombinations through collisions with the abundant molecular hydrogen) are much faster than radiative recombinations, and adsorption on to grains is even faster (Desch & Turner 2015).

Since we solve directly for the equilibrium ionization state, we do not have access to the time-scale $t_{\rm chem}$. However, we can estimate it as $t_{\rm chem} = n_{\rm e}/\mathcal{R}$, where $\mathcal{R}$ is the fastest of the above three recombination rates (in general, but not always, that is adsorption on to dust grains). Fig. 17 shows, for our fully self-consistent model with our full chemical network, that $t_{\rm chem}/t_{\rm dyn} < 1$ everywhere except in the uppermost, lowest-density layers of the disc. We thus conclude that our use of the ambipolar diffusion criterion in the strong-coupling regime is justified.

5 CONCLUSIONS

We present a steady-state model of the inner protoplanetary disc which accretes viscously, primarily due to the MRI. In this model, the disc is heated by viscous dissipation and stellar irradiation, and cools radiatively and convectively. The disc is ionized by thermal ionization, thermionic and ion emission from dust grains and by stellar X-rays, cosmic rays and radionuclides, and we also account for adsorption of charges on to dust grains. The disc's structure (density, temperature), viscosity due to the MRI, opacity, and ionization state are calculated self-consistently everywhere in the disc (both as a function of radius and height). To the best of our knowledge, this is the first model that self-consistently couples all these processes to describe the structure of the inner regions of a steadily accreting protoplanetary disc.

We investigate how these various processes affect the structure of the inner disc and the extent to which the MRI can drive efficient accretion, i.e. the locations of the inner edge of the dead zone and the gas pressure maximum. For the fiducial parameters considered in this work (stellar parameters: $M_* = 1\,{\rm M}_\odot$, $R_* = 3\,{\rm R}_\odot$, $T_* = 4400$ K, X-ray luminosity $L_{\rm X} = 10^{-3.5}L_{\rm bol}$; disc parameters: gas accretion rate $\dot{M} = 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$, viscosity in the MRI-dead zone $\alpha_{\rm DZ} = 10^{-4}$; dust parameters: dust-to-gas mass ratio $f_{\rm dg} = 10^{-2}$, maximum grain size $a_{\rm max} = 1\ \mu$m), we find that:

(i) Inwards of the pressure maximum, the MRI is active only around the disc mid-plane. This differs from the predictions of vertically isothermal models, and is possibly important for the evolution of dust grains in the inner disc.

(ii) Since the inner disc is optically thick, stellar irradiation only weakly influences the mid-plane temperature, and thus also only weakly affects the location of the dead zone inner edge.

(iii) Most of the inner disc is convectively unstable, which we show is a property of any optically thick disc in which viscous heating occurs near the mid-plane. This motivates further work into a coupled MRI-convective instability in the limit of non-ideal MHD.

(iv) As suggested by Desch & Turner (2015), dust controls the ionization state of the inner disc, and thus the onset of the MRI. Thermal ionization plays a secondary role, as thermionic and ion emission from dust grains ionize the hot dense regions.

(v) High above disc mid-plane stellar X-rays produce an MRI-active layer. However, the X-rays barely change the overall viscosity at short orbital distances, or the location of the pressure maximum.

(vi) The pressure maximum resides at ∼0.7 AU for our fiducial parameters, roughly the same location as in our previous work (Mohanty et al. 2018). This is a consequence of the high optical depth in the inner disc, and the similarity between the ionization potential of potassium and the work function of the dust grains, rather than physics setting the pressure maximum location being similar.

These conclusions are drawn for a disc with a fiducial dust-to-gas ratio of 10^{-2} and small dust grains ($a_{\rm max} = 1\ \mu$m), which may be expected in the early stages of dust evolution in the disc. How these results depend on the model parameters, including dust grain size and dust-to-gas ratio, is explored in a companion paper (Jankovic et al., in preparation), where we also use our new inner disc model to speculate on possible formation pathways for close-in super-Earths. Finally, it is important to note that our model is based on a simplifying assumption that the inner disc is at steady-state. Only time-dependent simulations will show whether steady-state is indeed achieved and, if it is, whether this steady-state is stable.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

ACKNOWLEDGEMENTS

We thank the reviewer for helpful suggestions that improved the manuscript. We thank Richard Booth, Thomas Haworth, Zhaohuan Zhu, Steven Desch, Neal Turner, and Colin McNally for helpful discussions. MRJ acknowledges support from the President's PhD scholarship of the Imperial College London and the UK Science and Technology research Council (STFC) via the consolidated grant ST/S000623/1. JEO is supported by a Royal Society University Research Fellowship. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 853022, ERC-STG-2019 grant, PEVAP). JCT acknowledges support from NASA ATP grant Inside-Out Planet Formation (80NSSC19K0010).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Bai X.-N., 2011, ApJ, 739, 50
Bai X.-N., Goodman J., 2009, ApJ, 701, 737
Bai X.-N., Stone J. M., 2011, ApJ, 736, 144
Balbus S. A., 2000, ApJ, 534, 420
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Baraffe I., Homeier D., Allard F., Chabrier G., 2015, A&A, 577, A42
Bitsch B., Raymond S. N., Izidoro A., 2019, A&A, 624, A109
Bodo G., Cattaneo F., Mignone A., Rossi P., 2013, ApJ, 771, L23
Boley A. C., Ford E. B., 2013, preprint (arXiv:1306.0566)
Calvet N., Magris G. C., Patino A., D'Alessio P., 1992, Rev. Mex. Astron. Astrofis., 24, 27
Cassen P., 1993, in Lunar and Planetary Science Conference. Lunar and Planetary Science Conference, Vol. 24, Houston. p. 261
Chatterjee S., Tan J. C., 2014, ApJ, 780, 53
Chiang E. I., Goldreich P., 1997, ApJ, 490, 368
Chiang E. I., Joung M. K., Creech-Eakman M. J., Qi C., Kessler J. E., Blake G. A., van Dishoeck E. F., 2001, ApJ, 547, 1077
Cossou C., Raymond S. N., Hersant F., Pierens A., 2014, A&A, 569, A56
D'Alessio P., Cantö J., Calvet N., Lizano S., 1998, ApJ, 500, 411
D'Alessio P., Calvet N., Hartmann L., Lizano S., Cantó J., 1999, ApJ, 527, 893
Desch S. J., Turner N. J., 2015, ApJ, 811, 156
Draine B. T., 2003, ApJ, 598, 1017
Draine B. T., Sutin B., 1987, ApJ, 320, 803
Dressing C. D., Charbonneau D., 2013, ApJ, 767, 95
Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45
Dullemond C. P., 2002, A&A, 395, 853
Dullemond C. P., Dominik C., Natta A., 2001, ApJ, 560, 957
Dullemond C. P., Juhasz A., Pohl A., Sereshti F., Shetty R., Peters T., Commercon B., Flock M., 2012, RADMC-3D: A multi-purpose radiative transfer tool, record ascl:1202.015
Ercolano B., Glassgold A. E., 2013, MNRAS, 436, 3446
Estrada P. R., Cuzzi J. N., Morgan D. A., 2016, ApJ, 818, 200
Flock M., Turner N. J., Mulders G. D., Hasegawa Y., Nelson R. P., Bitsch B., 2019, A&A, 630, A147
Frank J., King A., Raine D. J., 2002, Accretion Power in Astrophysics, Third Edition, Cambridge Univ. Press, Cambridge, UK
Fressin F., et al., 2013, ApJ, 766, 81
Gammie C. F., 1996, ApJ, 457, 355
Garaud P., Lin D. N. C., 2007, ApJ, 654, 606
Glassgold A. E., Najita J., Igea J., 1997, ApJ, 480, 344
Hansen B. M. S., Murray N., 2013, ApJ, 775, 53
Hartmann L., Calvet N., Gullbring E., D'Alessio P., 1998, ApJ, 495, 385
Held L. E., Latter H. N., 2018, MNRAS, 480, 4797
Hirose S., 2015, MNRAS, 448, 3105
Hirose S., Blaes O., Krolik J. H., Coleman M. S. B., Sano T., 2014, ApJ, 787, 1
Hsu D. C., Ford E. B., Ragozzine D., Ashby K., 2019, AJ, 158, 109
Hu X., Tan J. C., Zhu Z., Chatterjee S., Birnstiel T., Youdin A. N., Mohanty S., 2018, ApJ, 857, 20
Hubeny I., 1990, ApJ, 351, 632
Igea J., Glassgold A. E., 1999, ApJ, 518, 848
Ilgner M., Nelson R. P., 2006, A&A, 445, 205
Izidoro A., Ogihara M., Raymond S. N., Morbidelli A., Pierens A., Bitsch B., Cossou C., Hersant F., 2017, MNRAS, 470, 1750
Izidoro A., Bitsch B., Raymond S. N., Johansen A., Morbidelli A., Lambrechts M., Jacobson S. A., 2019, preprint (arXiv:1902.08772)
Jankovic M. R., Owen J. E., Mohanty S., 2019, MNRAS, 484, 2296
Keith S. L., Wardle M., 2014, MNRAS, 440, 89
Lesur G., Ogilvie G. I., 2010, MNRAS, 404, L64
Lin D. N. C., Papaloizou J., 1980, MNRAS, 191, 37
Lodders K., 2003, ApJ, 591, 1220
Malygin M. G., Kuiper R., Klahr H., Dullemond C. P., Henning T., 2014, A&A, 568, A91
Manara C. F., et al., 2016, A&A, 591, L3
Manara C. F., et al., 2017, A&A, 604, A127
Mathis J. S., Rumpl W., Nordsieck K. H., 1977, ApJ, 217, 425
McNeil D. S., Nelson R. P., 2010, MNRAS, 401, 1691
Mohanty S., Ercolano B., Turner N. J., 2013, ApJ, 764, 65
Mohanty S., Jankovic M. R., Tan J. C., Owen J. E., 2018, ApJ, 861, 144
Mulders G. D., Pascucci I., Apai D., Ciesla F. J., 2018, AJ, 156, 24
Natta A., Prusti T., Neri R., Wooden D., Grinin V. P., Mannings V., 2001, A&A, 371, 186
Ogihara M., Ida S., 2009, ApJ, 699, 824
Owen J. E., Wu Y., 2017, ApJ, 847, 29
Papaloizou J. C. B., Terquem C., 1999, ApJ, 521, 823
Pollack J. B., Hollenbach D., Beckwith S., Simonelli D. P., Roush T., Fong W., 1994, ApJ, 421, 615
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, Numerical recipes in C+ + : the art of scientific computing, Cambridge Univ. Press, Cambridge, UK
Rafikov R. R., 2007, ApJ, 662, 642
Ruediger G., Elstner D., Tschaepe R., 1988, Acta Astron., 38, 299
Salmeron R., Wardle M., 2008, MNRAS, 388, 1223
Sano T., Stone J. M., 2002, ApJ, 577, 534
Sano T., Miyama S. M., Umebayashi T., Nakano T., 2000, ApJ, 543, 486
Sano T., Inutsuka S.-i., Turner N. J., Stone J. M., 2004, ApJ, 605, 321
Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33
Shu F. H., 1992, The physics of astrophysics. Volume II: Gas dynamics, University Science Books, Mill Valley CA, USA
Suzuki T. K., Ogihara M., Morbidelli A., Crida A., Guillot T., 2016, A&A, 596, A74
Terquem C. E. J. M. L. J., 2008, ApJ, 689, 532
Terquem C., Papaloizou J. C. B., 2007, ApJ, 654, 1110
Umebayashi T., Nakano T., 1981, PASJ, 33, 617
Umebayashi T., Nakano T., 2009, ApJ, 690, 69
Van Eylen V., Agentoft C., Lundkvist M. S., Kjeldsen H., Owen J. E., Fulton B. J., Petigura E., Snellen I., 2018, MNRAS, 479, 4786
Wardle M., 1999, MNRAS, 307, 849
Wardle M., 2007, Ap&SS, 311, 35
Wright N. J., Drake J. J., Mamajek E. E., Henry G. W., 2011, ApJ, 743, 48
Wu Y., 2019, ApJ, 874, 91
Zink J. K., Christiansen J. L., Hansen B. M. S., 2019, MNRAS, 483, 4479

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/504/1/280/6207952 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **495,** 451–459 (2020)
doi:10.1093/mnras/staa1167
Advance Access publication 2020 April 30

A three-dimensional view of Gomez's hamburger

Richard Teague ,[1★] Marija R. Jankovic ,[2] Thomas J. Haworth ,[3] Chunhua Qi[1] and John D. Ilee [4]

[1] *Center for Astrophysics | Harvard, Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[2] *Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*
[3] *Astronomy Unit, School of Physics and Astronomy, Queen Mary University of London, London E1 4NS, UK*
[4] *School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK*

Accepted 2020 April 23. Received 2020 April 23; in original form 2020 March 4

ABSTRACT

Unravelling the three-dimensional physical structure, the temperature and density distribution, of protoplanetary discs is an essential step if we are to confront simulations of embedded planets or dynamical instabilities. In this paper, we focus on submillimeter array observations of the edge-on source, Gomez's Hamburger, believed to host an overdensity hypothesized to be a product of gravitational instability in the disc, GoHam b. We demonstrate that, by leveraging the well-characterized rotation of a Keplerian disc to deproject observations of molecular lines in position-position-velocity space into disc-centric coordinates, we are able to map out the emission distribution in the $(r,\ z)$ plane and $(x,\ |y|,\ z)$ space. We show that ^{12}CO traces an elevated layer of $z\,/\,r \sim 0.3$, while ^{13}CO traces deeper in the disc at $z\,/\,r \lesssim 0.2$. We identify an azimuthal asymmetry in the deprojected ^{13}CO emission coincident with GoHam b at a polar angle of $\approx 30^{\circ}$. At the spatial resolution of $\sim$1.5 arcsec, GoHam b is spatially unresolved, with an upper limit to its radius of $<$190 au.

Key words: accretion, accretion discs – circumstellar matter – stars: formation.

1 INTRODUCTION

High angular resolution observations of the dust in protoplanetary discs, both at mm and near-infrared (NIR) wavelengths, have shown a stunning variety of features such as concentric rings and spirals (Andrews et al. 2018; Avenhaus et al. 2018). These structures hint at highly dynamic environments where the dust distributions are sculpted by changes in the gas pressure distribution. The precise cause for the perturbations in the gas is hard to constrain, with multiple scenarios possible, including embedded planets (e.g. Dipierro et al. 2015b; Fedele et al. 2018; Keppler et al. 2018; Zhang et al. 2018), and (magneto-)hydrodynamical instabilities (Flock et al. 2015) or gravitational instabilities (Dong et al. 2015; Dipierro et al. 2015a; Hall et al. 2016; Meru et al. 2017). Differentiating between these scenarios requires an intimate knowledge of the underlying gas structure and, in particular, how that structure changes from the mid-plane, as traced by the mm continuum emission, to the disc atmosphere, populated by the small sub-μm grains which efficiently scatter stellar NIR radiation.

This is routinely attempted by using observations of different molecular species believed to trace distinct vertical regions in the disc. This is due to a combination of both optical depth effects and changes in physical conditions with height in the disc, which make certain regions more conducive to the formation of particular species. However, it is only with high spatial resolution data that we are beginning to be able to directly measure the height at which molecular emission arises (de Gregorio-Monsalvo et al. 2013; Rosenfeld et al. 2013; Pinte et al. 2018), verifying predictions from chemical models.

A more direct approach is the observation of high-inclination discs where the emission distribution can be mapped directly. Unlike continuum emission that suffers from extremely high optical depths due to the long path lengths for edge-on discs (Guilloteau et al. 2016; Louvet et al. 2018), the rotation of the disc limits the optical depth of molecular emission in a given spectral channel. This allowed Dutrey et al. (2017) to map the ^{12}CO $J = 2-1$ and CS $J = 5-4$ emission distribution in the $(r_{\rm disc},\ z_{\rm disc})$ plane, calling this a tomographically reconstructed distribution (TRD), for the edge-on disc colloquially known as the Flying Saucer (2MASS J16281370-2431391).

In addition to allowing access to the $(r_{\rm disc},\ z_{\rm disc})$ plane, Dent et al. (2014), but see also Matrà et al. (2017) and Cataldi et al. (2018), demonstrated how similar techniques can be used to deproject a cut across the disc major axis into the $(x_{\rm disc},\ |y_{\rm disc}|)$ plane. The absolute value of $y_{\rm disc}$ arises because it is impossible to distinguish between the near and far side of the disc ($\pm y_{\rm disc}$) from their projected line-of-sight velocities alone. Using this technique, the authors were able to extract the azimuthal emission distribution along the line of sight

---

★ E-mail: richard.d.teague@cfa.harvard.edu

© 2020 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

revealing a clump of CO emission. Application of this technique to a vertically extended source enables the extraction of a full three-dimensional (3D) emission distribution.

In this paper, we apply these techniques to submillimeter array (SMA) observations of Gomez's Hamburger, an edge-on circumstellar disc. In Section 2, we describe the observations and data reduction. In Section 3, we provide an overview of the deprojection techniques used and their application to Gomez's Hamburger. A discussion of these results and a summary conclude the paper in Sections 4 and 5, respectively.

2 SUMMARY OF OBSERVATIONS

At an inclination of $i \approx 86^\circ$ and a distance of 250 ± 50 pc, Gomez's Hamburger (GoHam, IRAS 18059-3211) offers a rare opportunity to study the chemical and physical structure of an edge-on disc. Although originally classified as an evolved A0 star surrounded by a planetary nebula, follow-up observations using the SMA showed CO emission in the distinct pattern of Keplerian rotation about GoHam. These and subsequent observations firmly establish GoHam as a $2.5 \pm 0.5\ \mathrm{M}_\odot$ A-type star at a distance of 250 ± 50 pc surrounded by a massive, $M_{\rm disc} \sim 0.2 M_{\rm sun}$, circumstellar disc (Bujarrabal, Young & Fong 2008; Wood et al. 2008; Bujarrabal, Young & Castro-Carrizo 2009; De Beck et al. 2010). This identification is further justified with the exquisite observations from the NICMOS instrument on the *Hubble Space Telescope*, which show the distinct flared geometry associated with protoplanetary discs (Bujarrabal et al. 2009).

2.1 Data reduction

The data were obtained from the SMA archive[1] and calibrated using the `MIR` software.[2] The interested reader is referred to the original papers, Bujarrabal et al. (2008, 2009), for a thorough overview of the calibration process. After calibration, the data were exported to `CASA v5.6.0` where two rounds of self-calibration were performed on the continuum, with phase solutions applied to the spectral line windows. The phase centre was adjusted so that the centre of the continuum was in the image centre.

After experimenting with various imaging properties, both the ^{12}CO and ^{13}CO transitions were imaged at their native channel spacing of 203 kHz (264 $\mathrm{m\,s^{-1}}$) with a Briggs weighting scheme and a robust parameter of 0.5. This resulted in synthesized beams of 1.53×1.11 arcsec^2 at $0\overset{\circ}{.}4$ for ^{12}CO and 1.57×1.15 arcsec^2 at $2\overset{\circ}{.}0$ for ^{13}CO. The measured rms in a line free channel was found to be 132 and 120 mJy beam^{-1} for the ^{12}CO and ^{13}CO. Channel maps were created both at the native channel spacing and downsampled by a factor of 2 to increase the signal-to-noise ratio.

Moment maps were also generated for the data using the PYTHON package `bettermoments` (Teague & Foreman-Mackey 2018). Integrated intensity maps were created using a threshold of 2σ for both molecules, while the rotation map used the quadratic method described in (Teague & Foreman-Mackey 2018) without the need for any σ-clipping. Rather than using the intensity weighted velocity dispersion (second moment), which is typically very noisy and incurs a large uncertainty (Teague 2019a), we use the 'effective linewidth' implemented in `bettermoments`. This calculates an effective line width using $\Delta V_{\rm eff} = M_0 / \sqrt{\pi}\, F_\nu^{\rm max}$, where M_0 is the integrated intensity and $F_\nu^{\rm max}$ is the line peak. For a Gaussian line profile, this returns the true Doppler width of the line. Both transitions show a peak at the disc centre, gradually decreasing in the outer disc. However, at this spatial resolution, the line profile is dominated by systematic broadening effects from the imaging.

2.2 Observational results

Using the 2D-Gaussian fitting tool `IMFIT` in `CASA` the integrated flux of the 1.3-mm continuum was found to be 293 ± 4 mJy, consistent with Bujarrabal et al. (2008). Integrating over an elliptical region with a major axis of 14 arcsec, a minor axis of 7 arcsec and a position angle of 175°, and clipping all values below 2σ, the ^{12}CO integrated flux was found to be 37.2 $\mathrm{Jy\,km\,s^{-1}}$. For the ^{13}CO, integrating over an elliptical mask with a major axis of 12 arcsec and a minor axis of 4.2 arcsec a position angle of 175°, again clipping all values below 2σ, resulted in an integrated flux of 16.5 $\mathrm{Jy\,km\,s^{-1}}$.

A summary of the moment maps alongside the continuum image is shown in Fig. 1. The continuum is clearly detected and considerably smaller in extent than the gas component. Assuming a source distance of 250 pc (Bujarrabal et al. 2008), the gaseous disc extends 1500 au in radius. For both transitions, the southern side of the disc is observed to be considerably brighter than the northern side, in addition to a slight north–south asymmetry in the continuum emission. In addition, the east–west asymmetry in the ^{12}CO integrated intensity suggests that the eastern side of the disc is tilted towards the observer.

Fig. 2 shows the channel maps, downsampled in velocity by a factor of 2, for the ^{12}CO emission, top panel, and the ^{13}CO emission, bottom panel. Both lines show the distinct 'butterfly' emission morphology characteristic of a rotating disc. The ^{12}CO emission is more extended, both in the radial and vertical directions, as would be expected given its larger abundance. The ^{12}CO emission also splits into two lobes, most clearly seen in the channels at 1.62 and 3.73 $\mathrm{km\,s^{-1}}$, due to the elevated emission surface, while the ^{13}CO appears more centrally peaked.

To find the systemic velocity of the disc, we fit the rotation maps, maps of the line centre, v_0, shown in Fig. 1, using the PYTHON package `eddy` (Teague 2019b). At these large inclinations, vertically extended emission, as expected for ^{12}CO and to a lesser extent, ^{13}CO, will result in rotation maps which are extended along the minor axis (see fig. 3 a from Dutrey et al. 2017), resulting in a v_0 distribution which deviates significantly from an inclined 2D disc model. Despite this, the rotation profile will be symmetric about the systemic velocity such that the inferred $v_{\rm LSR}$ from a fit of an inclined 2D disc will provide a good estimate of the true systemic velocity. We fix the source distance to 250 pc, and allow the source centre, inclination, position angle, stellar mass, and systemic velocity to vary. Using 64 walkers, which take 10 000 burn-in steps and an addition 5000 steps to estimate the posterior distributions, we find Gaussian-like posteriors for $v_{\rm LSR}$ for both transitions: $v_{\rm LSR}(^{12}{\rm CO}) = 2793 \pm 34\ \mathrm{m\,s^{-1}}$ and $v_{\rm LSR}(^{13}{\rm CO}) = 2778 \pm 38\ \mathrm{m\,s^{-1}}$. These uncertainties represent the statistical uncertainties, which do not consider the applicability of the model and so the true uncertainties are likely larger.

3 DEPROJECTION TO DISC-CENTRIC COORDINATES

If the velocity structure of the source is known, it is possible to deproject observations of an edge-on disc in position–position–velocity (PPV) space, $(x_{\rm sky},\ y_{\rm sky},\ v_0)$, into 3D disc-centred coordi-

---

[1] https://www.cfa.harvard.edu/cgi-bin/sma/smaarch.pl

[2] https://www.cfa.harvard.edu/~cqi/mircook.html

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

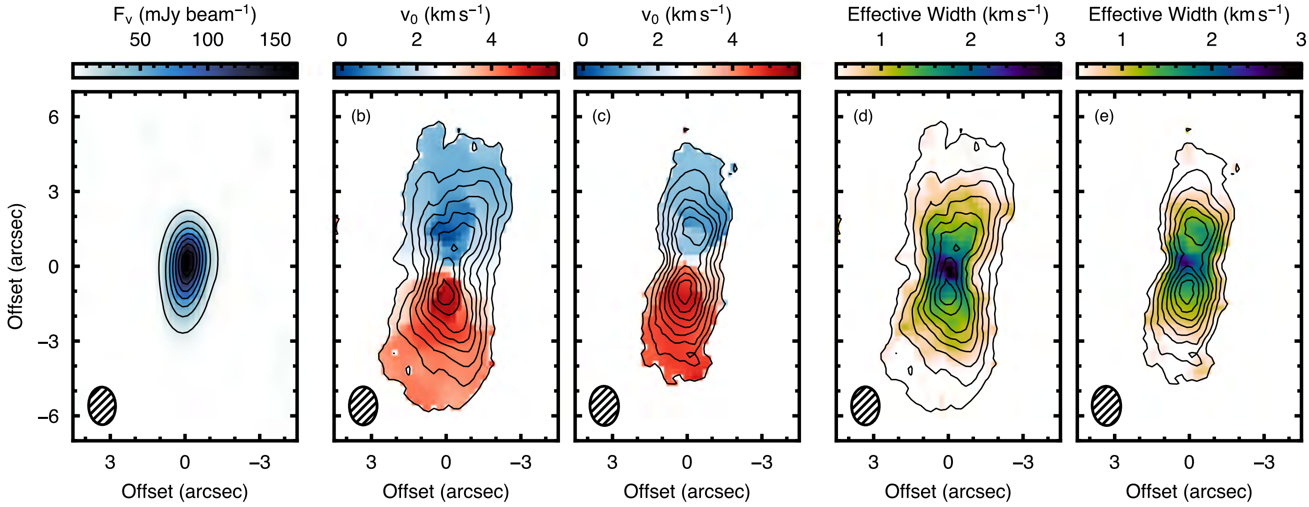


Figure 1. Summary of the observations. Panel (a) shows the 1.3-mm continuum emission. The black contours show steps of 20σ starting at 10σ, where $\sigma = 1.13$ mJy beam^{-1}. Central panels (b) and (c) show the rotation maps for the ^{12}CO and ^{13}CO emission, respectively. The black contours show the integrated intensities for the two lines in steps of 10 per cent of their peak values, 4.90 and 3.51 Jy beam^{-1} km s^{-1}. The right most panels, (d) and (e), show the effective width of the line, with the integrated intensity contours overlaid. The synthesized beams are shown in the bottom left-hand side of each panel.

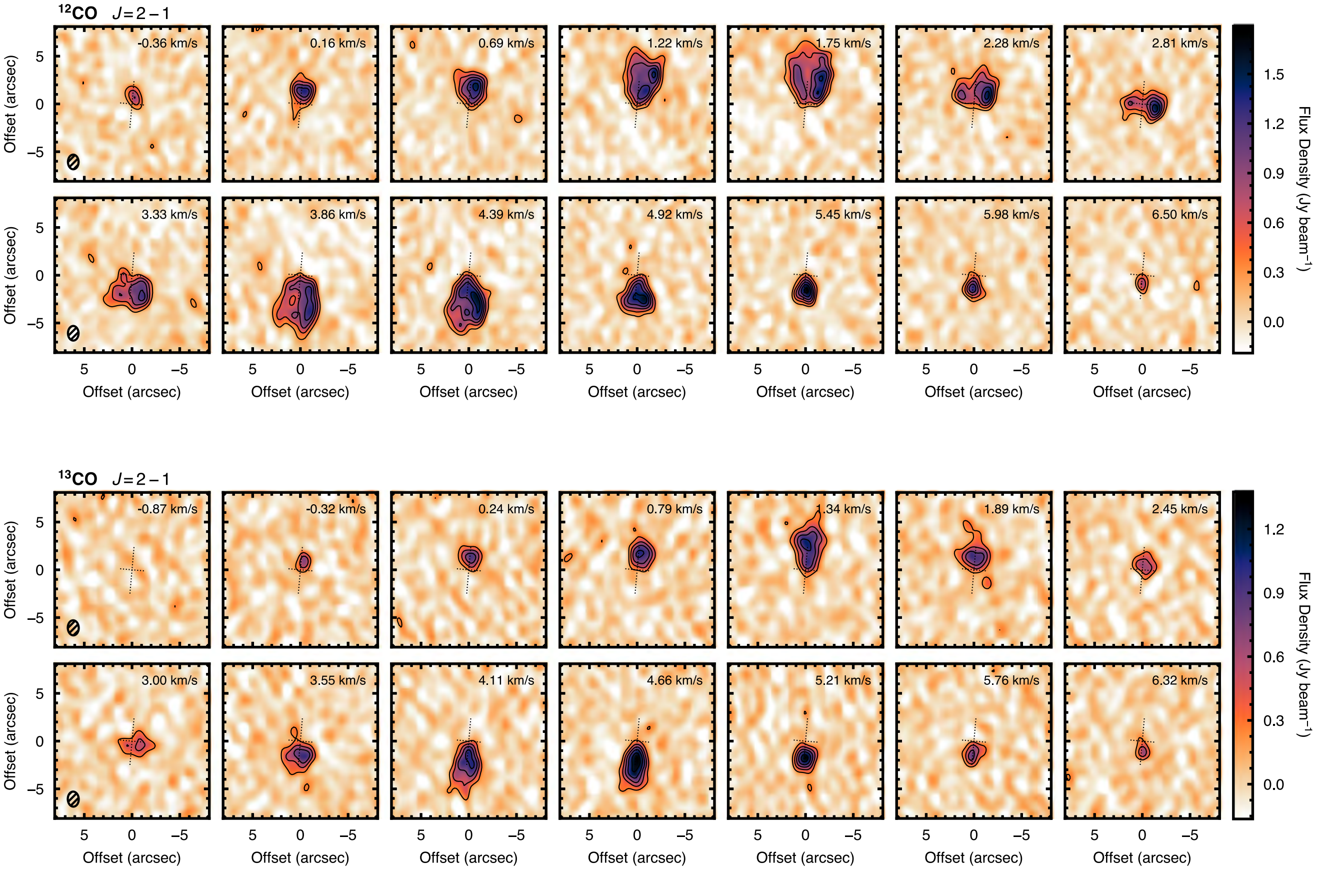


Figure 2. Top panel: Channel maps of the ^{12}CO emission, downsampled to 528 m s^{-1} channel spacing for presentation. Solid lines show contours starting at 3σ and increasing in steps of 3σ, where $\sigma = 87$ mJy beam^{-1}. The synthesized beams are shown in the left rows. The dotted lines show the orientation of the major and minor axes of the disc. The central velocity of the channel is shown in the top right-hand side of each panel. Bottom panel: as above, but for ^{13}CO emission. The solid lines are contours of 3σ ,where $\sigma = 75$ mJy beam^{-1}. Note the substantial increase in brightness in the southern side of the disc.

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

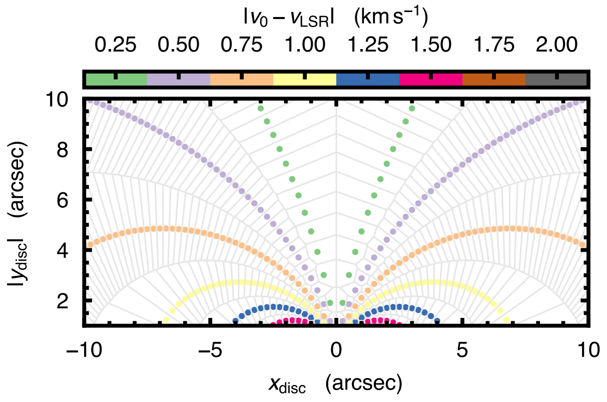


Figure 3. The deprojection of pixels assuming a 0.25 arcsec pixel size and a 250 m s^{-1} velocity spacing from equation (5). The velocity resolution sets the number of 'spokes' in the deprojection, while the pixel scaling sets the sampling along each spoke. The central channels close to the systemic velocities, $v_0 \sim v_{\rm LSR}$, populate regions, where $x_{\rm disc}$ is small. Thus, even with high-velocity resolution, it is hard to get an accurate deprojection for these regions (see, e.g. Dent et al. 2014; Matrà et al. 2017; Cataldi et al. 2018).

nates, $(x_{\rm disc},\ y_{\rm disc},\ z_{\rm disc})$. Both Dutrey et al. (2017) and Matrà et al. (2017) discuss similar deprojections, the former into an azimuthally averaged $(r_{\rm disc},\ z_{\rm disc})$ plane, and the latter into the $(x_{\rm disc},\ |y_{\rm disc}|)$ plane for a cut at a constant $z_{\rm disc}$ through the disc. In this section, we discuss both deprojections and include a correction due to changes in the rotation velocity as a function of height rather than assuming cylindrical rotation.

At any given voxel (a pixel in PPV space), the projected line of sight velocity, v_0, is given by

$$v_0 = v_\phi \,\cos\phi \,\sin i + v_{\rm LSR}, \tag{1}$$

where v_ϕ is the rotation velocity, ϕ is the azimuthal angle (not to be confused with the polar angle that is measured in the sky plane rather than the disc plane), i is the disc inclination, and $v_{\rm LSR}$ is the systemic velocity. For Keplerian rotation we know that

$$v_\phi(r_{\rm disc},\ z_{\rm disc}) = \sqrt{\frac{GM_{\rm star} r_{\rm disc}^2}{(r_{\rm disc}^2 + z_{\rm disc}^2)^{3/2}}}, \tag{2}$$

where $r_{\rm disc}$ and $z_{\rm disc}$ are the cylindrical radius and height in the disc, respectively, dropping the disc subscript for brevity. Substituting this into equation (1) and noting that for and edge-on disc, such that $i = 90^\circ$, $x_{\rm sky} = r_{\rm disc}\cos\phi$, and $z_{\rm disc} = y_{\rm sky}$, then we find

$$v_0 - v_{\rm LSR} = \sqrt{\frac{GM_{\rm star} x_{\rm sky}^2}{(r_{\rm disc}^2 + y_{\rm sky}^2)^{3/2}}}. \tag{3}$$

As both $x_{\rm sky}$ and $y_{\rm sky}$ are readily measured in the image plane, we can rearrange for $r_{\rm disc}$ giving

$$r_{\rm disc} = \sqrt{\left(\frac{GM_{\rm star} x_{\rm sky}^2}{(v_0 - v_{\rm LSR})^2}\right)^{2/3} - y_{\rm sky}^2}. \tag{4}$$

If cylindrical rotation is assumed, i.e. that there is no z dependence in v_ϕ in equation (2), the $y_{\rm sky}^2$ correction term vanishes, recovering the result from Dutrey et al. (2017).

Noting that $r_{\rm disc} = \sqrt{x_{\rm disc}^2 + y_{\rm disc}^2}$, where $y_{\rm disc}$ is the line-of-sight axis, we can additionally infer something about the line-of-sight distance of the emission,

$$|y_{\rm disc}| = \sqrt{\left(\frac{GM_{\rm star} x_{\rm sky}^2}{(v_0 - v_{\rm LSR})^2}\right)^{2/3} - y_{\rm sky}^2 - x_{\rm sky}^2}, \tag{5}$$

as used in Matrà et al. (2017). However, as there is a degeneracy in the side of the disc the emission arises, $\pm y$, this recovers an average of both sides of the disc. Again, if the cylindrical rotation is assumed, the $y_{\rm sky}^2$ correction term vanishes in equation (5). Fig. 3 shows how pixels would be deprojected into the $(x_{\rm disc},\ |y_{\rm disc}|)$ plane. It illustrates that the velocity resolution sets the 'azimuthal' sampling, i.e. how many spokes there are, while the pixel size (or spatial resolution) will set sampling along these spokes. As, such, both spatial and spectral resolution are required for an accurate deprojection of the data.

In addition to the transformation of the coordinates, it is essential to include the Jacobian such that integrated flux in the deprojected maps is conserved. Following appendix C of Cataldi et al. (2018), we find that the Jacobian for the transformation from $(x_{\rm sky},\ y_{\rm sky},\ v)$ to $(r_{\rm disc},\ z_{\rm disc},\ v)$, where $z_{\rm disc} = y_{\rm sky}$, is given by

$$J_{r_{\rm disc},\, z_{\rm disc}} = \frac{3 x_{\rm sky}\, r_{\rm disc}}{r_{\rm disc}^2 + z_{\rm disc}^2}, \tag{6}$$

which is a dimensionless transformation, meaning that the units are the same as in a channel map (e.g. Dutrey et al. 2017). Similarly, to transform the sky-plane coordinates into $(x_{\rm disc},\ |y_{\rm disc}|)$ coordinates, we find,

$$J_{x_{\rm disc},\, y_{\rm disc}} = \frac{3}{2}\sqrt{GM_{\rm star}}\, x_{\rm disc}\, y_{\rm disc}\, \left(x_{\rm disc}^2 + y_{\rm disc}^2 + z_{\rm disc}^2\right)^{-7/4}, \tag{7}$$

similar to equation (10) in Cataldi et al. (2018), but including an addition $z_{\rm disc}^2$ term owing to our definition of v_ϕ. This Jacobian has units of Hz, owing to the change from PPV space to position–position–position space. After correcting for the change in velocity with an addition $d\nu/c$ term, where d is the source distance and ν the frequency of the line, we have the final units of W m^{-2} sr^{-1}, i.e. a radiance along the z-axis.

We note that these derivations assume that the disc is completely edge-on and in Keplerian rotation. Dutrey et al. (2017) showed how changes in the inclination can affect the deprojection. The authors found that for only moderate deviations from edge-on, i.e. $i \gtrsim 80^\circ$, the TRD (see also Section 3.1) provided a good representation of the underlying physical structure. One half of the disc, either where $z > 0$ or $z < 0$, would be brighter, with this brighter half corresponding to the side of the disc, which is closer to the observer. In addition, the vertical extent of the emitting layer would broaden in the z-direction, before eventually splitting into two distinct arms when $i \lesssim 80^\circ$ and the near and far sides of the disc are spatially resolved.

3.1 TRD

As the disc is expected to be highly inclinated, $i \sim 85^\circ$ (Bujarrabal et al. 2008, 2009), we use the deprojection techniques described in Section 3 to explore the 3D structure of the disc, starting with the TRD as used for the Flying Saucer in Dutrey et al. (2017). We take the geometrical properties inferred from forward modelling a full 3D model presented in Bujarrabal et al. (2008, 2009), which assumed Keplerian rotation around a 2 M$_\odot$ central star and a disc inclined at 85°, observed at a position angle of 175°.

Using equation (4), each pixel is deprojected into $(r_{\rm disc},\ z_{\rm disc})$ space, before being binned into bins equal in size to the pixel. In

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

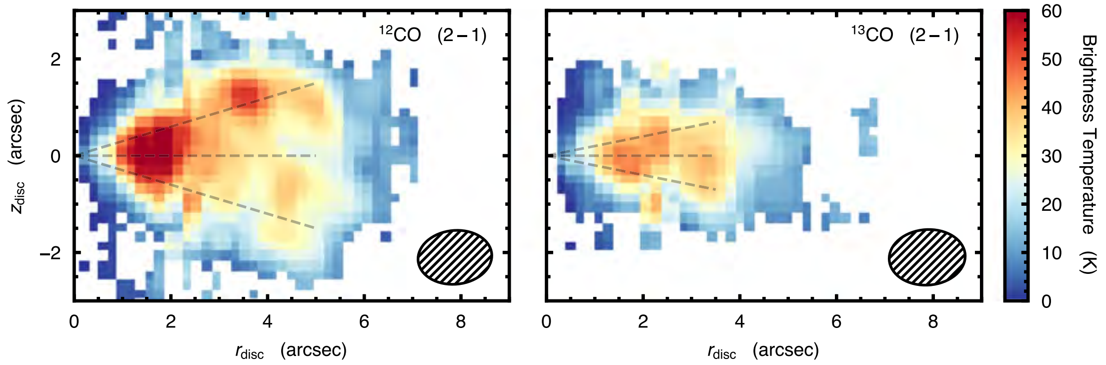


Figure 4. TRD using the method in Dutrey et al. (2017) for ^{12}CO, left-hand panel, and ^{13}CO, right-hand panel. The dashed lines show $z\,/\,r = 0.3$ to the left-hand panel and $z\,/\,r = 0.2$ to the right-hand panel. Note that asymmetry about the $z = 0$ line due to the deviation from a completely edge-on disc, as discussed in Dutrey et al. (2017).

each bin, we take the maximum value, equivalent to collapsing an image cube along the spectral axis by taking the maximum value along each pixel, e.g. a moment 8 map in `CASA`.

Fig. 4 shows the TRD for ^{12}CO, left-hand panel, and ^{13}CO, right-hand panel, taking the peak brightness temperature in each bin. Immediately, we see that ^{12}CO traces an elevated region of $z\,/\,r \sim 0.3$, while ^{13}CO appears to trace a region closer to the mid-plane, confined to $z\,/\,r \lesssim 0.2$. The drop off of signal within the inner 1 arcsec is due to convolution effects, as described in Dutrey et al. (2017). The asymmetry above the mid-plane is due to the deviation from a directly edge-on disc with a similar effect seen in the Flying Saucer, where the level of difference between the positive and negative values is consistent with the $i \approx 85^\circ$ inclination measured for the source. We note a bright point-source at $x_{\rm disc} \sim 3.5$ arcsec and $z_{\rm disc} \sim 1.5$ arcsec, likely associated with the peak in the ^{12}CO zeroth moment map in the north-west (see channel 1.22 km s^{-2} in Fig. 5). Higher resolution data are required to accurately disentangle this feature.

3.2 Line-of-sight deprojection

In Bujarrabal et al. (2009), it was argued that there was an enhancement of ^{13}CO at an offset of $r \approx 1$ arcsec. To explore whether this can be observed with the above techniques, we follow Matrà et al. (2017) and use equation (5) to deproject cuts along the major axis of the disc into the $(x_{\rm disc},\ |y_{\rm disc}|)$ plane.

The disc was split into six equally thick slices of 0.8 arcsec spanning ±2 arcsec about the disc mid-plane. For each slice, every PPV voxel above an signal-to-noise ratio of 2 was deprojected into disc coordinates then linearly interpolated on to a regular grid with the results shown in Fig. 5. The same procedure was performed for ^{13}CO, however, with narrower slices of 0.6 arcsec spanning ±1.5 arcsec with the results shown in Fig. 6.

As with the TRD, the western side of the ^{12}CO emission, positive z values, panels (e)–(g), is considerably brighter than the eastern side, negative z values, due to the slight deviation from a completely edge-on disc (Dutrey et al. 2017). It is also clear that at large separations from the disc mid-plane, the inner edge of the ^{12}CO emission moved outwards, most clearly seen in panels (b) and (g) of Fig. 5. Some azimuthal structure is tentatively observed at higher altitudes for ^{12}CO, namely in panel (f). Given the orientation of the disc, the gas rotates in a clockwise direction. Although the ^{13}CO data are noisier, some features are still observable. As with the ^{12}CO, at higher altitudes the emission peaks at $r \sim 3$ arcsec, while becoming more centrally peaks at lower z value.

Both ^{12}CO and ^{13}CO show an enhancement in emission close to the disc mid-plane, at $(x_{\rm disc},\ |y_{\rm disc}|,\ z_{\rm disc}) \approx$ (2 arcsec, 1 arcsec, 0 arcsec), marked in Figs 5 and 6 by the black-dashed circle. Bujarrabal et al. (2009) previously reported an enhancement in ^{13}CO emission at an offset position of $(\delta x_{\rm sky},\ \delta y_{\rm sky}) \approx$ (1.5 arcsec, −2.5 arcsec), with later observations of 8.6d and 11.2 μm PAH emission revealing a similar apparent overdensity (Berné et al. 2015). We note that, in principle, it is possible to subtract an azimuthally averaged model from each of these projected maps. However, we found that given the high azimuthal variability owing to the noise in the data and strong systematic feature due to the transformation from the limited spatial resolution of the data, these did not yield residual maps in which structure was readily distinguished, with higher spatial and spectral resolution data necessary for such an approach.

4 DISCUSSION

In the previous section, we have shown that assuming that an edge-on disc is in the Keplerian rotation allows one to deproject pixels in PPV space into disc-centric position–position–position space. In this section, we discuss the implication of these deprojections.

4.1 GoHam b

Previous studies of GoHam have detected a significant enhancement of emission in the southern half of the disc, dubbed GoHam b, seen in ^{13}CO emission and 8.6 and 11.2 μm PAH emission (Bujarrabal et al. 2009; Berné et al. 2015). They find that this excess emission could be explained with a gaseous overdensity containing a mass of 0.8 to $11.4 M_{\rm Jup}$, spread uniformly over a spherical region with a radius of ~0.6 arcsec (~150 au). Furthermore, based on models of the disc structure, it is estimated that the disc of GoHam is marginally gravitationally unstable, with Toomre parameter $Q \lesssim 2$ (Berné et al. 2015). In circumstellar discs, gravitational instabilities can lead to growth of local, gravitationally bound overensities (i.e. to disc fragmentation; Gammie 2001; Rice et al. 2003). It has been hypothesized that such self-gravitating overdensities could be precursors to giant planets (Boss 1997, 1998). In fact, formation by gravitational instability is favoured for giant planets on wide orbits (e.g. Morales et al. 2019). This poses the question of whether

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

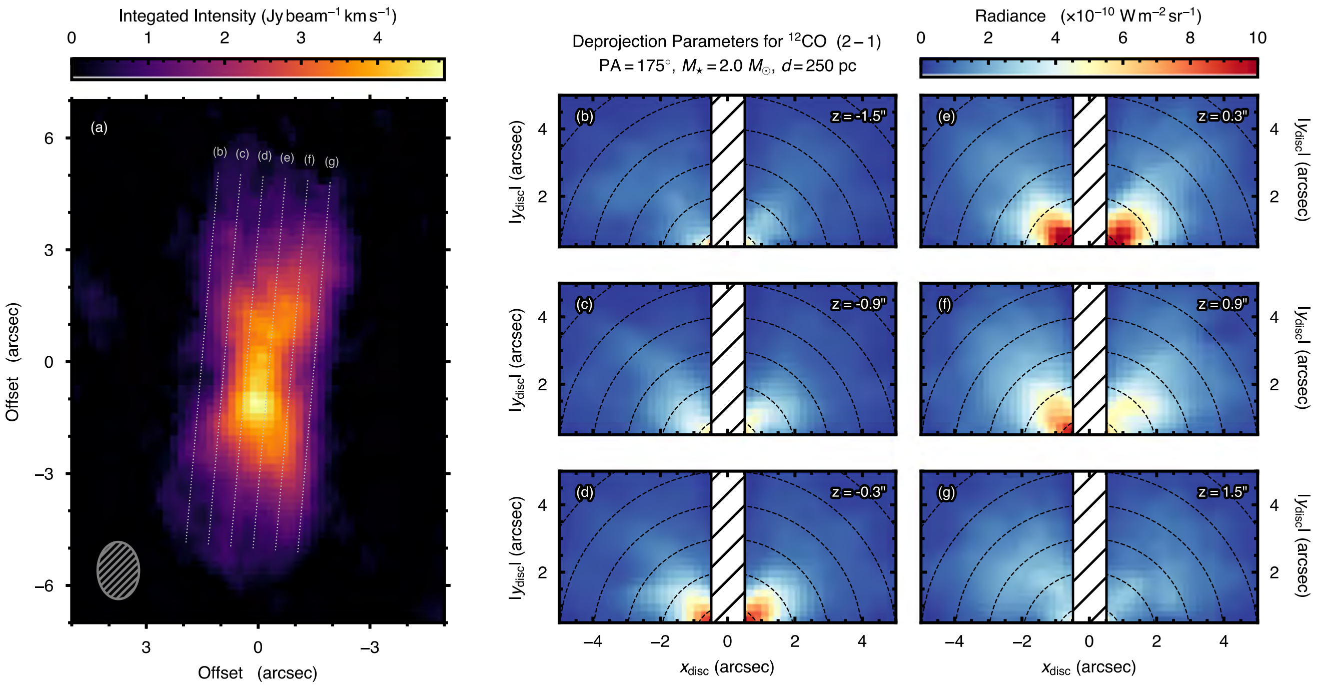


Figure 5. Deprojected ^{12}CO emission assuming Keplerian rotation. The left-hand panel shows the zeroth moment (integrated intensity) map of ^{12}CO. The six annotated slices, Panel (b)–(g), show the centre of the cuts which make up the two columns to the right-hand side. To the right-hand side, two columns, $y_{\rm disc}$ represents the line-of-sight axis. For the deprojected data, regions where $|x| < 0.5$ arcsec and $|y| < 0.5$ arcsec are masked. The height of each cut relative to the disc mid-plane is shown in the top right-hand of each panel. Note that negative x values are to the north of the disc centre. In panels (b)–(g), the black-dashed lines are lines of the constant cylindrical radius.

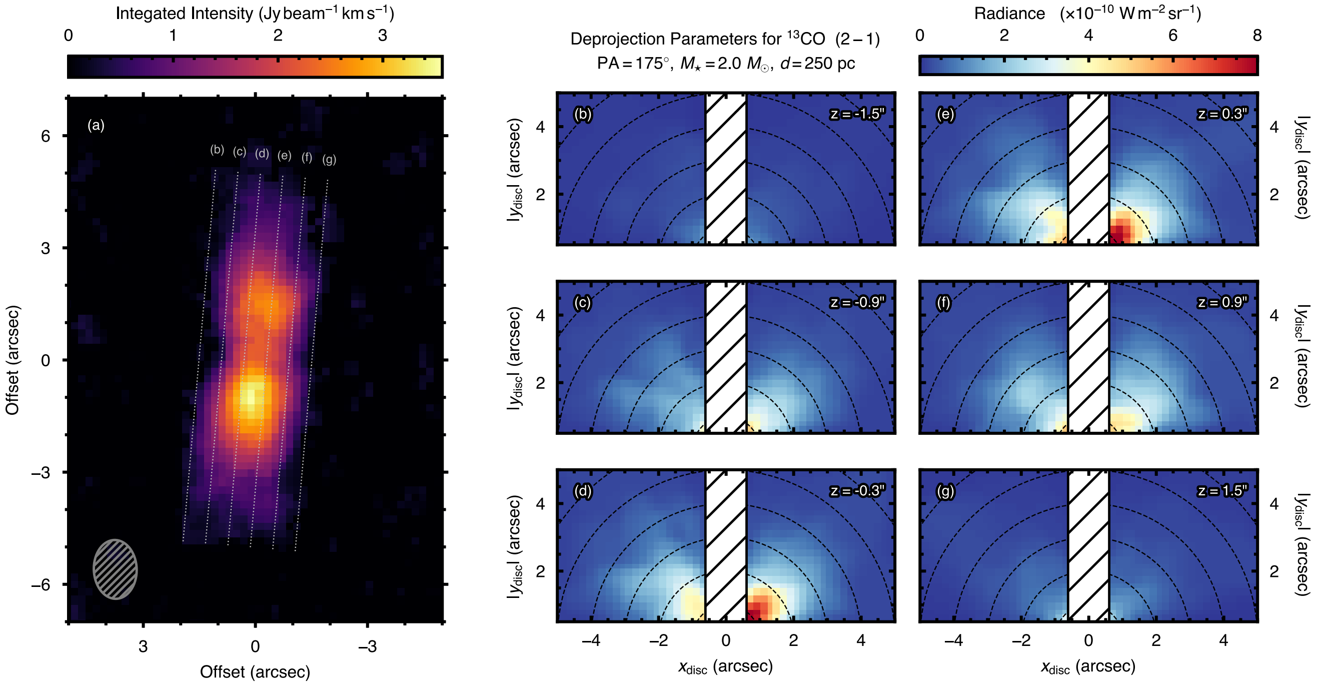


Figure 6. As in Fig. 5, but for ^{13}CO. In the integrated intensity map, left-hand panels, GoHam b manifests as a bright southern side of the disc. In the right-hand panels, GoHam b manifests as a significant asymmetry with the positive $x_{\rm disc}$ side of the disc being considerably brighter.

GoHam b may be a young protoplanet formed via gravitationally instability.

To test this hypothesis, we need to understand the 3D structure of the edge-on disc, which can be achieved using the deprojection techniques discussed above. In panels (c)–(f) of Fig. 6, the right half of the disc (corresponding to the southern half of the disc on the sky) is considerably brighter than the left half which we interpret as GoHam b. A similar asymmetry is seen in the ^{12}CO emission,

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

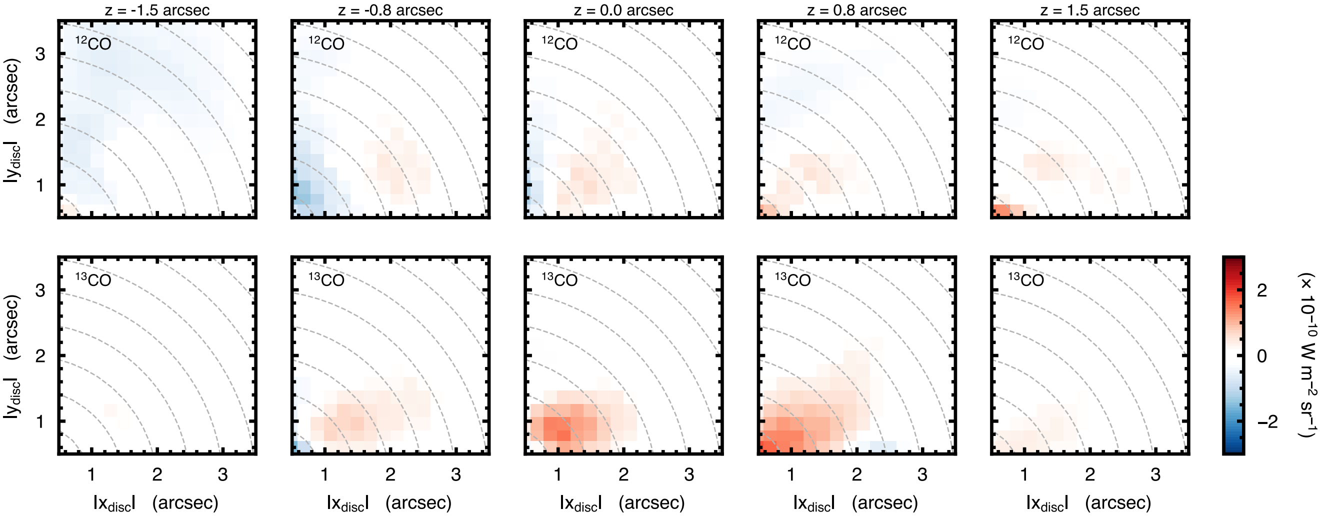


Figure 7. Residuals between the positive and negative $x_{\rm disc}$ panels from Figs 5, top row, and 6, bottom row. Each column represents a slice at a different height above or below the mid-plane, given at the top of each column. Positive values represent that the positive $x_{\rm disk}$ side is brighter, while negative values suggest that the negative $x_{\rm disc}$ side is brighter. The location of GoHam b is more readily seen for the ^{13}CO emission.

however at a much lower significance. This is more readily seen in Fig. 7, which shows the residuals between positive $x_{\rm disk}$ and negative $x_{\rm disc}$ quadrants of the deprojections shown in Figs 5 and 6. While this projection leave it ambiguous whether the feature is at positive or negative $x_{\rm disc}$, it is clear from the brighter southern side of the disc that these residuals are dominated by an excess of emission in the positive $x_{\rm disc}$ direction. The deprojection shows that the excess in emission is localized in all three dimensions, further confirming it as a local overdensity and not due to chance line-of-sight projection effects. These properties are consistent with what would be expected from an object formed via gravitational fragmentation of the disc.

GoHam b is also tentatively detected in the ^{13}CO panel of Fig. 4, consistent in location with the bright peaks seen in the zeroth moment maps in Figs 5 and 6. However, without the line-of-sight deprojection discussed above, it is hard to fully disentangle the contribution from GoHam b relative to the background. Future observations designed for these sort of analyses will benefit from first inferring a disc-averaged $(r,\ z)$ emission distribution, before using that as a background model to more readily identify deviations in the line-of-sight deprojections.

The enhancement in the brightness temperature shown in Fig. 6 is $\approx$ 20 per cent, comparable to that found in previous studies of this source (Bujarrabal et al. 2009). For an optically thin molecular line, the emission is linearly proportional to the product of the gas temperature and the column density of the emitting molecule, while an optically thick line is only proportional to the gas temperature. It is therefore tempting to assume that ^{13}CO is optically thin and thus offers a direct probe of the mass of GoHam b. However, given the unresolved nature of GoHam b and the lack of multiple transitions to infer the local excitation conditions, see Section 4.3, we do not have sufficient information to improve the estimates made previously regarding the mass of GoHam b, 0.8 to $11.4 M_{\rm Jup}$. Future observations of multiple transitions of optically thin lines will therefore provide the most accurate probe of the mass of GoHam b, leading to clues about its nature.

4.2 Utility in determining chemical stratification

As previously discussed in Dutrey et al. (2017), these deprojection techniques allow us to directly access the vertical stratification of molecular species, an essential data product with which to confront astrochemical models. This approach is hugely complementary to observations of moderately inclined discs which use the asymmetry of the line emission about the disc major axis in a moderately inclined disc to infer the height of the emission surface (e.g. Rosenfeld et al. 2013; Pinte et al. 2018).

First, the technique for moderately inclined discs can only be applied to bright lines such that the emission in any given channel is well defined. This criteria leaves only ^{12}CO and ^{13}CO as viable choices, meaning that less abundant molecules believed to arise from elevated regions, such as CH_3CN (Loomis et al. 2018), are unable to be tested. Conversely, for an edge-on disc there is no requirement on the significance of the detection; if the molecular emission can be detected in the channel maps, it can be deprojected into disc-centric coordinates.

Secondly, the deprojection techniques described in Section 3 do not require any assumptions about the optical depth of the lines to be made as all pixels can be deprojected to fill in the $(r_{\rm disc},\ z_{\rm disc})$ plane. This can be clearly seen in Fig. 8 where the ^{12}CO emission is detected in the mid-plane, where usually it is hidden due to high optical depths. For face-on or low inclination discs, an optical depth of 1 is quickly reached along the line of sight such that the disc regions behind this optical surface (the mid-plane) are hidden from view. Conversely, for an edge-on disc, the line of sight to the disc mid-plane is unobstructed, allowing us to directly probe the mid-plane emission without confusion from the upper layers. We additionally note that Dullemond et al. (2020) showed it is possible access similar information for moderately inclined sources if the spatial resolution of the data allowed for the top and bottom half of the disc to be spatially resolved.

Fig. 8 demonstrates these advantages using TRDs of ^{12}CO and ^{13}CO emission from GoHam. Note that this figure, unlike Fig. 4, has the colour scales normalized to the peak value for each molecule to bring out the structure of the emission. The ^{12}CO and likely ^{13}CO, being optically thick, will be tracing the local gas temperature. In the outer disc, the ^{13}CO emission may become optically thin, at which point the brightness is proportional to the gas temperature and local CO density. We observe that both molecules peak at elevated regions due to the chemical stratification of the disc rather than an optical depth effect as found in less inclined sources.

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

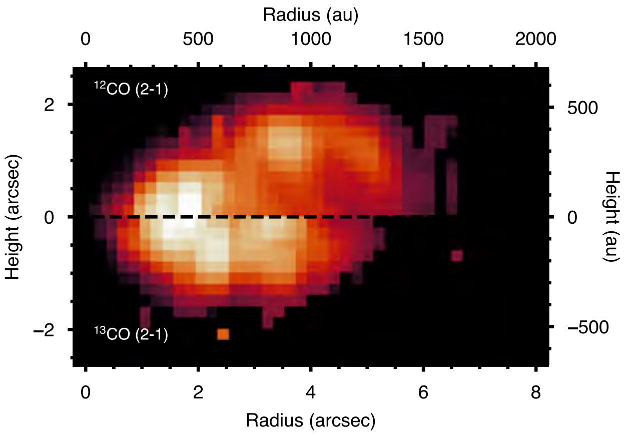


Figure 8. The molecular layers of ^{12}CO, top, and ^{13}CO, bottom. Note that the ^{13}CO has been flipped about $z = 0$ in order to provide a fair comparison due to the asymmetry due to the slight deviation from an edge-on disc. The conversion to linear scales assumed a distance of $d = 250$ pc. Note that the colour scaling has been normalized to the peak value of each to highlight structure.

^{12}CO will inhabit a more elevated region than that of ^{13}CO due to the lower abundance of ^{13}CO relative to ^{12}CO, resulting in less efficient shielding from photodissociated ultraviolet photons. A lower bound for the emission distribution will be given by the freeze-out temperature, ∼21 K for CO. The low mid-plane temperatures will not completely remove all gaseous-phase molecules, but will significantly reduce their abundance resulting in the very low level of emission seen in Fig. 8.

These observations demonstrate the utility of edge-on sources in terms of characterizing the chemical structures. Moving towards larger samples sizes, observed at higher angular and spectral resolutions will uncover the distribution of molecules currently unable to be constrained with moderately inclined discs. These deprojection techniques are readily combined with line-stacking techniques used to boost the significance of weak lines (e.g. Walsh et al. 2016), enabling studies of the molecular distribution of weak complex species, studies of which are currently hindered by their lack of bright emission.

4.3 The prospect for mapping the disc mass

With multiple transitions of a molecule observed in an edge-on source, it is possible to go beyond merely mapping out the emission distribution. For example, excitation analyses can be performed to extract local excitation temperatures and volume densities (e.g. Bergner et al. 2018; Loomis et al. 2018; Teague et al. 2018). By first deprojecting the data into 3D disc-centric coordinates, one can be certain that the emission being compared arises from the same location; an assumption always made but extremely hard to verify in less inclined sources. In other words, highly inclined sources provide access to the disc vertical structure, without losing access to the disc azimuthal structure.

Rarer CO isotopologues are less affected by optical depth issues, and therefore may be more accurate probes of the disc gas mass (e.g. ^{13}C^{17}O Booth et al. 2019). However, total gas masses derived in this way are sensitive to the assumed CO abundance. With the deprojection, it is also possible to calculate the volume of the emitting area. Thus, if the local H_2 density can be constrained using molecules which are not in non-local thermodynamic equilibrium (such as CS in the outer disc, e.g. Teague et al. 2018), this can be then be mapped to a total gas mass.

With a gas temperature and local gas mass to hand, it would then be possible to determine whether regions of the GoHam disc are gravitationally unstable (e.g. Toomre 1964). If the region around GoHam b is found to be at (or close to) instability, then this would favour its formation via the gravitational fragmentation of the disc. Such an interpretation is also supported by recent observations showing that star–disc systems similar to GoHam also appear to be unstable (e.g. HL Tau Booth & Ilee 2020).

5 SUMMARY AND CONCLUSIONS

We have used the deprojection techniques previously presented in Dutrey et al. (2017), Dent et al. (2014), Matrà et al. (2017), and Cataldi et al. (2018) to provide a 3D view of the massive disc, Gomez's Hamburger using archival SMA observations of ^{12}CO and ^{13}CO.

The deprojected data reveal a clear difference between the ^{12}CO and ^{13}CO emission regions with the ^{12}CO tracing a considerably elevated region of $z\,/\,r \sim 0.3$, while the ^{13}CO arises from much lower regions, $z\,/\,r \lesssim 0.2$, as expected from the higher abundance of ^{12}CO compared to ^{13}CO.

When deprojecting the data into the $(x_{\rm disc},\ |y_{\rm disc}|)$ plane, a clear feature in the southern side of the disc in ^{13}CO which is interpreted as the previously detected over density, GoHam b. With this deprojection, it is possible to localize the emission to $(r_{\rm disc},\ \phi_{\rm disc}) \approx (500\ {\rm au},\ \pm 30^{\circ})$, with the accuracy ultimately limited by the spatial and spectral resolution of the data.

We conclude with a discussion on the utility of these observational techniques in mapping the physical and chemical structure in protoplanetary discs. With access to the full 3D structure of the disc, future observations will be able to map out the gas temperature and density as has never been done before.

ACKNOWLEDGEMENTS

RT thanks Gianni Cataldi for discussions on unit transformations. We thanks the referee for a helpful and constructive report. RT acknowledges support from the Smithsonian Institution as an SMA Fellow. TJH is funded by a Royal Society Dorothy Hodgkin Fellowship. JDI acknowledges support from the STFC under ST/R000549/1. MRJ is funded by the President's PhD scholarship of the Imperial College London and the 'Dositeja' stipend from the Fund for Young Talents of the Serbian Ministry for Youth and Sport.

REFERENCES

Andrews S. M. et al., 2018, ApJ, 869, L41
Avenhaus H. et al., 2018, ApJ, 863, 44
Bergner J. B., Guzmán V. G., Öberg K. I., Loomis R. A., Pegues J., 2018, ApJ, 857, 69
Berné O. et al., 2015, A&A, 578, L8
Booth A. S., Ilee J. D., 2020, MNRAS, 493, L108
Booth A. S., Walsh C., Ilee J. D., Notsu S., Qi C., Nomura H., Akiyama E., 2019, ApJ, 882, L31
Boss A. P., 1997, Science, 276, 1836
Boss A. P., 1998, ApJ, 503, 923
Bujarrabal V., Young K., Fong D., 2008, A&A, 483, 839
Bujarrabal V., Young K., Castro-Carrizo A., 2009, A&A, 500, 1077
Cataldi G. et al., 2018, ApJ, 861, 72
De Beck E., Decin L., de Koter A., Justtanont K., Verhoelst T., Kemper F., Menten K. M., 2010, A&A, 523, A18

Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

de Gregorio-Monsalvo I. et al., 2013, A&A, 557, A133
Dent W. R. F. et al., 2014, Science, 343, 1490
Dipierro G., Pinilla P., Lodato G., Testi L., 2015a, MNRAS, 451, 974
Dipierro G., Price D., Laibe G., Hirsh K., Cerioli A., Lodato G., 2015b, MNRAS, 453, L73
Dong R., Hall C., Rice K., Chiang E., 2015, ApJ, 812, L32
Dullemond C. P., Isella A., Andrews S. M., Skobleva I., Dzyurkevich N., 2020, A&A, 633, A137
Dutrey A. et al., 2017, A&A, 607, A130
Fedele D. et al., 2018, A&A, 610, A24
Flock M., Ruge J. P., Dzyurkevich N., Henning T., Klahr H., Wolf S., 2015, A&A, 574, A68
Gammie C. F., 2001, ApJ, 553, 174
Guilloteau S. et al., 2016, A&A, 586, L1
Hall C., Forgan D., Rice K., Harries T. J., Klaassen P. D., Biller B., 2016, MNRAS, 458, 306
Keppler M. et al., 2018, A&A, 617, A44
Loomis R. A., Cleeves L. I., Öberg K. I., Aikawa Y., Bergner J., Furuya K., Guzman V. V., Walsh C., 2018, ApJ, 859, 131
Louvet F., Dougados C., Cabrit S., Mardones D., Ménard F., Tabone B., Pinte C., Dent W. R. F., 2018, A&A, 618, A120
Matrà L. et al., 2017, MNRAS, 464, 1415
Meru F., Juhász A., Ilee J. D., Clarke C. J., Rosotti G. P., Booth R. A., 2017, ApJ, 839, L24
Morales J. C. et al., 2019, Science, 365, 1441
Pinte C. et al., 2018, A&A, 609, A47
Rice W. K. M., Armitage P. J., Bonnell I. A., Bate M. R., Jeffers S. V., Vine S. G., 2003, MNRAS, 346, L36
Rosenfeld K. A., Andrews S. M., Hughes A. M., Wilner D. J., Qi C., 2013, ApJ, 774, 16
Teague R. et al., 2018, ApJ, 864, 133
Teague R., 2019a, Res. Notes AAS, 3, 74
Teague R., 2019b, J. Open Source Softw., 4, 1220
Teague R., Foreman-Mackey D., 2018, Res. Notes AAS, 2, 173
Toomre A., 1964, ApJ, 139, 1217
Walsh C. et al., 2016, ApJ, 823, L10
Wood K., Whitney B. A., Robitaille T., Draine B. T., 2008, ApJ, 688, 1118
Zhang S. et al., 2018, ApJ, 869, L47

This paper has been typeset from a TEX/LATEX file prepared by the author.

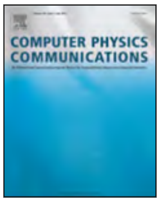
Downloaded from https://academic.oup.com/mnras/article/495/1/451/5827536 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Computer Physics Communications 250 (2020) 107052

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc



A guide to hunting periodic three-body orbits with non-vanishing angular momentum[☆]

Marija R. Janković[a], V. Dmitrašinović[b,*], Milovan Šuvakov[b,c]

[a] Department of Physics, Imperial College London, 1010 Blackett Lab., Prince Consort Rd., London SW7 2AZ, UK
[b] Institute of Physics Belgrade, Belgrade University, Pregrevica 118, Zemun, P.O. Box 57, 11080 Beograd, Serbia
[c] Department of Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA[1]

ARTICLE INFO

Article history:
Received 17 June 2019
Received in revised form 6 November 2019
Accepted 13 November 2019
Available online 21 November 2019

Keywords:
Numerical methods
Three-body systems
Newtonian gravity

ABSTRACT

A large number of periodic three-body orbits with vanishing angular momentum have been found in Newtonian gravity over the past 6 years due to a simple search method and to the contribution from practitioners outside the Celestial Mechanics community. Extension of such orbits to non-vanishing angular momentum has been lacking due to *inter alia* the absence of a sufficiently simple and widely known search method. We present a method, i.e., a general strategy plus detailed tactics (but not a specific algorithm, or a code), to numerically search for relative periodic orbits in the Newtonian three-body problem with three equal masses and non-vanishing angular momentum. We illustrate the method with an application to a specific, so-called Broucke–Hadjidemetriou–Hénon (BHH) family of periodic 3-body orbits: Our search yielded around 100 new "satellite" orbits, related to the original BHH orbits by a topological relation (defined in the text), with infinitely many orbits remaining to be discovered. We used the so-obtained orbits to test the period vs. topology relation that had previously been established, within a certain numerical accuracy, for orbits with vanishing angular momentum. Our method can be readily: (1) applied to families of periodic 3-body orbits other than the BHH one; (2) implemented using various standard algorithms for solving ordinary differential equations, such as the Bulirsch–Stoer and the Runge–Kutta–Fehlberg ones; (3) adapted to 3-body systems with distinct masses and/or coupling constants, including, but not limited to, Coulomb interaction. Our goal is to enable numerical searches for new orbits in as many families of orbits as possible, and thus to allow searches for other empirical relations, such as the aforementioned topology vs. period one.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

The three-body problem, as formulated by Newton, is to predict the motion of the Sun–Earth–Moon system [1–5]. Euler [6] and Lagrange [7] found their respective (analytic) solutions in the mid- to late 18th century, but that was of no immediate astronomical significance — the first Jovian satellites were only discovered 150 years after Lagrange's calculation, and even today such Lagrangian systems comprise less than 1% of all known three-body systems, the remaining 99% being the so-called "hierarchical" systems, such as the Sun–Earth–Moon one, [5,8].[2]

In the late 19th century Bruns showed that the general Newtonian three-body problem is not integrable [1], which explained the absence of new solutions at that time. That (should have) made it clear that there would be no further progress without numerical investigations. The first new periodic orbit in the unrestricted three-body problem arrived in 1956 when Schubart [9] found his collinear, and colliding periodic orbit using a mechanical computer.[3]

The numerical studies of periodic orbits in the general, i.e., unrestricted three-body problem (3BP) began in earnest about 50 years ago: The first new orbits after Schubart's one were announced in 1967 by Szebehely and Peters [13,14] who found several "free-fall" periodic orbits using electronic computers.

[☆] The review of this paper was arranged by Prof. Hazel Andrew.

[*] Corresponding author.
E-mail address: dmitrasin@ipb.ac.rs (V. Dmitrašinović).

[1] On sabbatical leave of absence.

[2] What is meant by "hierarchical" 3-body system is one in which two bodies move around each other, and thus form a "binary", and the third moves around the binary, or *vice versa*.

[3] At about the same time in mid-20th century, doubts about the existence of any further solution, i.e., other than the Eulerian and Lagrangian ones, were formally cast [10], and then equally formally refuted: Arenstorf [11] published an existence proof for periodic solutions of the general three-body problem, albeit without examples, and Jefferys and Moser [12] had also published existence proofs for "almost periodic solutions" in the three-dimensional case, also without examples. Without at least one explicit example of a new periodic orbit, that would have been just another academic controversy.

https://doi.org/10.1016/j.cpc.2019.107052
0010-4655/© 2019 Elsevier B.V. All rights reserved.

Subsequently, Standish [15] published several other free-fall orbits. While these publications definitely settled the question of whether the general problem had non-trivial periodic solutions other than the Euler and Lagrange ones, they did not begin to address the question of hierarchical orbits. Moreover, Szebehely and Peters [13,14] suggested that the periodic orbits were isolated.[4] This led to some confusion, which was resolved in 1974 by Hénon [16], who extended Szebehely and Peters' free-fall, i.e., zero-angular-momentum solutions to non-zero values of angular momentum, and showed that they form one-parameter continuous families of orbits. This is a general property of (relative)[5] periodic orbits and will be called Hénon's first theorem.

Further progress was based on this fact and the subsequent numerical discoveries of hierarchical periodic three-body orbits by Broucke's [17,18], Hadjidemetriou's [19–21], and by Hénon's [22–24] groups, working separately, though aware of each other's work, and using different methods. We shall refer to these three groups of authors collectively as BHH. BHH had not only confirmed the existence, but found two kinds ("prograde" and "retrograde", the latter with three branches) of stable periodic solutions, both for equal and unequal masses, as well as for different values of the angular momentum. The stable solutions among those found by BHH are in agreement with observed hierarchical systems [8,17,19]. As the scale-invariant angular momentum[6] is reduced, the hierarchical nature of these solutions is lost, and all three bodies become equally involved in the motion (which is sometimes called "interplay" of three bodies). Such orbits have not been observed, as yet, even though some of them are (linearly) stable. This "lost branch" of BHH solutions remains a challenge for both observational and theoretical astronomy. In the meantime, rigorous existence proofs, at least at certain discrete values of the angular momentum, have been supplied for some of the BHH orbits [26,27]. All of this ought to make it clear that the BHH family of solutions is important, both for astronomical applications and for mathematical purposes, more about which later.

In the meantime, more than 2000 new topologically distinct zero-angular-momentum three-body orbits have been reported [28–43]. By virtue of Hénon's first theorem, each and every one of these orbits defines (the beginning of) a distinct family of orbits with non-zero angular momentum, only a few of which have been studied [30,44–48] to any extent. Though it is practically impossible, at least for one group of investigators, to study all 2000 families, around 45 linearly stable orbits deserve to be further looked into. For such an endeavor, one needs a general method that has been lacking heretofore: the papers [44–47] rely on different techniques designed specifically for one, or another, particular family of orbits. A few years ago we extended a previously established search method for zero-angular momentum orbits (see Refs. [33,34,37]) to the BHH family of orbits with non-zero angular momentum [39]. It must be emphasized that this search method differs significantly from all three original BHH methods. Indeed, the three BHH methods were designed such that only one (what we now call the progenitor) orbit (at given energy and angular momentum) in the BHH family could be found — the first (and only) satellite orbit (a related orbit, defined in Section 2) of the BHH family, prior to [39], was discovered and reported in Ref. [44], perhaps unwittingly.

As there was no known reason why the number of such satellite orbits ought to be limited – indeed the Birkhoff–Lewis theorem [49] decrees the opposite – the search for BHH satellite orbits had to be conducted, for both practical astronomical and for theoretical reasons. We found around 100 such new orbits, [39], with (infinitely) many more waiting to be discovered, limited only by one's strength and/or patience and availability of computer resources. A similar situation may hold in other families of orbits.

In the meantime, we have realized that our method is sufficiently wide to accommodate searches for periodic orbits in some other, though not all (see Section 3.2), families of the Newtonian gravitational three-body problem, with minimal modifications also for distinct masses; but also in other non-relativistic three-body problem involving homogeneous potentials, such as the Coulombic one [50]. Therefore, the purpose of this work is two-fold:

- methodological: we have extended our previous search and scanning method [34] to periodic orbits with non-zero angular momentum. This was a non-trivial endeavor because one new continuous parameter (angular momentum L) enters the calculation, and the search is necessarily in a three-dimensional subspace of the full six-dimensional phase space of initial conditions, but one with a successful outcome.
  We have modified the method of minimizing the return proximity function in the phase space of initial conditions to the present search in two-dimensional "slices" of a three-dimensional sub-space. This, of course, has the consequence that a detailed search of the complete subspace would last too long to be practically implemented in a reasonably short time. Consequently, we searched only in the immediate vicinity of (previously discovered) progenitor BHH orbits, and, even with this limitation, we found around 100 new orbits.
- particular: to find as many as possible new satellite orbits of the BHH family and then to investigate (any, new) topological regularities among them, such as those discovered among orbits with vanishing angular momentum [36,37]. Davoust and Broucke [44] had found the first (k=3) satellite orbit in the retrograde branch of the BHH family. We extended the search for BHH satellite orbits systematically, at first up to values $k \leq 19$, where k is the so-called topological exponent, defined in Section 2.1 and Ref. [32], and then less systematically up to $k = 84$.[7]

The completion of the long-term goal of exploring the stable families of three-body orbits, as described above, can only be accomplished by a concerted effort by several teams using ever-more-powerful computing facilities. For this reason, here we publish technical details of our search method in the hope that someone else will take over the torch.

This paper is organized into six sections: After the present Introduction, we present necessary preliminaries in Section 2. In Section 3 we discuss our search method and, in particular the subspace of initial conditions. In Section 4 we present our numerical results together with our estimates of numerical uncertainties. Then in Section 5 we discuss the scaling laws for three body orbits and the expected topological dependence of the scaling-law "constant" for three bodies. In Section 6 we discuss the open questions and suggest future searches. Finally, in Section 7 we summarize and draw our conclusions.

---

[4] "Recent numerical investigations [13–15] have led to the conjecture that periodic solutions of the planar general problem of three bodies are isolated, for given masses of the bodies. To quote, for instance, Szebehely (1973): 'The periodic orbits of the general problem do not seem to form families in the same sense we know families in the restricted problems'; and 'to establish families of periodic orbits according to what is known today, requires changes in the participating masses as well as in the initial conditions'. This conjecture is made for arbitrary distances between the participating bodies"., a quote from [16].

[5] What is meant here by a relative periodic orbit is one that returns to its initial position after a period, though rotated by a finite angle, see Section 3.1.

[6] The change of angular momentum, while keeping the same form of the orbits, generally implies a change of energy, or of size. Due to the scaling rules [25] for orbits in the Newtonian potential, one can define scale-invariant angular momentum, see Section 2.3.

[7] This is not to say that this sequence stops at 84, but rather, that we have arrived close to the limits imposed by the precision of our codes (and algorithms) and the computing power available to us.

2. Preliminaries

In this section we provide some preliminary information, such as the motivation for this work, as well as basic background information necessary to follow the rest of the paper. There is nothing fundamentally new in this section, though it should give a brief pedagogical introduction to matter written for readers unfamiliar with celestial mechanics in general, and the three-body problem in particular.

2.1. Testing new topological laws

Historically, periodic three-body orbits were classified into families and named according to Strömgren's nomenclature used in the restricted three-body problem, see Ref. [44]. Such a definition of families does not always correspond with the unambiguous topological definition of families, provided by Montgomery [51]: For example out of approximately 20 families discussed by Davoust & Broucke [44], only 3 are topologically distinct.

Using the topological classification method, Ref. [33] gave a precise definition of "satellites" of an arbitrary progenitor orbit w, as orbits that are the kth-power of their progenitor, i.e., with the homotopy/free-group elements that have the form w^k, where $k = 2, 3, \ldots$. Thereupon this definition was applied to the study of figure-8 satellites, which were first observed in Ref. [30], and investigated in detail in Ref. [33]. The latter study led to the discovery of remarkable topological Kepler's third law-like regularities ("laws") for orbits with vanishing angular momentum, Refs. [36,37]. An immediate question is if such regularities persist when the angular momentum does not vanish?

Ref. [39] was the first step in an attempt to answer that question, *viz.* that of finding the satellite orbits in the BHH family, as there was no guarantee that they had to exist. The present paper is an elaboration of the brief first report [39]. Indeed, it was only in Ref. [37] that the existence of satellite orbits is related to the stability of progenitor orbit (with vanishing angular momentum) was understood,[8] in terms of the Birkhoff–Lewis theorem [49].

2.2. Basic facts

Broucke [17,18,44], Hadjidemetriou [19–21] and Hénon [22,23] (BHH) explored a set of periodic planar three-body orbits with bodies that have the same mass and wherein the initial coordinates form a collinear configuration (or a "syzygy", as it is known in the astronomical literature),[9]

$$\mathbf{r_1} = (x_1, 0), \mathbf{r_2} = (x_2, 0), \mathbf{r_3} = (x_3, 0)$$

and the initial velocities are orthogonal to the vector determined by the collinear position vectors, i.e., of the form:

$$\mathbf{v}_1 = (0, \dot{y}_1), \mathbf{v}_2 = (0, \dot{y}_2), \mathbf{v}_3 = (0, \dot{y}_3)$$

In the following we shall call this a collinear orthogonal configuration. Such configurations are special insofar as they lead to discrete symmetries of the orbit, when they appear in an orbit twice [53,54]. This is not the most general Ansatz for initial velocities: collinear, i.e., x-components, of relative velocities $\mathbf{v}_i$, $i = 1, 2, 3$, need not vanish in general, but allowing for that freedom would increase the dimensionality of the search phase space by two and thus greatly increase the difficulty of search.

These orbits, an example of which is shown in Fig. 1(a), form two continuous curves of relative periodic orbits in the phase space of initial conditions, whose termini ("ends") include a collinear collision (Schubart) orbit (retrograde), and both approach the limit of a two-body problem with masses m and $2m$ at the common upper terminus ("end") of their $L(T)$ curves, Fig. 2. Concerning the latter limit, the first periodic orbits that Broucke found contained the so-called double Keplerian motion, which means that two bodies revolve tightly around each other, while the pair together revolves around the third body, therefore representing the inner and the outer binary system. If the two binaries' revolve in the same direction, e.g. (both) clock-wise, then the orbit is called direct. If they revolve in opposite directions, then the orbit is retrograde. Some of these orbits have been proven to exist in a mathematically rigorous manner in Refs. [26,27].

Broucke [17,18], Hadjidemetriou [19–21] and Hénon [22,23] talk of two families[10] of orbits – direct and retrograde – but all of these orbits belong to a single topological family: during one period the orbit completes a single "loop" around one of the poles on the shape sphere (for definition see Ref. [32,34]), see Fig. 1.b. This "loop" is described by the conjugacy class of the fundamental group/free group element $\mathtt{a}$, according to the topological classification explained in Refs. [32,34]. It turns out, however, that there are numerous relative periodic orbits with topology $\mathtt{a}^k$, where $k = 2, 3..$, that have the same form of initial conditions. Such orbits are (sometimes) called "satellites" [30,33], whereas other authors call them "bifurcation orbits" [44].

2.3. Scaling laws for Newtonian trajectories

For the sake of clarity and completeness, here we review some elementary scaling laws, indeed so elementary that they are explicitly presented in only one graduate-school level texbook on classical mechanics Ref. [25] that we know of. Nevertheless, these scaling rules have significant non-trivial consequences in the three-body problem. Here we follow our own presentation(s) from Refs. [36,39].

It is well known that Kepler's third law follows from the spatio-temporal (mechanical) scaling laws, which, in turn, follow from the homogeneity of the Newtonian gravity's potential. Under spatial scaling $\mathbf{r} \to \lambda\mathbf{r}$, the time must scale as $t \to \lambda^{3/2}t$, and consequently $\mathbf{v} \to \mathbf{v}/\sqrt{\lambda}$. The (total) energy scales as $E \to \lambda^{-1}E$, the period T as $T \to \lambda^{3/2}T$ and the angular momentum as $L \to \lambda^{1/2}L$, the last one behaving differently than either the period T, or the hyper-radius $R = \sqrt{\frac{1}{3}\sum_{i<j}^{3}(\mathbf{r}_i - \mathbf{r}_j)^2} = \sqrt{\sum_i^3(\mathbf{R}_{\text{g.b.}} - \mathbf{r}_i)^2}$, which is proportional to the root-mean-square distance of the three particles from their geometrical barycenter $\mathbf{R}_{\text{g.b.}} = \frac{1}{3}\sum_i^3 \mathbf{r}_i$,[11] and scales linearly with λ: $R \to \lambda R$ thus presenting a measure of the overall "size" of the triangle. The angular momentum L, though conserved by virtue of the equations of motion, changes ("scales") as a function of the total energy E, or of the "size" R. For this reason only vanishing angular momentum

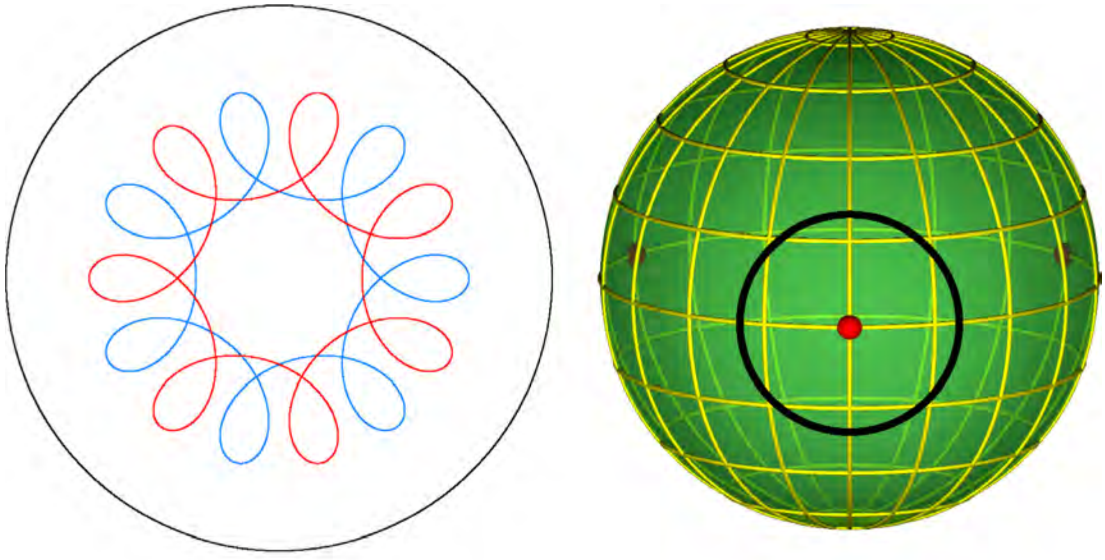
---

[8] In Ref. [39] the (unnecessarily strong) KAM theorem was invoked when the (weaker) Birkhoff–Lewis one would have sufficed.

[9] Which is in conformity with Montgomery's theorem, [52], that, with the exception of Lagrange's solution, every periodic solution to the Newtonian three-body problem passes through syzygies.

[10] whereas Davoust and Broucke [44] relate further three families (D_2, D_3, D_4) to the BHH family D_1 by way of analytic continuation through binary collisions; see "... although it is very likely that the four families D_1 through D_4 are, in fact, one single complex family" in Sect. 7. of Ref. [44].

[11] which equals the physical center-of-mass $\mathbf{R}_{\text{g.b.}} = \mathbf{R}_{CM}$ when all three masses are equal.



(a) (b)

Fig. 1. (a) One retrograde BHH orbit, an absolute periodic one. It can be obtained from a particular relative periodic orbit whose period is multiplied by 7, after which time the orbit closes its trajectory in real space. (b) the same orbit on the shape sphere.

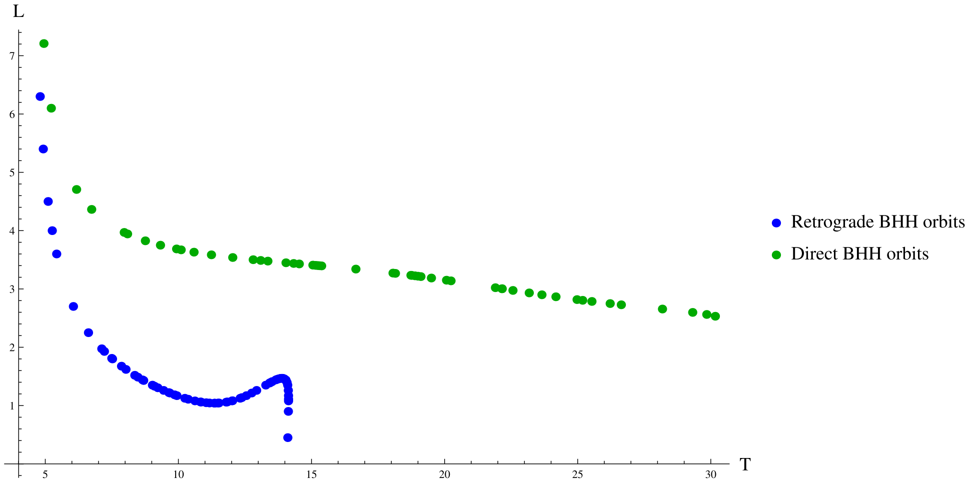


Fig. 2. $L(T)$ curves for (a) direct (prograde) orbits (green dots); and (b) retrograde BHH orbits (blue dots). All orbits are scaled to have total energy $E = -\frac{1}{2}$. Figure reproduced from Ref. [39] with permission from the publisher. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$L = 0$ is a "fixed point" under scaling transformations. Therefore, in the following we shall use the scale-invariant angular momentum $L_r = L|E|^{1/2}$[12] and, for simplicity's sake, equal masses. Thus, we (may) replace the (typical, or mean, or maximum) "size" $\bar{R}$ of the three-body system in Kepler's third law $T \propto \bar{R}^{3/2}$ with the inverse absolute value of energy $|E|^{-1}$, i.e., $T \propto |E|^{-3/2}$, or equivalently $T|E|^{3/2} =$ const. . These scaling laws hold for any number of bodies interacting by Newtonian gravity.

The "constant" on the right-hand-side of the equation $T|E|^{3/2} =$ const. is not universal in the three-body case, as it is in the two-body case — it may depend on all or any one of the following: the topological family w of the three-body orbit, described by the free-group word w, on the mass ratios, and on the scale-invariant angular momentum $L_r = L|E|^{1/2}$, see Refs. [22,23]

$$T^{(w)}|E|^{3/2} = f(L^{(w)}|E|^{1/2}),$$

or as an inverse function:

$$L^{(w)}|E|^{1/2} = f^{-1}(T^{(w)}|E|^{3/2}).$$

Thus, the curve $L_r^{(w)}(T_r^{(w)}) = L^{(w)}|E|^{1/2}(T^{(w)}|E|^{3/2})$ as a function of $T_r^{(w)} = T^{(w)}|E|^{3/2}$ is a fundamental property of any family w of periodic three-body orbits. For the BHH family the $L(T)$ curve shown in Fig. 2 is based on the data from Refs. [17–23].

---

[12] Davoust and Broucke [44] used the combination of variables $27L^2E$, which is effectively the negative of 27 times the square of L_r.

2.4. The return proximity function

The return proximity function $d(\mathbf{X}_0, T_0)$ in phase space is defined as the absolute minimum of the $d(\mathbf{X}_0, T_0) = \min_{t_m < t \le T_0} \left|\mathbf{X}(t) - \mathbf{X}(0)\right|$, of the 12-dimensional state vector $\mathbf{X}_0 = \mathbf{X}(t) - \mathbf{X}(0)$ evolving from the initial time 0 to the time t, where

$$\left|\mathbf{X}(t) - \mathbf{X}(0)\right| = \sqrt{\sum_i^3 [\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2 + \sum_i^3 [\mathbf{p}_i(t) - \mathbf{p}_i(0)]^2} \tag{1}$$

is the Euclidean norm ("distance" between two 12-vectors) in the 12-dimensional Euclidean phase space consisting of the Cartesian coordinates and velocities of all three bodies without removing the center-of-mass motion, and t_m is the shortest non-zero time such that $\left.\frac{d\left|\mathbf{X}(t)-\mathbf{X}(0)\right|}{dt}\right|_{t=t_m} = 0$. The recurrence time $\tau(\mathbf{X}_0, T_0)$ is the time t at which a minimum of $\left|\mathbf{X}(t) - \mathbf{X}(0)\right|$ is reached. Searching for periodic solutions with a period $T < T_0$ is equivalent to finding zeros of the return proximity function.

A definition analogous to Eq. (1) holds for the 8-vector $\mathbf{Y}(t)$, made up of Jacobi relative vectors $(\boldsymbol{\rho}, \boldsymbol{\lambda})$:

$$\boldsymbol{\rho} = \frac{1}{\sqrt{2}}(\mathbf{r}_1 - \mathbf{r}_2)$$

$$\boldsymbol{\lambda} = \frac{1}{\sqrt{6}}(\mathbf{r}_1 + \mathbf{r}_2 - 2\mathbf{r}_3)$$

and their time derivatives $(\dot{\boldsymbol{\rho}}, \dot{\boldsymbol{\lambda}})$; here we have eliminated the center-of-mass vector and its corresponding linear momentum from the phase space. Similarly, the 6-vector $\mathbf{Z}(t) = (x, y, z, \dot{x}, \dot{y}, \dot{z})$ consisting of "three-vectors" (x, y, z)

$$x = \frac{2\boldsymbol{\rho} \cdot \boldsymbol{\lambda}}{R^2}, \quad y = \frac{\boldsymbol{\lambda}^2 - \boldsymbol{\rho}^2}{R^2}, \quad z = \frac{2(\boldsymbol{\rho} \times \boldsymbol{\lambda}) \cdot \mathbf{e}_z}{R^2} \tag{2}$$

where $R = \sqrt{\boldsymbol{\rho}^2 + \boldsymbol{\lambda}^2}$ is the hyper-radius, which are scalar products of Jacobi vectors, and thus rotation-invariant, and their time derivatives $(\dot{x}, \dot{y}, \dot{z})$.

3. The space of initial conditions for BHH orbits

In the preceding section we showed some of the arguments used by BHH, as well as some of their results. In the following we present our method and describe its advantages.

3.1. Relative periodic orbits

Relative periodic orbits are such that the system returns, after one period, to its initial configuration, albeit rotated through some angle. When this total rotation angle equals zero, or is a multiple of 2π, the solution is called absolute periodic. A relative periodic orbit with a total rotation angle that is commensurate with 2π yields, after a certain number of periods, an absolute periodic orbit. All of the BHH orbits are relative periodic, and some of them also lead to absolute periodic orbits.

Therefore, we shall search only for relative periodic orbits, by eliminating the rotations, i.e., by using the 6-vector, $\mathbf{Z}(t) = (x, y, z, \dot{x}, \dot{y}, \dot{z})$, defined in Eq. (2). The overall (total) rotation angle Φ of the system can be reconstructed from the trajectory in this hyper-space and the equation for angular momentum conservation. By using the relative return proximity function, $d(\mathbf{Z}_0, T_0) = \min_{t_m < t \le T_0} |\mathbf{Z}(t) - \mathbf{Z}_0|$, in the minimization procedure, Refs. [32,34,39], we are effectively searching for relative periodic orbits.

The collinear orthogonal configuration is a fixed point of a reversing symmetry [17,53,54] and as such has consequences for periodic orbits in the problem of three-bodies, as noted already by Broucke and Boggs [17] and by Bengochea et al. [55]. If an orbit passes through two, or more, collinear orthogonal configurations, then it possesses a higher (discrete) symmetry, as formalized in the "mirror theorem", Refs. [17,53–55]. We shall find multiple examples of such discrete symmetries among our three-body orbits.

3.2. Parametrization of initial conditions

Montgomery [52] has shown that all periodic 3-body orbits, with the exception of the Lagrange one, must encounter a syzygy, or equivalently cross the equator on the shape sphere, at least once during one period. Thus, by choosing a collinear configuration (a syzygy), though not necessarily a collinear orthogonal configuration for the initial one, one does not lose generality. Thus one reduces the number of independent Cartesian components of initial relative vectors and velocities from eight down to six. Therefore, the most general parametrization of initial conditions involves six parameters: (1) two for the initial configuration; and (2) four for the initial velocities. Fixing one of two initial configuration parameters can be thought of as constraining the initial size of the system, which reduces the number of free parameters to five.

The energy and the angular momentum conservation impose two further linearly independent, but non-linearly related constraints on the initial vectors and velocities. Indeed, fixing the energy at a particular value is equivalent to choosing a particular initial size of the system, so we shall leave these intricacies for Section 3.2.2. Five-dimensional search space is still too large for practical implementation. Therefore, we chose to set two collinear components of the initial relative velocities equal to zero, and keep only the orthogonal ones. This choice reduces the number of parameters to three.

The same choice is consistent with, i.e., sufficient but not necessary for, vanishing of the time derivative of the hyperradius R at the initial (syzygy) time $t = 0$: $\dot{R}(t = 0) = 0$, see Section 3.2.2. In other words $\dot{R}(t = 0) = 0$, does not imply orthogonality of the initial velocities to the initial vector. The additional initial constraint $\dot{R}(t = 0) = 0$ is satisfied by all of the (non-zero-angular-momentum) BHH orbits [17–23], and by all of the (non-zero-angular-momentum) Davoust–Broucke orbits [44], as well as by all of the (zero-angular-momentum) solutions in Refs. [33,37,42,43]. Thus, the collinear orthogonal initial configuration is sufficiently wide to encompass many, though not all of presently known families of periodic orbits. One exception is Martin Grant's Rosette orbit, Ref. [56], for his initial conditions see footnote.[13] This set of i.c.s appears to have 8 independent variables. By evolving to a collinear configuration this number is reduced to six, see Section 6.2.4. One of two initial configuration parameters may be fixed (e.g. to unity, as in Section 3.3) which leaves us with five free parameters in total, which is more than the three parameters allowed in our search subspace. This means that this solution does not fit into the class with collinear orthogonal initial conditions. Nevertheless, this orbit passes through 3 "reversor" isosceles configurations, which fact guarantees additional discrete symmetry of the orbit.

---

[13] Martin Grant's i.c.s are: $x_1(0) = 0.7812$, $y_1(0) = -0.2465$, $x_2(0) = y_1(0) = -0.2465$, $y_2(0) = x_1(0) = 0.7812$, $x_3(0) = -x_1(0) - x_2(0) = -0.5347$, $y_3(0) = x_3(0) = -0.5347$, $\dot{x}_1(0) = -\dot{y}_2(0) = -0.6087$, $\dot{y}_1(0) = -\dot{x}_2(0) = -0.286$, $\dot{x}_3(0) = -\dot{x}_1(0) - \dot{x}_2(0) = 0.3227$, $\dot{y}_3(0) = -\dot{x}_3(0) = -0.3227$, $T = 17.0874$, $E = -1.89451$, $L = -0.40184$.

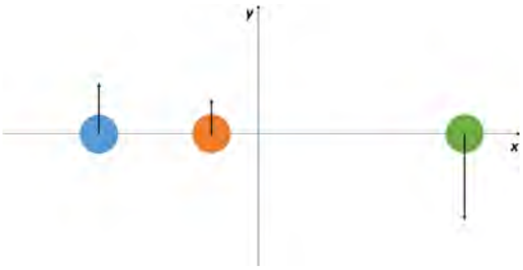


Fig. 3. Geometry of initial conditions.

3.2.1. *Initial configuration*

The initial configuration is collinear, (or recti-linear) orthogonal, though generally asymmetric, as shown in Fig. 3, so as to yield non-zero angular momentum, and can be written in terms of (initial configuration) Jacobi vectors $\boldsymbol{\rho}_0, \boldsymbol{\lambda}_0$ as:

$$\boldsymbol{\rho}_0 = (a, 0), \quad \boldsymbol{\lambda}_0 = (b, 0),$$

and the initial position of the center-of-mass (CM)

$$\mathbf{R}_0^{\mathrm{CM}} = \frac{1}{3}\left(\mathbf{r}_0^1 + \mathbf{r}_0^2 + \mathbf{r}_0^3\right) = \frac{1}{3}\left(x_1 + x_2 + x_3\right)\hat{\mathbf{x}} = 0$$

From these three linear algebraic equations one can easily find

$$x_3 - x_1 = -\frac{a + \sqrt{3}b}{\sqrt{2}}, \quad x_3 - x_2 = \frac{a - \sqrt{3}b}{\sqrt{2}}, \quad x_1 - x_2 = \sqrt{2}a, \tag{3}$$

the initial potential energy equals

$$V = -\sum_{i<j}\frac{1}{r_{ij}} = -\left(\frac{\sqrt{2}}{a + \sqrt{3}b} + \frac{\sqrt{2}}{|a - \sqrt{3}b|} + \frac{1}{\sqrt{2}a}\right), \tag{4}$$

and the initial hyper-radius is $R_0 = \sqrt{\boldsymbol{\rho}_0^2 + \boldsymbol{\lambda}_0^2} = \sqrt{a^2 + b^2}$.

3.2.2. *Initial velocities*

We demand that the CM velocity vanishes

$$\dot{\mathbf{R}}_{CM} = \frac{1}{3}\left(\dot{\mathbf{r}}_1 + \dot{\mathbf{r}}_2 + \dot{\mathbf{r}}_3\right) = 0$$

which leaves four independent components of two two-dimensional Jacobi velocity vectors, as advertised earlier.

Next, we impose $\dot{R}(0) = 0$ as an initial condition. The hyperradius R can be expressed as $R = \sqrt{\boldsymbol{\rho}^2 + \boldsymbol{\lambda}^2}$, a function of the Jacobi vectors $\boldsymbol{\rho}, \boldsymbol{\lambda}$. This leads to the requirement

$$R(0)\dot{R}(0) = \boldsymbol{\rho}_0 \cdot \dot{\boldsymbol{\rho}}_0 + \boldsymbol{\lambda}_0 \cdot \dot{\boldsymbol{\lambda}}_0 = 0,$$

as $R(0) \neq 0$. Since $(\boldsymbol{\rho}_0, \boldsymbol{\lambda}_0)$ have only components in one (the x-) direction, the above equation imposes a constraint only on a linear combination of the x-components of the velocity vectors, and no condition on the orthogonal (y-) components, which are, in turn, determined by the angular momentum

$$L = m\left(\boldsymbol{\rho}_0 \times \dot{\boldsymbol{\rho}}_0 + \boldsymbol{\lambda}_0 \times \dot{\boldsymbol{\lambda}}_0\right).$$

This is most simply solved by the requirement that the initial velocity vectors have no x-component, i.e., that they are parallel, and orthogonal to the initial spatial separation vectors:

$$\dot{\boldsymbol{\rho}}_0 = c\hat{\mathbf{y}} = \frac{1}{\sqrt{2}}\left(\dot{y}_1 - \dot{y}_2\right)\hat{\mathbf{y}},$$

$$\dot{\boldsymbol{\lambda}}_0 = d\hat{\mathbf{y}} = \frac{1}{\sqrt{6}}\left(\dot{y}_1 + \dot{y}_2 - 2\dot{y}_3\right)\hat{\mathbf{y}},$$

which leaves us with two additional free parameters (c, d).

Of course, this is not the most general set of initial conditions satisfying our additional constraint $\dot{R}(0) = 0$ – it can be augmented/extended by adding two new parameters, the x-components of the velocity vectors, $e = \dot{\boldsymbol{\rho}}_x, f = \dot{\boldsymbol{\lambda}}_x$, that satisfy

$$\boldsymbol{\rho}_x\dot{\boldsymbol{\rho}}_x + \boldsymbol{\lambda}_x\dot{\boldsymbol{\lambda}}_x = 0,$$

thus leaving (only) one independent new parameter.

Our Ansatz $e = f = 0$ is sufficiently wide to cover the BHH family and its topological satellites, as well as other families of orbits (such as, though not limited to, the non-zero-angular-momentum BHH orbits [17–23] and by (some of) three non-zero-angular-momentum Davoust–Broucke orbits [44], and, last but not least, all of the zero-angular-momentum solutions in Refs. [33,37,42,43]).

Thus our space of initial conditions is nominally four-dimensional, with an additional (non-linear) constraint between the angular momentum and energy. The energy E is determined by

$$E = \frac{1}{2}\left(\mathbf{v}_\rho^2 + \mathbf{v}_\lambda^2\right) - \sum_{i<j}\frac{1}{r_{ij}} = \frac{1}{2}\left(c^2 + d^2\right) - \left(\frac{\sqrt{2}}{a + \sqrt{3}b} + \frac{\sqrt{2}}{|a - \sqrt{3}b|} + \frac{1}{\sqrt{2}a}\right) \tag{5}$$

(we take all three masses to equal $m = 1$, and set the gravitational constant $G = 1$, which sets the units in our system). The total angular momentum is given by

$$L = |\boldsymbol{\rho} \times \dot{\boldsymbol{\rho}}| + |\boldsymbol{\lambda} \times \dot{\boldsymbol{\lambda}}| = ac + bd,$$

both of which are constants of motion. As briefly mentioned above, one might think that these two integrals of motion could be used to effectively remove one of the three parameters. The energy is related to the overall size (hyper-radius R) by scaling rules, Section 2.3. So it will be used up when we fix the size see Section 3.3. The angular momentum, on the other hand, scales with size differently than the energy, so that, in effect, it remains an independent variable. Thus, our space of initial conditions of BHH orbits remains effectively three-dimensional, even when these two conservation laws are explicitly implemented.

3.3. *The search sub-space*

We use these four parameters (a, b, c, d), together with the following additional constraint to parametrize the search-space for periodic orbits. The "size" of the orbits must be fixed in order to avoid finding the same orbits, only rescaled by a size factor, see Section 2.3. Therefore, we fix the "size" by setting $b = 1$. Henceforth we shall not rescale the size R, or energy E of the system.[14]

Then, for any fixed value of L, we have $d = L - ac$, and with energy $E = \frac{1}{2}\left(c^2 + (L - ac)^2\right) - \left(\frac{\sqrt{2}}{a+\sqrt{3}} + \frac{\sqrt{2}}{|a-\sqrt{3}|} + \frac{1}{\sqrt{2}a}\right)$, a function of three independent variables: a, c and L. Now we are left with (only) a three-dimensional "phase sub-space", spanned by a, c and L, to be searched for periodic orbits. For a fixed value of L the search can be further restricted to certain regions of parameters a and c (see the following).

Firstly, as three equal masses imply permutation symmetry, one only needs to look at one permutation of the three bodies, thus allowing us to search only a limited set of configurations *viz.* starting from those in which the "inner body" is closer to the "left-hand-side" body, to the configurations when the "inner

---

[14] Only at the end of the day, one may change to scale-invariant period and angular momentum.

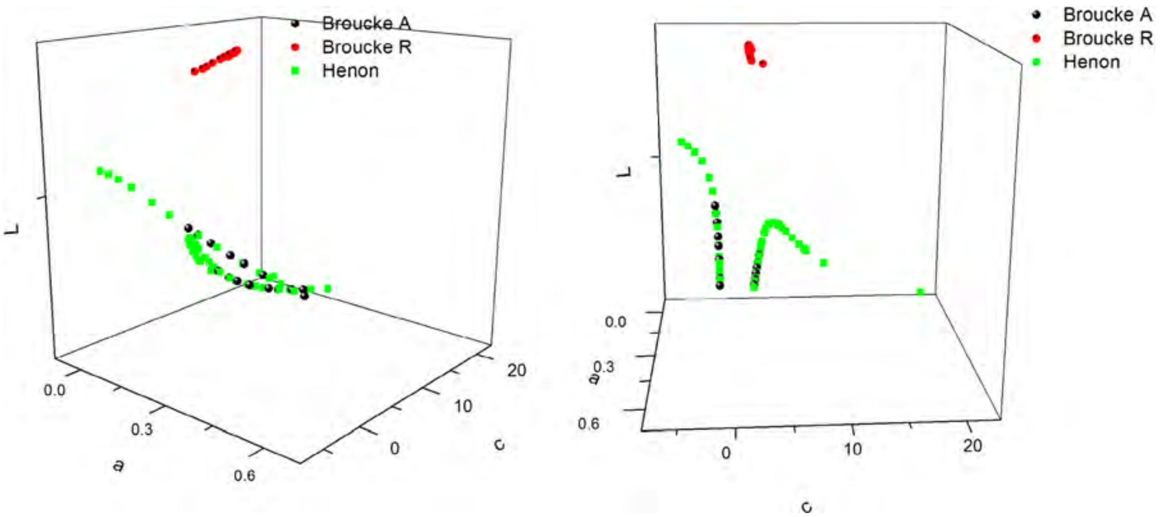


Fig. 4. Two views of the three-dimensional parameter space of the initial conditions of BHH orbits. Green points denote Hénon's (retrograde) orbits [22], black points are Broucke's "A family" (retrograde) orbits and the red points denote Broucke's "R family" (prograde) orbits [18]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

body" is exactly in the middle between the other two bodies, i.e., for $a \in (0, \frac{1}{\sqrt{3}}]$, we have $|a - \sqrt{3}| = \sqrt{3} - a$ and leads to

$$E = \frac{1}{2}\left(c^2 + (L - ac)^2\right) - \frac{1}{\sqrt{2a}} + \frac{2\sqrt{6}}{a^2 - 3}.$$

Secondly, we confine ourselves to the values of free parameters that lead to negative total energies, $E < 0$, so as to avoid escape of one of the bodies to infinity.[15] For example, having fixed the values of L and a, the allowed values of c lie within a circle defined by the inequality:

$$\frac{1}{2}\left(c^2 + (L - ac)^2\right) < \frac{1}{\sqrt{2a}} + \frac{2\sqrt{6}}{3 - a^2}. \tag{6}$$

This inequality has non-trivial implications: it divides the (a, c, L) space into disjoint regions, some of which are allowed, and others are not.

The inequality (6) can be rewritten in various ways, keeping one pair of parameters fixed while solving for the third one; thus it is quartic in a, and quadratic in c, and L. As some of its roots need not be real, or within the allowed regions of (real-valued) parameters the actual number of relevant roots may be anywhere between 0 and $4 \times 2 \times 2 = 16$, which determines the number of allowed, disjoint regions in the (a, c, L) space. E.g. depending on whether e.g. $c \in (c_-, c_+)$, or $c \in (0, c_-)$, $c \in (c_+, \infty)$

$$c_\pm = \frac{aL}{(a^2+1)} \mp \frac{1}{(a^2+1)} \times \sqrt{\frac{\left(a\left(\left(a^2-3\right)L^2 + \sqrt{2}\left(a\left(-a^2 + 4\sqrt{3}a + 2\right) + 4\sqrt{3}\right)\right) + 3\sqrt{2}\right)}{-a\left(a^2-3\right)}}$$

and, e.g. $L \in (L_-, L_+)$, or $L \in (0, L_-)$, $L \in (L_+, \infty)$, where

$$L_\pm = ac \pm \sqrt{\frac{\left(a\left(a\left(\sqrt{2} - ac^2\right) + 3c^2 - 4\sqrt{6}\right) - 3\sqrt{2}\right)}{a\left(a^2-3\right)}}$$

It should be clear that the number and form of these restrictions depends on the functional dependence $L(a, c)$ for any given family of periodic orbits. As these $L(a, c)$ functional dependences are generally not known, except in one (the BHH), or perhaps two cases (the figure-8 orbits), in the following we shall eschew a general analysis of all possible cases, such as the one above, but concentrate on the case of BHH orbits at hand.

Before we continue, a few comments are in order: (1) it should be clear that the present method can be readily modified to include distinct masses; (2) at first we shall apply this method to the BHH family, as that is the most thoroughly studied family of three-body orbits in the literature and the $a(L, c)_{BHH}$ is fairly well known; (3) the present method holds in general, and not only for the BHH family. It can and should be applied in the vicinity of all known (linearly) stable orbits with vanishing angular momentum.

3.4. The region of BHH orbits' initial conditions and the scanning method

As can be seen in Fig. 4 all of the previously known BHH orbits fall on two continuous "curves", the green–black one depicting the retrograde, and the red one the prograde orbits, in the three-dimensional space of initial conditions parametrized by L, a and c.[16] This is a consequence of Hénon's first theorem [16]. Therefore, we choose the vicinity of this structure as the starting point of our search. Our search was conducted at fixed values of angular momentum L, in a region of a two-dimensional subspace of initial conditions parametrized by a and c. The "scanning" method for numerical searching for relative periodic orbits, which is described in more detail in Ref. [34] consists of two steps: (1) a "brute force" scan that produces the "initial candidates"; and (2) use the "initial candidates" as starting points in a minimization method, such as coordinate descent, or gradient descent. The crux of the matter is to have a good idea where to look for candidates, which depends on the parametrization of the initial conditions. The resulting orbit is accepted as periodic if the minimized r.p.f. is sufficiently close to zero.

---

[15] This condition alone does not guarantee that there can never be an escape to infinity, but only that there will be none as long as there are no two-body collisions.

[16] N.B. The retrograde orbits' "one continuous" curve appears as two here due to our choice of the domain of the parameters a and c, and of symmetries; another choice might have given a single curve, which, however, would have been less auspicious for the purposes of orbit hunting. Similarly, the prograde and the retrograde curves would merge in the (extreme) limit of two "inner" bodies merging into one.

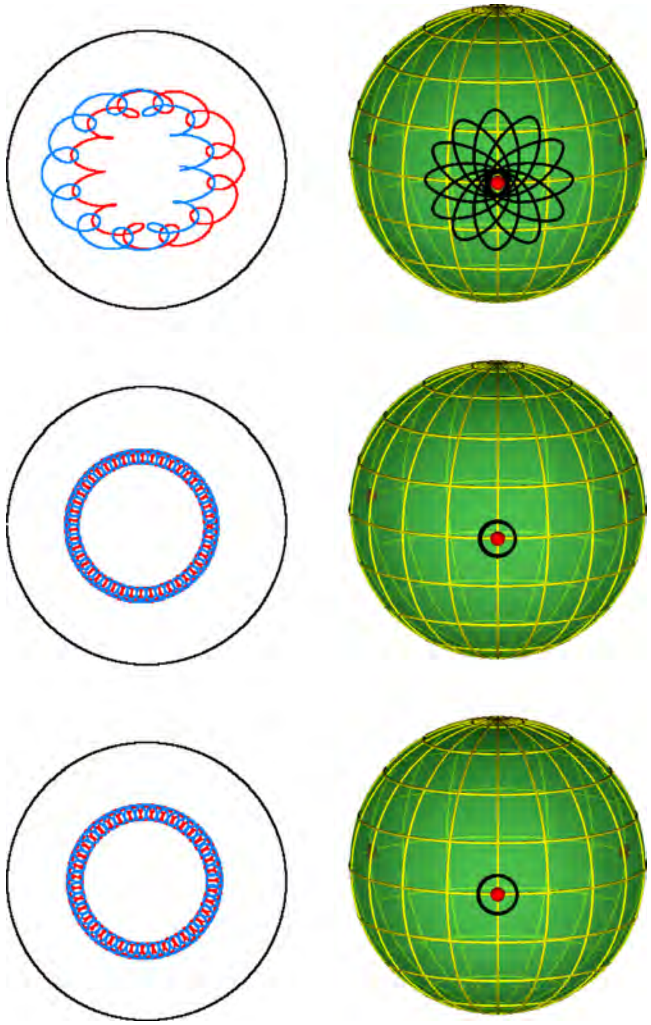


Fig. 5. Initially discovered satellite orbits of the BHH family, shown in real space and on the shape sphere, $k = 12, 45, 39$.

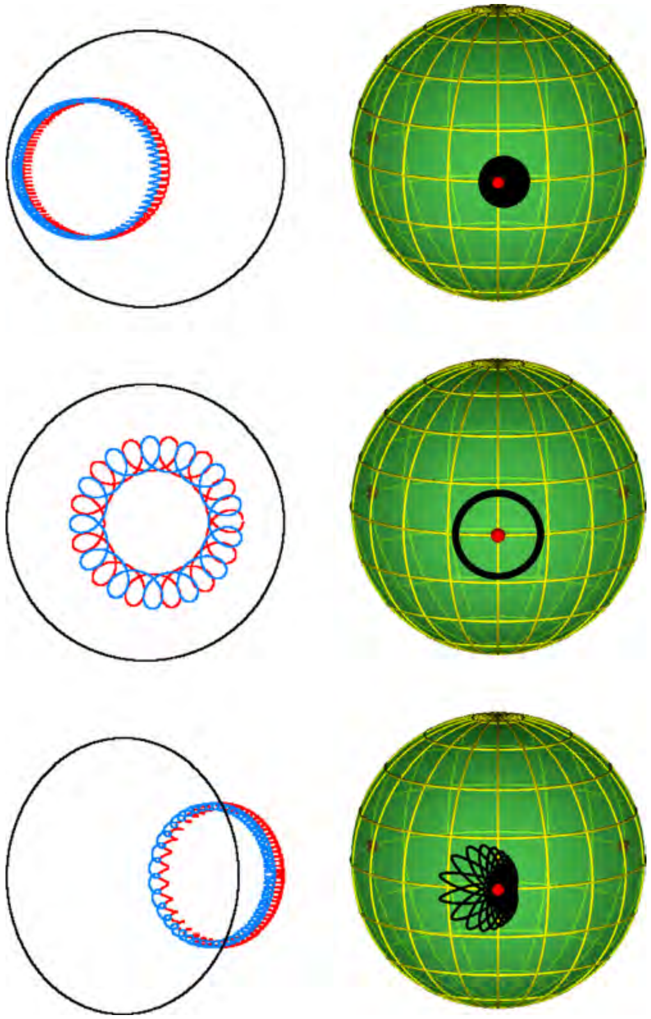


Fig. 6. Initially discovered satellite orbits of the BHH family, shown in real space and on the shape sphere, $k = 84, 13, 58$.

4. Results

Our search was conducted at fixed values of angular momentum L, in a region of a two-dimensional subspace of initial conditions parametrized by a and c. The "scanning" method of numerical searching for periodic orbits was briefly explained in Section 3.4 and described in pedagogic detail in Ref. [34].

We performed two preliminary searches at angular momentum $L = 0.935549$, where there is one (Hénon's) orbit, and at $L = 1.5$, where there are no BHH orbits. Parameters a and c took values in the following intervals: $a \in [0.05, 0.6]$, $c \in [-5.5, 5.5]$, although it was unnecessary to go above $a = 1/\sqrt{3} = 0.57735\cdots$, see Section 3.3. The resolution of the search was 1000×1000, which determined the duration of search. Naturally, one expects that longer searches would have produced more copious results, as longer periods would have been probed.

Some local minima of the return proximity function (r.p.f.) were extracted from the results and refined using additional minimization methods, which led to several new orbits, some of which were direct, whereas others were retrograde. After plotting the newly found orbits' trajectories on the shape sphere, it was easy to see that all of these orbits are topological powers of BHH orbits, the so-called topological satellites.

In the first two (preliminary) searches we found six BHH satellite orbits. These orbits make k loops about a single collision point on the shape sphere, with $k = 84, 12, 45, 39, 58, 13$ (see Figs. 5 and 6), some of which are retrograde and others are direct. One notices immediately the diversity of patterns on the shape sphere: some (e.g. panels 2, 3, 4, 6 in Fig. 5) are symmetric, whereas others (e.g. panels 1, 5 in Fig. 5) are asymmetric.

However, the level of detail in these results is very low: the resolution is not high enough to resolve some minima. Therefore we decided to perform such "scans" in smaller sections of the $a-c$ plane and in the vicinity of BHH orbits (see Figs. 7 and 8).

4.1. High resolution searches for BHH satellite orbits

The maps of the return proximity function at $L = 0.8, 0.85, 0.9, 0.935$ are shown in Figs. 7 and 8. Note the "interference-like" dark and bright regions (patterns) in this map. The brighter regions correspond to higher values of the negative logarithm of the return proximity function. Each bright (yellow) "stripe" in Figs. 7, 8 contains a particular satellite of topological order

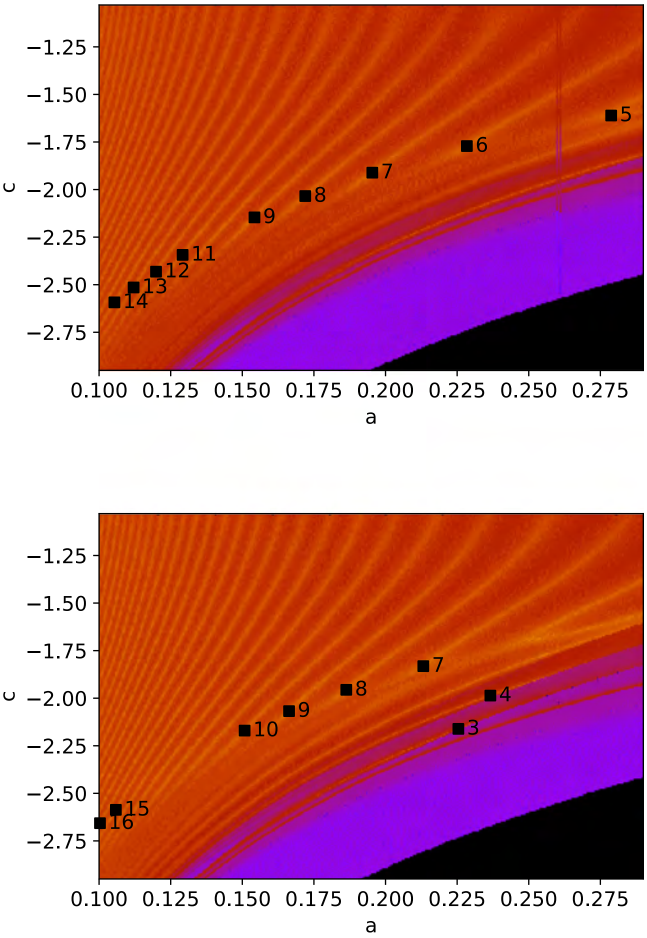


Fig. 7. The map of the negative logarithm of the return proximity function at angular momentum values $L = 0.8$ (upper panel), and $L = 0.85$ (lower panel). The brighter regions correspond to higher values of the negative logarithm of the return proximity function. Each black square denotes a local minimum of the return proximity function for which the value of the return proximity function is sufficiently close to zero: a satellite orbit, and the number is the k-value of that satellite. The black region in the lower right-hand corner is forbidden by the negative energy condition.

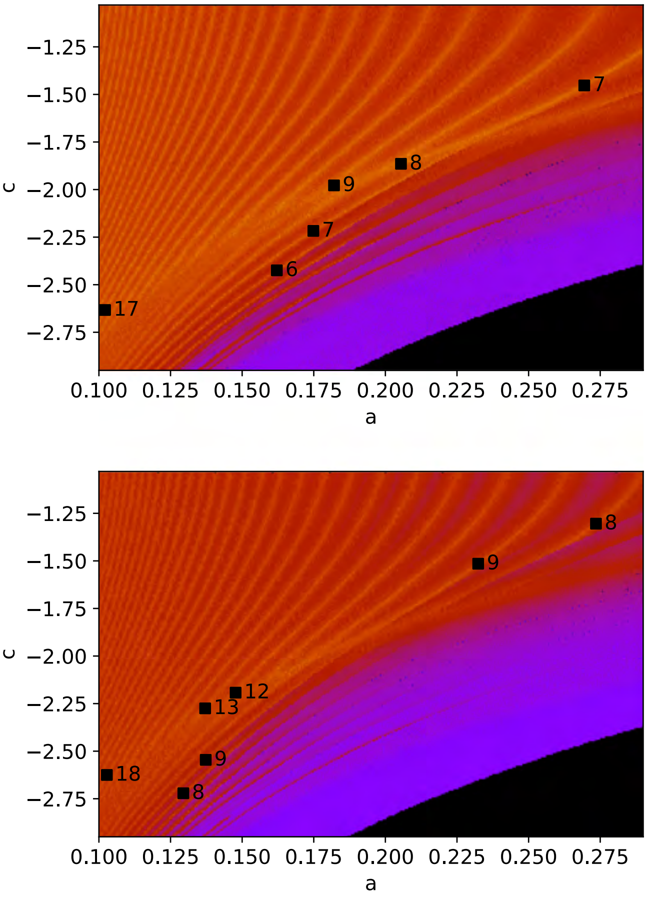


Fig. 8. The map of the negative logarithm of the return proximity function at angular momentum values $L = 0.9$ (upper panel), and $L = 0.935549$ (lower panel). The brighter regions correspond to higher values of the negative logarithm of the return proximity function. Each black square denotes a local minimum of the return proximity function for which the value of the return proximity function is sufficiently close to zero: a satellite orbit, and the number is the k-value of that satellite. The black region in the lower right-hand corner is forbidden by the negative energy condition.

k, arranged in an increasing order of k from the right to the left-hand side (see Figs. 10 and 11).

The same region in the $a - c$ plane was also explored at four other values of $L \in [0.8, 1]$. There we found satellite orbits with the same values of k, at slightly different initial conditions. For example, their trajectories on the shape sphere were wider or narrower, and for one of the two visually distinct types in Fig. 12, the satellites of the same k differ in eccentricity.

Some shape-sphere orbits (e.g. panels 1, 2, 3 in Fig. 12) are symmetric, whereas others (e.g. panels 4, 5 in Fig. 12) are asymmetric. In Fig. 12 one can also recognize that there are orbits with significant eccentricities of the inner and outer binaries, where these two kinds of binaries can (still) be recognized. In contrast, we challenge the reader to recognize any kind of binary in the satellite orbits shown in panel 1 of Fig. 12].

All of these satellites at the above-mentioned values of L are "arranged" in a way similar to those at $L = 0.9$, and in an overall manner/shape that are in accordance with Hénon's first theorem, i.e., with the expectation that there are continuous families (curves in the parameter space) of satellite orbits with a particular value of k, and most probably several different families/curves for the same k.

In addition to this, we searched for satellites with higher powers k, at higher values of angular momentum L, where we found examples of k up to 58.

4.2. Detailed searches near BHH retrograde orbits

Summary of detailed searches that were performed in the vicinity of BHH retrograde orbits is given in Table 1, with the searched segments of parameter space illustrated in Fig. 9. Comments made in the right-hand-most column of Table 1 indicate segments open to future improvement. One, particularly interesting segment is the low L-values, which we could not access due to lack of regularization of collisions in our code. Orbits in this region ought to have satellites, as well, assuming relevance of the Birkhoff–Lewis theorem to this system, see Section 6.1. A detailed study of this segment would constitute a test of this assumption.

The complete set of orbits is shown in the text below; the topological power k takes values $k = 3$–48, 58, 84, see Fig. 12. For more, see the web-site [38].

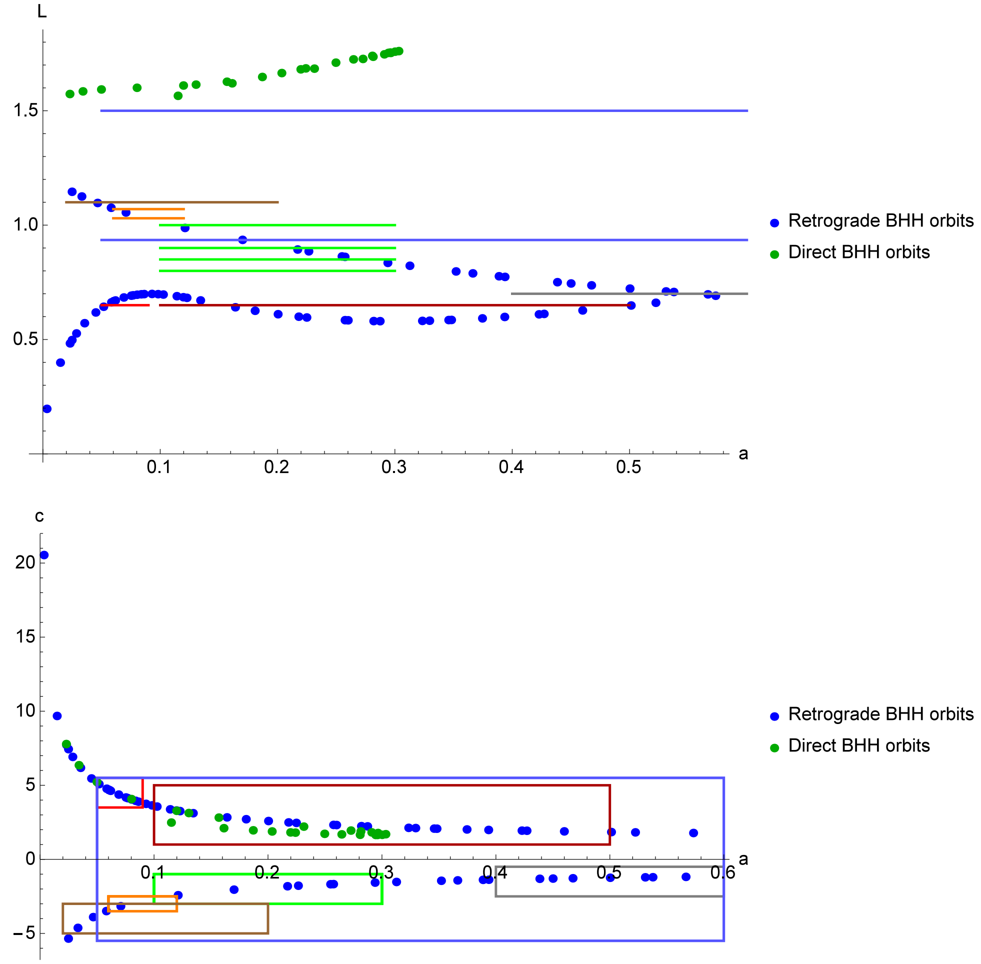


Fig. 9. Illustration of searched segments ("scans") in the parameter spaces (a, L) (upper panel), and (a, c) (lower panel). Each different colored line/box corresponds to a scan we performed in one a–c region, see Table 1, as follows. Upper panel (from top to bottom): blue (1.5)(0.935549), light brown (1.1), orange (1.03)(1.07), green (1.0)(0.9)(0.85)(0.8), gray (0.7), red(0.65), dark brown (0.65). Lower panel: red (0.65), dark brown (0.65), gray (0.7), orange (1.03)(1.07), light brown (1.1), green (0.935549), blue (1.5)(0.935549). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Searches ("scans") performed in segments of parameter space.

L	a_{min}	a_{max}	c_{min}	c_{max}	Comment
0.65	0.05	0.09	3.5	5.5	Few local minima
0.65	0.1	0.5	1	5	Several local minima
0.7	0.4	0.6	−2.5	−0.5	Unstructured set of local minima
0.8	0.1	0.3	−3	−1	Good results
0.85	0.1	0.3	−3	−1	Good results
0.9	0.1	0.3	−3	−1	Good results
0.935548917	0.1	0.3	−3	−1	Good results
1	0.1	0.3	−3	−1	Good results
1.03	0.06	0.12	−3.5	−2.5	Good results
1.07	0.06	0.12	−3.5	−2.5	Good results
1.1	0.02	0.2	−5	−3	Many local minima packed densely

5. Kepler-third-law-like regularities

In this section we follow Ref. [39] closely, for two reasons: (1) for the sake of completeness: Kepler-third-law-like regularities are perhaps the most exciting news that have emerged from the discovery of new (satellite) orbits and thus deserve a proper presentation; (2) in order to reveal our methods and way of thinking: we suspect that similar, though perhaps not exactly identical regularities may hold in other families of periodic orbits.

As stated in the Introduction 1, families of periodic three-body orbits can be characterized by the topology of their trajectories in the real configuration space ("braid group"), or on the (so-called) shape sphere ("free group"), as described in Refs. [32–34,51], the latter can be specified by the conjugacy classes of elements, or "words" w, for short, consisting of letters $\mathtt{a}$, $\mathtt{b}$, $\mathtt{A} = \mathtt{a}^{-1}$, $\mathtt{B} = \mathtt{b}^{-1}$, that define the free group on two letters ($\mathtt{a}$, $\mathtt{b}$).

Thus, here we must study the dependence of the constant on the right-hand-side of the scaling law $T(w)|E(w)|^{3/2} = \text{const}(w)$

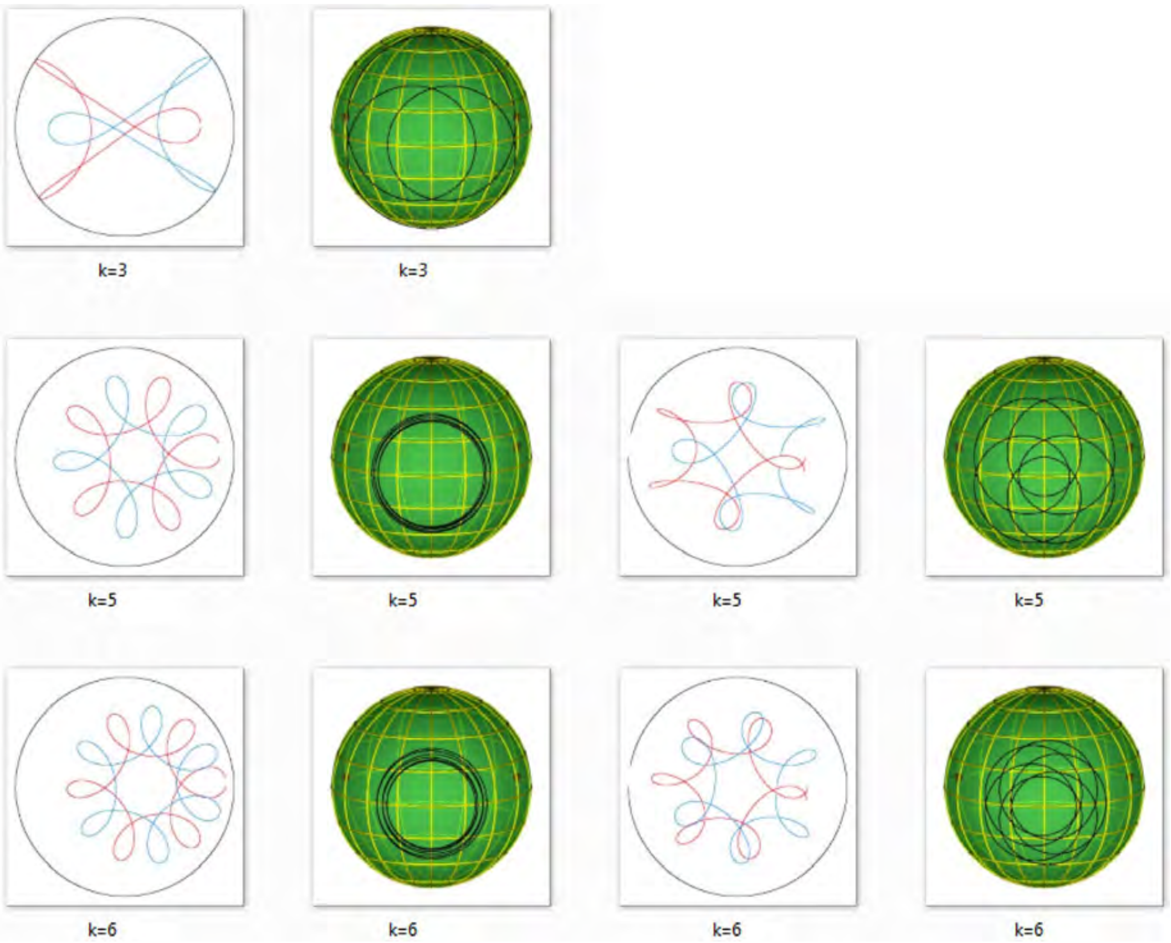


Fig. 10. Examples of satellite orbits of the BHH family, shown in real space and on the shape sphere, $k = 3, 4, 5, 6$. Each row contains two different satellites for each value of $k = 3, 4, 5, 6$, starting from the top, except for $k = 3$, where there is only one.

on the structure of the word $w(\mathtt{a}, \mathtt{b}, \mathtt{A}, \mathtt{B})$ that characterizes a periodic three-body orbit with zero angular-momentum. In Ref. [36] we have shown that Kepler's third law “constant” $T(w^k)|E(w^k)|^{3/2}$ of the “kth satellite orbit” with zero angular-momentum (specified by the free-group element w^k where k is an integer) of the “progenitor orbit” w equals k times the Kepler's third law constant $T(w)|E(w)|^{3/2}$ of the progenitor orbit w: $T(w^k)|E(w^k)|^{3/2} = kT(w)|E(w)|^{3/2}$. More simply, periodic orbits with zero angular-momentum normalized to a common energy E have periods related by $T(w^{k_i}) = k_i T(w)$. We wish to see if this, or some similar statement holds also at non-zero angular momentum?

Then the analogon of Eq. $T(w^k)|E(w^k)|^{3/2} = kT(w)|E(w)|^{3/2}$ for orbits with non-zero angular momenta is

$$\begin{aligned} T(w^k)|E(w^k)|^{3/2} &= f\left(L(w^k)|E(w^k)|^{1/2}\right) = kT(w)|E(w)|^{3/2} \\ &= kf\left(L(w)|E(w)|^{1/2}\right), \end{aligned}$$

or

$$\begin{aligned} L(w^k)|E(w^k)|^{1/2} &= f^{-1}\left(T(w^k)|E(w^k)|^{1/2}\right) = f^{-1}\left(T(w)|E(w)|^{1/2}/k\right) \\ &= L(w)|E(w)|^{1/2}\left(T(w)|E(w)|^{1/2}/k\right), \end{aligned} \tag{7}$$

or yet more simply, a relation between $L(T)$ curves for the progenitor orbit $L_r(T_r)$ and its k'th satellite $L_r^{(w^k)}(T_r^{(w^k)})$:

$$L_r^{(w)}(T_r^{(w)}) = L_r^{(w^k)}(T_r^{(w^k)}/k). \tag{8}$$

In the brief report [39] we tested this relation on the BHH family of solutions using the orbits presented in the previous section. In order to ensure a precise check of Eq. (8) it is necessary to precisely determine orbital periods. The periods T of the orbits could not be reliably established to better than the seven significant digits (decimal places), as shown in Table I in the Supplementary information for Ref. [39].

Our solutions are numerical, hence they necessarily contain some, small, but finite numerical “error”, i.e., difference between the particles' spatial positions after one period and their initial values. This error is perhaps best quantified by the value of the “return proximity function”, $d(\mathbf{Z}_0, T_0) = \min_{t \leq T_0} |\mathbf{Z}(t) - \mathbf{Z}_0|$, see Refs. [33,34], evaluated after one period $t = T$. The minimal values of $d(\mathbf{Z}_0, T_0)$, $d_{\min}$, for our solutions are typically of the order of $\mathcal{O}(10^{-10}) - \mathcal{O}(10^{-9})$, which also indicates the order of magnitude of the expected error in the values of kinematical variables, such as the period T. The relation between the expected error in the period T and the minimal “return proximity function”, $d_{\min}$, is not a linear one, however.

We have undertaken four independent evaluations (denoted by Roman capital numerals I–IV) of period T: two $T(\mathrm{I}) = T_{RKF_{rpf}}$, $T(\mathrm{III}) = T_{RKF_z}$), are based on the Runge–Kutta–Fehlberg (RKF) algorithm of fourth order, and another two ($T(\mathrm{II}) = T_{BS_{rpf}}$, $T(\mathrm{IV}) = T_{BS_z}$) are based on the Bulirsch–Stoer (BS) algorithm, each with two different definitions of the period T: T_{rpf} is based on the minimum of the return proximity function (rpf) and T_z is

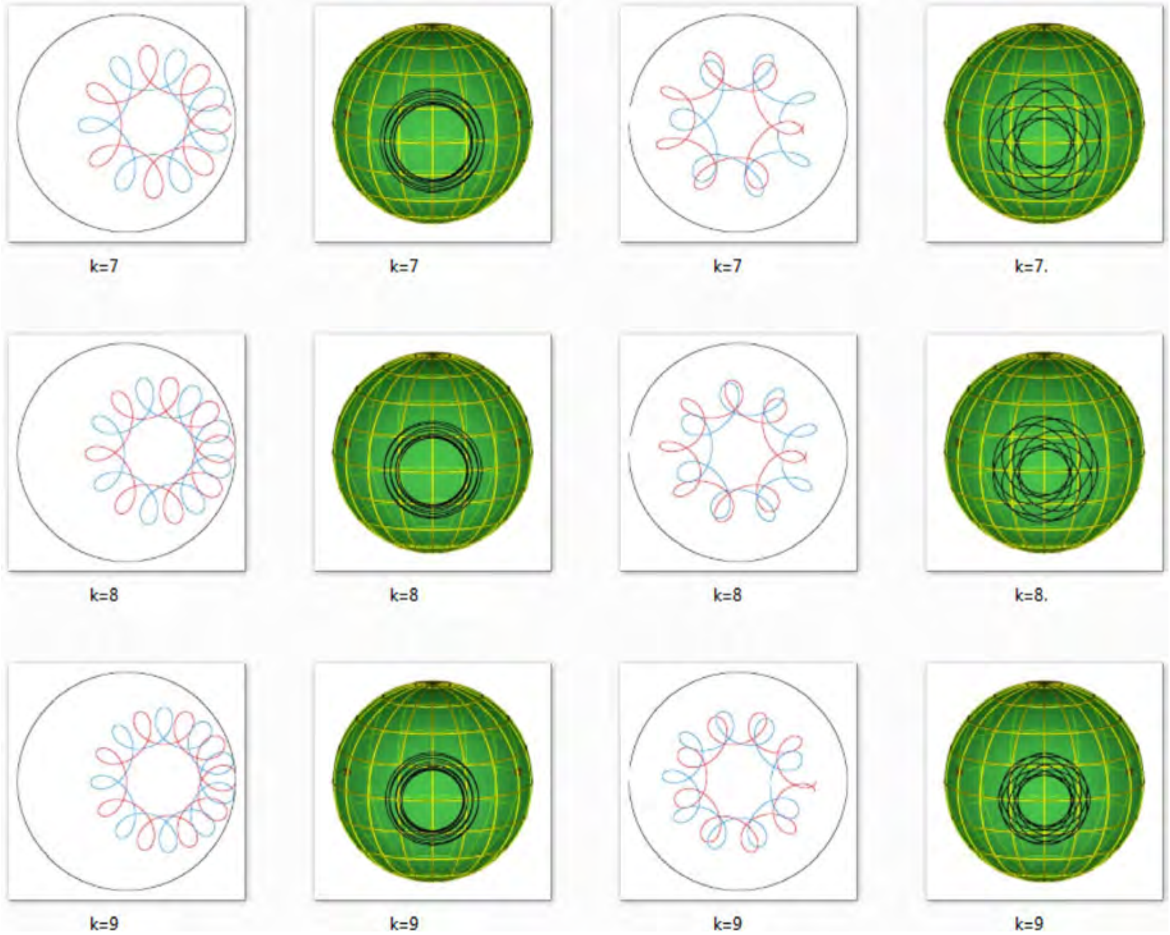


Fig. 11. Examples of satellite orbits of the BHH family, shown in real space and on the shape sphere, $k = 7, 8, 9$. Each row contains two different satellites for each value of $k = 7, 8, 9$, starting from the top.

based on the crossing of the equator on the shape sphere. These measurements may, but need not, agree in general. They must agree only when the initial conditions are "perfect", i.e., when the value of $d(\mathbf{Z}_0, T_0)$ is zero (which does not happen in actual numerical calculations).

5.1. *Observed topological dependence of the scaling laws for three bodies*

The $L(T)$ curves of different-k satellite orbits are scattered over a large region and do not "touch"/intersect the BHH progenitor family of orbits' $L(T)$ curve when plotted as a function of the (un-divided) period T, see Fig. 13. Note the huge span/scatter of periods in the data.

After dividing the period T by the topological exponent/index k, $T' = T/k$, we can see in Fig. 14 that the satellite orbits' $L(T/k)$ curve (the angular momentum as a function of topologically-rescaled period T/k) approximately coincides with the $L(T)$ curve of BHH retrograde orbits.

6. Open questions

Here we present a list of open questions related to our paper. They range from general questions about astronomical existence and some, perhaps abstract mathematical questions about the deeper underlying causes for these orbits, to entirely practical suggestions as to which specific subspace of i.c.s ought to be explored in a numerical search.

6.1. *Astronomical and mathematical questions*

Our study also opens up several mathematical and astronomical questions:

1. The BHH family is one of only two families, another being the Lagrange one, of periodic three-body orbits that have been observed in the skies: all "hierarchical" triple star systems belong to BHH orbits, though the converse is not true. The Sun–Earth–Moon system may be viewed as a BHH solution [5,17,20], with unequal masses. Therefore, it seems important to extend the present study to the case of three different masses: some work in this direction has already been done in Refs. [17,20,55], but more needs to be done, and our method lends itself to the task. A number of 3-body systems have been discovered by the *Kepler* space telescope, all of which are of the hierarchical type, see §5.2 and §5.4 in Ref. [5].[17] Are there BHH topological satellites among astronomically observed three-body systems?
2. We have observed satellite orbits only in the stable region of BHH progenitor orbits' $L(T)$ curve: why? What precisely is the relation between satellites' existence and stability of

---

[17] see §5.2 Searches for Exomoons in Ref. [5]: "There are many dynamical processes that affect circumbinary planets, but we will focus on Kepler-16b, the first circumbinary planet confirmed by the Kepler mission (Doyle et al. 2011). This was a huge discovery because previous circumbinary planets had been posited through the post-common envelope binaries (e.g., Beuermann et al. 2010)".

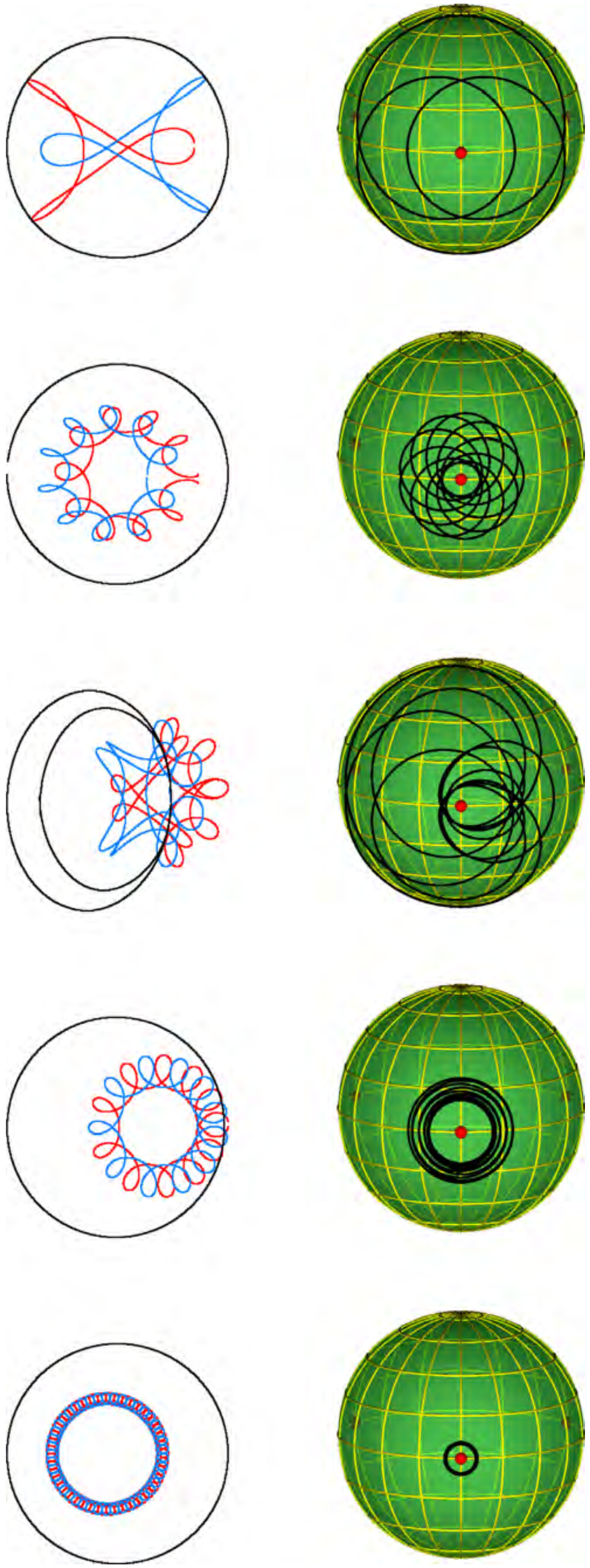


Fig. 12. Overview of some of the satellite orbits in the BHH family, $k = 3, 10, 10, 11, 43$.

progenitor's orbit? In 1976 Hénon [22] stated that "... the stable periodic orbits which we have found are very probably surrounded by a region of finite measure in phase space in which the orbits possess non-linear stability", i.e., that the linearly stable orbits in the BHH three-body family are also nonlinearly, or perpetually, or Kolmogorov–Arnold–Moser (KAM) stable. In other words, the KAM theorem [57–59] is believed to hold for stable BHH orbits. From the KAM theorem it follows that there must be infinitely many "conditionally periodic" solutions near non-linearly stable BHH orbits. In Ref. [37] it was noted that, according to the Birkhoff–Lewis theorem [49], even the (weaker) linear stability may lead to an infinity of periodic satellite orbits. It would be good to verify this conjecture.

3. Prograde orbits are unstable at these values of L, which makes them unlikely to have satellites, according to the Birkhoff–Lewis theorem. One ought to perform further searches in the parameter space at higher values of angular momentum L before one can draw any conclusions about the (non)existence of satellite orbits and their properties on the $L(T)$ plot.
4. In recent years there has been progress in providing formal "proofs of existence" for some BHH orbits, Refs. [26,27]. The obvious question is: can one "prove the existence" of the satellite orbits, and when?
5. Several different types of BHH satellites with identical values of k were reported in Ref. [39]. The question naturally arises: just how many such satellites are there for each value of k?
6. The above point (2) would account for the existence of satellite orbits, though not for the relation between their periods and topologies. Some suggestions about the cause of this relation based on the (complex-variable) analytic properties of the action integral were presented in Appendix E of Ref. [37], but it would be good to make those arguments rigorous, or to repudiate them (for example by finding counterexamples).

6.2. *Suggestions for future numerical work*

If true, the aforementioned Hénon's conjecture implies existence of infinitely many satellite orbits for all linearly stable three-body orbits, of which there are many. Of course, one may object that a search for all such orbits must be without an end.

However, there is by now sufficient reason to believe that there is some, perhaps not fully revealed as yet, structure in the spectrum of periodic 3-body orbits (*viz.* discrete multiples of periods, relation(s) between periods and topology), and the goal would be to reveal this structure to the fullest extent possible.

Thus, the goal of a search would/ought to be to find (all) linearly stable orbits with the periods shorter than some (finite upper) bound (i.e., with the simplest topologies) — whereas the subsequent extension of sequences (generated by such short orbits — progenitors) towards infinity may indeed be pointless.

6.2.1. *Extension of zero-angular-momentum solutions to nonzero values*

There are around 20 linearly stable, out of grand total around 200 zero-angular-momentum orbits in Ref. [37] and at least 23 among roughly 2000 orbits from Refs. [35,40,42,43,60].

As explained above, all linearly stable orbits deserve a thorough investigation in the sense of extending them to non-zero angular momenta. Reference [40] commands special attention, because it is the only work that has reported periodic orbits outside of the i.c. subspace defined in Refs. [32,37].

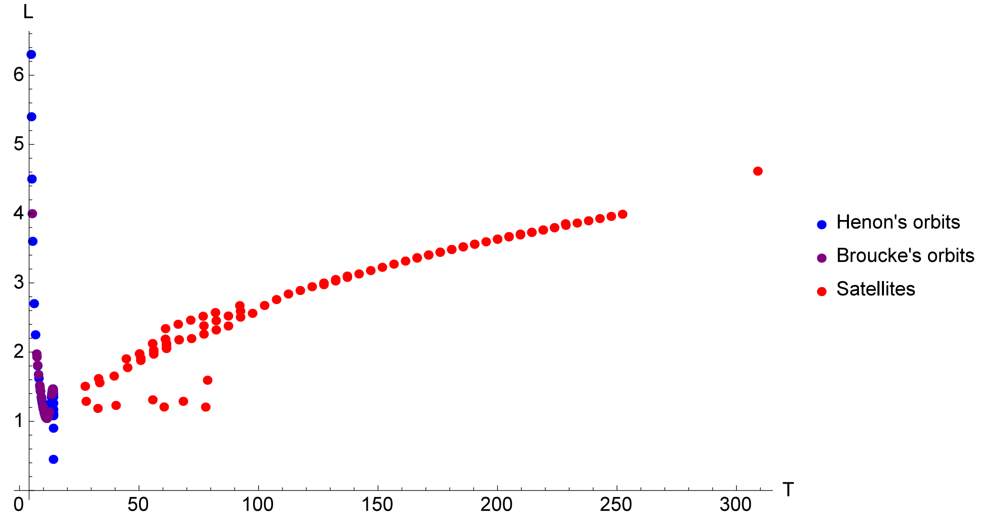


Fig. 13. Retrograde BHH orbits and the BHH satellite orbits with various values of k, discovered thus far, $L(T)$ dependence at fixed energy $E = -0.5$. Adapted from Ref. [39].

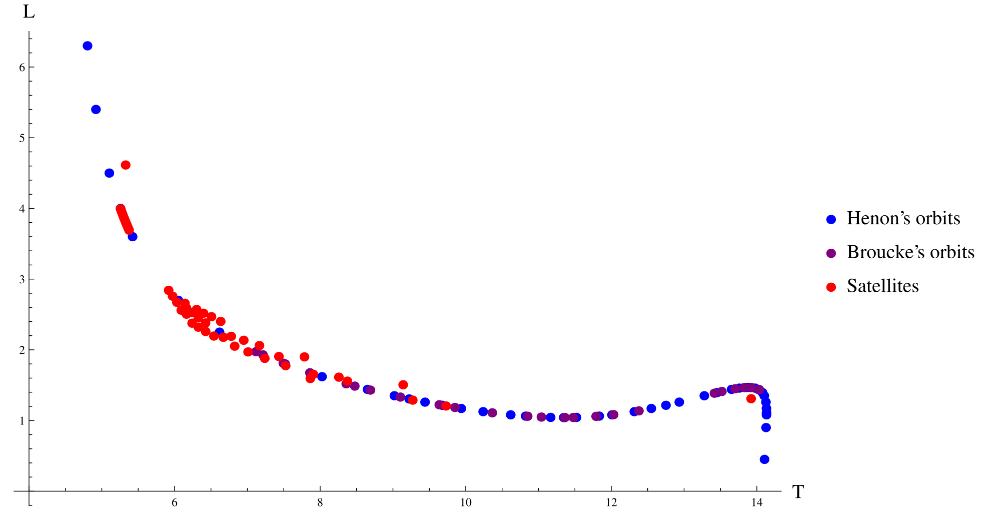


Fig. 14. Retrograde BHH orbits (black and blue dots) and their satellites (red dots) with various values of k, discovered thus far, $L(T' = T/k)$ dependence at fixed energy $E = -0.5$. Adapted from Ref. [39]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Last, but not least, the same method can be applied to 100, or so, zero angular momentum orbits in the Coulomb potential [50].

6.2.2. Searches around (other) known nonzero-angular-momentum orbits

Of course, search for other satellites can and should be continued in the BHH family, specifically in the prograde orbit branch, where only a few satellites have been found thus far, as well as in the retrograde branch, near the Schubert orbit, where further satellites are expected, but regularization is required [61–63].

There are other known, linearly stable orbits that fall into the present class of initial conditions: Davoust and Broucke reported a number (136) of (generally topologically unidentified) nonzero-angular-momentum orbits in Ref. [44]. All of these orbits' initial conditions are described by four parameters, just as ours, and a number of these orbits are linearly stable. Unfortunately, the tables of i.c.s in Ref. [44] are (very) difficult to read (due to a bad font) – which prevented us from examining them all[18] – here we discuss only those orbits which we managed to reconstruct.

These orbits were not classified into families according to their topologies, but, rather, using Strömgren's (restricted 3-body problem) classification into 14 *simple symmetry* (denoted by small roman and greek letters) and 11 *double symmetry* families (denoted by capital letters).

Many of these Davoust–Broucke (Strömgren) families fall into one of only three topological families: (1) The (Strömgren) double-symmetry A_1, B and single-symmetry a, b, c, d, α, β families contain orbits with the same topology as the Lagrange–Euler solutions, i.e., the identity element/zero of the free group; (2) The

---

[18] And when we read them, we could not be sure that we did it correctly.

Table 2
The initial conditions in terms of parameters A, B, C, D, as defined in the text, for three semichoreographic orbits. T is the period, L the angular momentum, and E is the energy. The shortest period orbit is Moore's [28].

A	B	C	T	$T\|E\|^{3/2}$	$L\|E\|^{1/2}$	E
0.6789245	−1.992169	1.3677159	3.147316	3.6087	−1.7668	−1.0955
0.2369355	−0.720445	1.2540211	8.004531	13.700	0.131881	−1.4308
0.400045	1.415959	1.1166540	8.413701	22.945	6.14072	−1.9519

Table 3
Satellite orbits in the retrograde BHH family, some of which are shown in Fig. 12. Minimal return proximity $d_{\min}$ for these orbits is $d_{\min} < 10^{-7}$. k is the topological power of the orbit, T is its period and E its energy.

N_r	L	a	c	T	k	$d_{\min}$
1	0.7	0.427052524289	−1.336907801590	4.46125383	3	1.12E−10
2	0.85	0.225423376709	−2.161172667330	4.93326124	3	1.72E−10
3	0.65	0.226748054608	2.494853501883	5.38634911	3	3.22E−11
4	0.85	0.236597473885	−1.986209145030	4.79387855	4	1.09E−10
5	0.7	0.410445264670	−0.901755581763	3.04224052	4	1.16E−10
6	0.65	0.089473243424	3.863464841380	6.80175586	4	3.88E−10
7	0.65	0.351891702719	1.961519344674	5.95736952	4	1.26E−10
8	0.8	0.278773932080	−1.610811894950	4.47639401	5	2.26E−10
9	0.65	0.131937959644	3.165458555994	8.78754480	5	6.12E−08
10	0.8	0.228399714670	−1.771422978980	4.12780456	6	1.12E−10
11	0.9	0.162112751455	−2.424555946410	4.72091689	6	1.09E−10
12	0.65	0.335276789538	2.040219149776	9.43988995	6	1.38E−10
13	0.8	0.195389423297	−1.911145472440	3.90877460	7	2.84E−10
14	0.85	0.213186587101	−1.831539642360	4.45339868	7	1.61E−10
15	0.9	0.174854596011	−2.217240443720	4.74829390	7	9.79E−11
16	0.9	0.269439116363	−1.452760161970	4.74215052	7	1.87E−10
17	0.8	0.171971993864	−2.034163513840	3.75598535	8	1.38E−10
18	0.935548917	0.129471314426	−2.721144023250	4.76584900	8	1.01E−10
19	0.7	0.537026752182	−1.208756213130	15.23729361	8	1.49E−10
20	0.85	0.186299773074	−1.955937487530	4.23559872	8	1.45E−10
21	0.9	0.205445523859	−1.864664426590	4.90547644	8	1.26E−10
22	0.935548917	0.273518188668	−1.304218790410	4.75908509	8	1.34E−10
23	0.8	0.154250567982	−2.146266402400	3.64212942	9	3.05E−10
24	0.935548917	0.232402133831	−1.514749892810	4.78046677	9	9.95E−11
25	0.85	0.166323483739	−2.067914324720	4.07742799	9	2.21E−10

(Strömgren) double-symmetry family A_2 contains quasi-isosceles orbits (with angular momentum); (3) The (Strömgren) double-symmetry families D_1, E, G and single-symmetry g, h, i are (direct) parts of the BHH topological family; (4) The (Strömgren) double-symmetry D_2, D_3, D_4, F and single-symmetry e families are (retrograde) parts of the BHH topological family, in particular solutions no. 90 and no. 91 are the $k = 3$ satellites.

6.2.3. *Semichoreographies with nonzero-angular-momentum*

Davoust and Broucke [44] were apparently the first ones to find a semichoregraphic solution — a periodic 3-body orbit wherein two bodies move on the same trajectory, whereas the third one moves on its own. In Ref. [28] Moore rediscovered Davoust and Broucke's orbit. We have found two other such orbits, see Table 2, and called them semichoreographies; each (semi)choreography defines a continuous family of orbits, as a function of angular momentum, whose orbits are not (necessarily) semichoreographies themselves, and which have not been explored thus far, to our knowledge.

We list several such orbits' i.c.s in Table 2, with the following definition of i.c. parameters A, B, C, D in terms of our parameters a, b, c, d,

$$a = \sqrt{2}, \quad b = -\sqrt{\frac{2}{3}}A, \quad c = \sqrt{\frac{1}{2}}(B - C), \quad d = \sqrt{\frac{3}{2}}(B + C).$$

6.2.4. *Isolated orbits*

Last, but not least there are a number of isolated, generally topologically unidentified periodic orbits, often unpublished, or published only in Ph.D. or M.Sc. theses, and/or on the internet; here we list the ones we knew at the time of writing.

- The "Celtic knot" choreographic orbit of Montaldi & Steckles [64] (no i.c.s supplied), which appears to be equivalent to the "Rosette" orbit of Grant [56] with nonvanishing angular momentum, the i.c.s are supplied in footnote [91]. This orbit does not fit into the (sub)space of collinear orthogonal i.c.s, for proof, see footnote,[19] and therefore requires relaxation of conditions imposed in Section 3.2.
- Danya Rose's many previously unknown orbits with vanishing angular momentum [40], some of which do not pass through an equidistant collinear ("Eulerian") initial configuration.
- A number of as yet topologically unidentified orbits presented in Refs. [55,65] and references therein.

All of this indicates: (1) a need to complete the families with other non-zero angular momentum orbits; (2) a probable abundance of new satellite orbits waiting to be discovered.

Free-fall orbits generally do not satisfy $\dot{R} = 0$ at equator crossings,[20] so generally they do not fall into the present search space.

---

[19] In order to check if its i.c.s fall into the class of collinear orthogonal, i.e., if $\mathbf{r}_i \cdot \dot{\mathbf{r}}_i = 0$, (for all $i = 1, 2, 3$), we note that the first of six syzygies is reached after $t = 0.7032783$, where we have $(x_1(t), y_1(t), x_2(t), y_2(t), x_3(t), y_3(t)) = (0.00077832, -0.469023, \ -0.00165945, 0.999997, \ 0.000881126, -0.530975)$, which still has 4 independent variables. Of course, two of these four can be eliminated by an appropriate rotation of the reference frame. Similarly, the velocities at $t = 0.7032783$ are $(\dot{x}_1(t), \dot{y}_1(t)) = (-3.03236, -2.77587)$; $(\dot{x}_2(t), \dot{y}_2(t)) = (0.385354, 0.00774803)$; $(\dot{x}_3(t), \dot{y}_3(t)) = (2.647, 2.76812)$, of which there are 4 independent ones. Evaluating the scalar products, which are rotation-invariant, we see that $\mathbf{r}_i(t) \cdot \dot{\mathbf{r}}_i(t) \neq 0$ (for all $i = 1, 2, 3$), i.e., this is not an orthogonal collinear configuration. Nevertheless, this orbit passes through three reversible configurations of another kind: the isosceles one, see Ref. [65].

[20] With the exception of isosceles triangle ones.

Table 4
Satellite orbits in the retrograde BHH family — Table 3 continued.

N_r	L	a	c	T	k	$d_{\min}$
26	0.9	0.182069791972	−1.978004135720	4.67040604	9	1.17E−10
27	0.935548917	0.137285145946	−2.545654735050	4.78370980	9	9.67E−11
28	1	0.294303286736	−1.008322699800	4.76801186	9	1.23E−10
29	1	0.238218402625	−1.735467337160	6.99274465	9	1.40E−10
30	0.85	0.150808938415	−2.170190077810	3.95629338	10	1.48E−10
31	1	0.209243455808	−1.847936324290	6.46969425	10	1.57E−10
32	1	0.266925744815	−1.097311225480	4.78597293	10	1.23E−10
33	0.7	0.442402892100	−0.700265953090	6.66684025	10	2.26E−09
34	0.8	0.129160165070	−2.343250256240	3.48168164	11	1.12E−10
35	1	0.243617171755	−1.190238664360	4.79985962	11	1.09E−10
36	0.8	0.119882517495	−2.431498122390	3.42239489	12	1.27E−10
37	0.935548917	0.147716034231	−2.192473293620	4.68169607	12	1.18E−10
38	1	0.223340706817	−1.288029535680	4.81093626	12	2.30E−10
39	0.8	0.112051269923	−2.514553762900	3.37232423	13	1.43E−10
40	0.935548917	0.137088743692	−2.274985930540	4.56392294	13	1.52E−10
41	1	0.205347639306	−1.392000632830	4.81998871	13	1.62E−10
42	0.8	0.105352054293	−2.592836811550	3.32933694	14	1.97E−10
43	1	0.146284147722	−2.203021298400	5.44909516	14	1.06E−10
44	0.85	0.105854763654	−2.586702862280	3.60999927	15	1.05E−10
45	1	0.136829854037	−2.277089258430	5.30678948	15	2.09E−10
46	1	0.173982588477	−1.627461935580	4.83391624	15	5.74E−09
47	0.85	0.100315012262	−2.656853706660	3.56736722	16	9.88E−11
48	0.9	0.102148274757	−2.632996062570	3.90923742	17	1.52E−10
49	0.935548917	0.102773178918	−2.625007735230	4.19114405	18	1.04E−10
50	1	0.110009104810	−2.537603649510	4.91429079	19	1.42E−08

7. Summary, conclusions

Here we have presented details of our method, originally designed to search for periodic orbits within the Broucke–Hadjidemetriou–Hénon (BHH), Refs. [17–23], family of orbits. However, we have realized that the method has a (much) wider scope.

We have numerically found 99 new satellite orbits in the family of BHH relative periodic solutions to the planar three body problem, and analyzed their properties and compared them with the properties of the original BHH orbits. An approximate relation between their kinematic and topological properties was reported in Ref. [39].

BHH orbits form a family with a very simple topology, and their satellites are orbits with topology that can be described as the kth power of BHH one. The BHH orbits' scale-invariant angular momenta L and scale-invariant periods T form a continuous curve $L(T)$, whereas our satellite orbits form a scattered set of points on the same $L(T)$ plot. The latter exhibit the property that when their period T is divided by their "topological power" k, they approximately fall on the $L(T)$ curve of the original ($k = 1$) BHH orbits. The deviation from exact identity of the two curves, can be quantified by a mean-square-root deviation of the observed satellite-orbit data from a fit to the BHH progenitor-orbit data.

Our study was motivated by the discovery of satellite orbits at vanishing angular momentum and of the proportionality of their scale-invariant period to their topological power [33]. The Kepler-like topological regularities have been found to hold more generally in sequences of orbits, albeit thus far only at vanishing angular momenta [36,37]. This report shows, however, that this regularity persists even when orbits with $L \neq 0$ are considered, albeit approximately, i.e., within some tolerance.

These results are even more striking when one remembers that there are several distinct types of satellite orbits of the same topological power k, some with quite different values of L and T, all of which display this property. Furthermore, more than one satellite of the same power k and the same type have been found for several progenitor orbits presented in this report; our results (not shown here) suggest that satellites form continuous curves in the parameter space of initial conditions. That "fine structure" should be investigated in greater detail, however, as in the examples set by Davoust and Broucke [44].

Table 5
Satellite orbits in the retrograde BHH family — Tables 3 and 4 continued.

N_r	L	a	c	T	k	$d_{\min}$
51	1.03	0.111843109779	−2.516815645433	5.31721934	20	9.76E−11
52	1.03	0.106999360702	−2.572864154118	5.23675314	21	1.10E−10
53	1.03	0.102629340641	−2.626824252565	5.16461551	22	1.04E−10
54	1.07	0.112532299117	−2.509130097108	5.93058960	22	9.37E−11
55	1.03	0.098663558847	−2.678884782752	5.09949046	23	1.13E−10
56	1.07	0.107950445783	−2.561552366744	5.83679015	23	1.62E−10
57	1.03	0.095045815614	−2.729207988143	5.04033295	24	1.24E−10
58	1.07	0.103792166702	−2.612125153377	5.75238033	24	9.83E−10
59	1.07	0.099998313989	−2.661009860717	5.67592271	25	1.93E−10
60	1.07	0.096520577947	−2.708345753387	5.60626742	26	1.77E−10
61	1.07	0.093319072567	−2.754253774713	5.54248254	27	1.28E−10
62	1.07	0.090360532835	−2.798839977984	5.48380387	28	1.27E−10
63	1.07	0.087616970029	−2.842197991476	5.42959820	29	1.09E−10
64	1.03	0.078546062730	−3.001449136918	4.77275161	30	1.18E−10
65	1.07	0.085064647578	−2.884410942736	5.37933595	30	1.04E−10
66	1.03	0.076424987284	−3.042734364817	4.73851381	31	1.48E−10
67	1.07	0.082683286816	−2.925553119986	5.33257043	31	1.02E−10
68	1.03	0.074436173120	−3.083042211579	4.70642108	32	1.04E−10
69	1.07	0.080455449983	−2.965691193228	5.28892198	32	9.96E−11
70	1.03	0.072567055606	−3.122428961620	4.67626201	33	1.51E−10
71	1.07	0.078366051119	−3.004885281315	5.24806564	33	1.39E−10
72	1.03	0.070806647363	−3.160945841855	4.64785297	34	9.95E−11
73	1.07	0.076401964229	−3.043189884203	5.20972146	34	1.05E−10
74	1.03	0.069145295338	−3.198639607087	4.62103356	35	1.66E−10
75	1.07	0.074551712461	−3.080654510153	5.17364685	35	1.59E−10

From the methodological point of view, we have shown that a systematic search for periodic solutions is possible in a three-dimensional subspace of initial conditions, although it is more challenging and time-consuming than in the case of orbits with vanishing angular momentum. We have found satellite orbits up to topological power $k = 58$ (84); but we feel that we have not found all satellite orbits with values $k \leq 58$ (84), despite there being no known theorem stating how many orbits there ought to be. It should be noted that our search was time-limited: the same method can be used without modifications to complete the

Table 6
Satellite orbits in the retrograde BHH family — Tables 3–5 continued.

N_r	L	a	c	T	k	$d_{\min}$
76	1.03	0.067574480868	−3.235553106562	4.59566306	36	1.02E−10
77	1.03	0.066086655842	−3.271725713345	4.57161742	37	1.18E−10
78	1.07	0.071153562285	−3.153240705933	5.10748717	37	1.01E−10
79	1.03	0.064675107247	−3.307193655949	4.54878679	38	1.05E−10
80	1.07	0.069588876352	−3.188441664622	5.07705416	38	1.04E−10
81	1.1	0.072454170036	−3.124853635110	5.53965642	39	1.05E−10
82	1.03	0.063333841711	−3.341990448546	4.52707358	39	1.68E−10
83	1.07	0.068104140192	−3.222962078095	5.04818753	39	1.04E−10
84	1.1	0.070912224143	−3.158592700200	5.50480009	40	9.94E−11
85	1.03	0.062057492484	−3.376147044334	4.50639068	40	1.13E−10
86	1.1	0.069446525539	−3.191702128170	5.47170776	41	1.07E−10
87	1.03	0.060841236077	−3.409692239688	4.48666015	41	1.03E−10
88	1.07	0.065350113732	−3.290088012685	4.99465663	41	9.97E−11
89	1.07	0.065350112192	−3.290088080655	4.99465659	41	1.15E−10
90	1.1	0.068051285222	−3.224210823950	5.44023825	42	1.04E−10
91	1.03	0.059680728422	−3.442652640917	4.46781196	42	1.24E−10
92	1.1	0.066721298871	−3.256145631480	5.41026552	43	8.88E−11
93	1.07	0.069589283429	−2.997571148119	4.88748351	43	1.26E−10
94	1.1	0.065451870057	−3.287531720240	5.38167667	44	1.02E−10
95	1.1	0.064238759898	−3.318392179570	5.35437026	45	1.03E−10
96	1.1	0.063078117801	−3.348748949910	5.32825481	46	2.14E−10
97	1.1	0.061966455334	−3.378621962290	5.30324770	47	9.76E−11
98	1.1	0.060900580973	−3.408030582910	5.27927394	48	9.86E−11
99	0.93555	0.061515814146	1.086851979730	2.57459301	58	1.96E−10

search. We concentrated on a search for satellites of retrograde BHH orbits, yet in this process we inadvertently found (only) four satellites of direct BHH orbits.

There is no reason to believe, however, that a finite maximum value of k exists. As k increases the satellite orbits seem to be packed more densely, however, so the search for higher values of k will be limited by inevitable numerical inaccuracies.

Our method can be also used without modification to extend this search to higher angular momenta and to direct BHH orbits. An extension of our search into unexplored regions of the $L - T$ plane ought to provide (new) data that will further test our hypothesis.

Last, but not least, the same method can be used to search for non-vanishing angular momentum families of orbits other than the BHH one, as explained in Section 6.2. That, of course, would represent a major new research program.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We wish to thank the anonymous referee for useful comments. The present paper is based on M.R.J.'s M.Sc. thesis, defended in June 2016 at Belgrade University and partially published as Ref. [39]. M.R.J. was a recipient of the "Prof Dr Djordje Živanović" scholarship for the academic year 2013/14, awarded jointly by the Faculty of Physics and the Institute of Physics, Belgrade University, and was also supported by a City of Belgrade studentship, Serbia (Gradska stipendija grada Beograda) in 2013, as well as the "Dositeja" stipend from the Fund for Young Talents of the Serbian Ministry for Youth and Sport for the academic year 2015/2016 (Fond za mlade talente — stipendija "Dositeja"). M.R.J. acknowledges current support from the President's Ph.D. scholarship of the Imperial College London, UK. The work of M.Š. and V.D. was supported by the Serbian Ministry of Science and Technological Development under grant numbers OI 171037 and III 41011.

Appendix A. Tables of initial conditions

For the sake of reproducibility, here we list the initial conditions of our 99 orbits.

A.1. Arbitrary k satellites

Tables 3, 4, 5, 6 show initial conditions for un-normalized orbits.

References

[1] C. Marchal, The Three-Body Problem, Elsevier Science Publishers, Amsterdam, 1990.
[2] Mauri Valtonen, Joanna Anosova, Konstantin Kholshevnikov, Aleksandr Mylläri, Victor Orlov, Kiyotaka Tanikawa, The Three-Body Problem from Pythagoras To Hawking, Springer International Publishing, Switzerland, 2016.
[3] M. Valtonen, H. Karttunen, The Three-Body Problem, Cambridge University Press, 2005.
[4] Kenneth R. Meyer, Glen R. Hall, Dan Offin, Introduction To Hamiltonian Dynamical Systems and the N-Body Problem, second ed., Springer, New York, 2009.
[5] Zdzislaw Musielak, Billy Quarles, Three Body Dynamics and Its Applications To Exoplanets, in: SpringerBriefs in Astronomy, 2017.
[6] L. Euler, Nov. Comm. Acad. Imp. Petropolitanae 10 (1740) 207–242; L. Euler, Nov. Comm. Acad. Imp. Petropolitanae 11 (1740) 152–184; Mémoires de l'Acad. de Berlin 11 228–249.
[7] J.L. Lagrange, Oeuvres tome 6; Miscellanea Taurinensia 4 (1772) 118–243; Oeuvres 2 67–121; Mécanique Analytique 262–286; second ed., vol. 2, pp. 108–121; Oeuvres 12 101–114.
[8] A. Tokovinin, in: N. St-Louis, A.F.J. Moffat (Eds.), Massive Stars in Interacting Binaries, in: ASP Conference Series, vol. 367, 2007, pp. 615–619.
[9] J. Schubart, Astron. Nachr. 283 (1956) 17.
[10] R. Vernić, Hrvatsko Prirod, Društvo Clas. Mat. Fiz. Astron. Ser. 2 8 (1953) 247–266.
[11] R.F. Arenstorf, in: J.K. Hale, J.P. Lasalle (Eds.), Differential Equations and Dynamical Problems, Academic Press, 1967, pp. 55–68.
[12] W.H. Jefferys, J. Moser, Astron. J. 71 (1966) 568–578.
[13] V. Szebehely, C.F. Peters, Astron. J. 72 (1967) 876–883.
[14] V. Szebehely, C.F. Peters, Astron. J. 72 (1967) 1187–1190.
[15] E.M. Standish, in: G.E.O. Giacaglia (Ed.), Periodic Orbits, Stability and Resonances, D. Reidel Publ. Co., Dordrecht, Holland, 1970, pp. 375–381.
[16] M. Hénon, Celest. Mech. 10 (1974) 375.
[17] R. Broucke, D. Boggs, Celest. Mech. 11 (1975) 13.
[18] R. Broucke, Celest. Mech. 12 (1975) 439.
[19] J.D. Hadjidemetriou, Celest. Mech. 12 (1975) 155.
[20] J.D. Hadjidemetriou, Th. Christides, Celest. Mech. 12 (1975) 175.
[21] J.D. Hadjidemetriou, Celest. Mech. 12 (1975) 255.

[22] M. Hénon, Celest. Mech. 13 (1976) 267.
[23] M. Hénon, Celest. Mech. 15 (1977) 243.
[24] M. Hénon, Celest. Mech. 15 (1977) 99.
[25] L.D. Landau, E.M. Lifshitz, Mechanics, third ed., Butterworth-Heinemann, Oxford, 1976.
[26] Kuo-Chang Chen, Ann. of Math. 167 (2008) 325–348.
[27] Kuo-Chang Chen, Yu-Chu Lin, Commun. Math. Phys. 291 (2009) 403–441.
[28] C. Moore, Phys. Rev. Lett. 70 (1993) 3675.
[29] A. Chenciner, R. Montgomery, Ann. of Math. 152 (2000) 881–901.
[30] C. Simó, in: Alain Chenciner, Richard Cushman, Clark Robinson, Zhihong Jeff Xia (Eds.), Celestial Mechanics, in: Contemporary Mathematics, vol. 292, AMS, Providence, R.I, 2002, p. 209.
[31] A.I. Martynova, V.V. Orlov, A.V. Rubinov, Astron. Rep. 53 (2009) 710.
[32] M. Šuvakov, V. Dmitrašinović, Phys. Rev. Lett. 110 (2013) 114301.
[33] Milovan Šuvakov, Celestial Mech. Dynam. Astronom. 119 (2014) 369–377.
[34] M. Šuvakov, V. Dmitrašinović, Amer. J. Phys. 82 (2014) 609–619.
[35] P.P. Iasko, V.V. Orlov, Astron. Rep. 58 (11) (2014) 869–879.
[36] V. Dmitrašinović, M. Šuvakov, Phys. Lett. A 379 (2015) 1939–1945.
[37] V. Dmitrašinović, Ana Hudomal, Mitsuru Shibayama, Ayumu Sugita, J. Phys. A 51 (2018) 315101, physics.
[38] http://three-body.ipb.ac.rs/ and http://three-body.ipb.ac.rs/sequences.php.
[39] M.R. Janković, V. Dmitrašinović, Phys. Rev. Lett. 116 (5) (2016) 064301.
[40] Danya Rose, Geometric Phase and Periodic Orbits of the Equal-Mass, Planar Three-Body Problem with Vanishing Angular Momentum (Ph.D. thesis), University of Sydney, 2016, Available at https://ses.library.usyd.edu.au/h{and}le/2123/14416.
[41] Milovan Šuvakov, Mitsuru Shibayama, Celestial Mech. Dynam. Astronom. 124 (2016) 155–162.
[42] X. Li, Y. Jing, S. Liao, Publ. Astron. Soc. Japan 057 (2018) 1–7.
[43] Xiaoming Li, Shijun Liao, Sci. China A 60 (12) (2017) 129511.
[44] E. Davoust, R. Broucke, Astron. Astrophys. 112 (1982) 305–320.
[45] M. Nauenberg, Phys. Lett. A 292 (2001) 93–99.
[46] M. Nauenberg, Celest. Mech. 97 (2007) 1.
[47] A. Chenciner, J. Fejoz, R. Montgomery, Nonlinearity 18 (2005) 1407–1424.
[48] R. Broucke, A. Elipe, A. Riaguas, Chaos Solitons Fractals 30 (2006) 513–520.
[49] G.D. Birkhoff, D.C. Lewis, Ann. Mat. 12 (4) (1933) 117–133.
[50] M. Šindik, A. Sugita, M. Šuvakov, V. Dmitrašinović, Phys. Rev. E 98 (2018) 060101.
[51] R. Montgomery, Nonlinearity 11 (1998) 363–376.
[52] R. Montgomery, Ergodic Theory Dynam. Systems 27 (6) (2007) 1933–1946.
[53] A.E. Roy, M.W. Ovenden, Mon. Not R. Astron. Soc. 115 (1955) 296–309.
[54] A. Vanderbauwhede, Discrete Contin. Dyn. Syst. Ser. A 33 (2013) 359.
[55] Abimael Bengochea, Jorge Galán, Ernesto Pérez-Chavela, Astrophys. Space Sci. 348 (2013) 399–415.
[56] Martin Eric Grant, of Las Positas College 3000 Campus Hill Drive Livermore CA 94551, private communication, Jan 8, 2018, <martin@ramblemuse.com>. See also https://math.stackexchange.com/questions/1851339/possibly-new-solution-to-equal-mass-three-body-problem-refinement-required.
[57] A.N. Kolmogorov, Dokl. Akad. Nauk. SSSR 98 (1954) 527–530 (in Russian).
[58] V.I. Arnold, A.N. Proof of, Russian Math. Surveys 18 (1963) 9–36.
[59] J. Moser, Nachr. Akad. Wiss. Göttingen Math. Phys. Kl. IIa 1 (1962) 1–20.
[60] P.P. Iasko, V.V. Orlov, Astron. Rep. 59 (5) (2015) 404–413.
[61] T. Levi-Civita, Acta Math. 30 (1906) 305–327, Opere Matematiche, vol. 2, Bologna, 1956, p. 419.
[62] T. Levi-Civita, Acta Math. 42 (1920) 99–144.
[63] P. Kustaanheimo, E.L. Stiefel, Reine Angew. Math. 218 (1965) 204.
[64] James Montaldi, Katrina Steckles, Classification of symmetry groups for planar n-body choreographie, arXiv:1305.0470, see also: http://www.maths.manchester.ac.uk/jm/Choreographies/.
[65] Abimael Bengochea, Jorge Galán, Manuel Falconi, Astrophys. Space Sci. 333 (2011) 40–408.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **484,** 2296–2308 (2019)
Advance Access publication 2019 January 5

doi:10.1093/mnras/stz004

Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc

Marija R. Jankovic ,[⋆] James E. Owen and Subhanjoy Mohanty
Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK

Accepted 2018 December 12. Received 2018 December 11; in original form 2018 July 11

ABSTRACT

We explore *in situ* formation and subsequent evolution of close-in super-Earths and mini-Neptunes. We adopt a steady-state inner protoplanetary gas disc structure that arises from viscous accretion due to the magneto-rotational instability (MRI). We consider the evolution of dust in the inner disc, including growth, radial drift, and fragmentation, and find that dust particles that radially drift into the inner disc fragment severely due to the MRI-induced turbulence. This result has two consequences: (1) radial drift of grains within the inner disc is quenched, leading to an enhancement of dust in the inner regions that scales as dust-to-gas-mass-flux-ratio at ∼1 au; (2) however, despite this enhancement, planetesimal formation is impeded by the small grain size. Nevertheless, assuming that planetary cores are present in the inner disc, we then investigate the accretion of atmospheres on to cores and their subsequent photoevaporation. We then compare our results to the observed exoplanet mass–radius relationship. We find that (1) the low gas surface densities and high temperatures in the inner disc reduce gas accretion on to cores compared to the minimum mass solar nebula, preventing the cores from growing into hot Jupiters, in agreement with the data; (2) however, our predicted envelope masses are still typically larger than observed ones. Finally, we sketch a qualitative picture of how grains may grow and planetesimals form in the inner disc if grain effects on the ionization levels and the MRI and the back reaction of the dust on the gas (both neglected in our calculations) are accounted for.

Key words: planets and satellites: formation – protoplanetary discs.

1 INTRODUCTION

Recent advances in exoplanet detection, led primarily by the *Kepler* mission, have uncovered several new classes of exoplanets (e.g. Borucki et al. 2011; Batalha et al. 2013). These are the close-in super-Earths and mini-Neptunes, planets with radii of 1–4 $R_\oplus$ and periods of up to 100 d, and they are found to be abundant around solar and sub-solar mass stars (e.g. Fressin et al. 2013; Dressing & Charbonneau 2015). How (and where in their parent protoplanetary discs) these planets form is a subject undergoing intense study.

One suggestion is that these super-Earths/mini-Neptunes form at larger separations, as more solids are potentially available outside the ice line, and then migrate inwards through the disc (e.g. Ida & Lin 2008; Kley & Nelson 2012; Cossou, Raymond & Pierens 2013; Cossou et al. 2014). However, the migration scenario predicts that planets in multiplanet systems should typically end up in mean motion resonances, whereas such orbital resonances are rare among the *Kepler* planets (e.g. Baruteau et al. 2014; Winn & Fabrycky 2015). Although several mechanisms have been explored to either break the resonances or prevent capture into them, (e.g. Goldreich & Schlichting 2014; Izidoro et al. 2017; Liu, Ormel & Lin 2017), this discrepancy has not yet been fully resolved. Moreover, the radius distribution of the *Kepler* planets, shaped by atmospheric photoevaporation, appears consistent with the planetary cores having a rock/iron (Earth-like) composition (Owen & Wu 2017), implying formation inside the ice line and arguing against significant migration.

An alternative scenario is that these planets form *in situ*, close to their present orbits. In this case, planetary cores form in the inner protoplanetary disc. They can still be subjected to planet–disc interactions, and the two are not mutually exclusive. If the cores are subjected to the fast type I migration, the innermost planet could stall at the inner disc edge (Masset et al. 2006; Ogihara, Morbidelli & Guillot 2015), explaining why the observed period distribution of the innermost planet peaks around ∼ 10 d (Mulders et al. 2018; see also Mulders, Pascucci & Apai 2015, Lee & Chiang 2017). Alternatively, type I migration could be suppressed if the surface density profile is flat or has a positive slope in the inner disc (Ogihara et al. 2018), or stalled by the core opening a gap (e.g. Hu et al. 2016). In fact, there is evidence suggesting that the

---

[⋆] E-mail: m.jankovic16@imperial.ac.uk

© 2019 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Kepler planets could have been massive enough to open gaps in the inner disc (Wu 2018).

If the close-in planets do form *in situ*, a large amount of solids is necessary in the inner protoplanetary disc compared to the amount in the minimum mass solar nebula (Chiang & Laughlin 2013). These solids may be delivered to the inner disc from the outer disc prior to planet formation (Hansen & Murray 2012, 2013; Chatterjee & Tan 2014), through the radial drift of pebbles and rocks (Weidenschilling 1977; Takeuchi & Lin 2002; Armitage 2018). The growth of dust grains in the outer disc and their subsequent radial drift inwards have been confirmed by observations (e.g. Panić et al. 2009; Andrews et al. 2012; Isella, Pérez & Carpenter 2012; Rosenfeld et al. 2013; Powell, Murray-Clay & Schlichting 2017).

To create a dust-rich inner disc in which to form planets, the radial drift of dust particles needs to be stopped or slowed down. The radial drift of particles in the Epstein drag regime slows down closer to the star in conventional disc models, and this can concentrate dust in the inner disc to some extent (Youdin & Shu 2002; Youdin & Chiang 2004; Birnstiel, Dullemond & Brauer 2010; Birnstiel, Klahr & Ercolano 2012; Drążkowska, Alibert & Moore 2016). Another way to halt the radial drift and enrich the inner disc with dust is to trap the solids inside an axisymmetric local gas pressure maximum that is expected to form if the accretion in the inner disc is driven by the magneto-rotational instability (MRI; Kretke et al. 2009; Dzyurkevich et al. 2010; Drążkowska, Windmark & Dullemond 2013; Chatterjee & Tan 2014). A gas pressure maximum acts as a trap for the marginally coupled solids as the gas inwards of the pressure maximum is super-Keplerian, reversing the direction of the radial drift (e.g. Pinilla, Benisty & Birnstiel 2012). Furthermore, in a steady-state disc accreting due to the MRI a pressure maximum forms at the boundary between the thermally ionized innermost disc in which the MRI-induced viscosity is high, and the low-viscosity *dead zone* in which the MRI is suppressed due to low ionization levels (Gammie 1996). This local pressure maximum is expected to form at few tenths of au around solar and sub-solar stars (Chatterjee & Tan 2014), which is consistent with the orbital distances of the close-in super-Earths and mini-Neptunes.

Mohanty et al. (2018) presented a semi-analytic steady-state inner disc model in which the disc structure, thermal ionization, and the viscosity due to the MRI were determined self-consistently. The location of the pressure maximum inferred from this model is similarly in general agreement with the orbital distances of the close-in planets. An important insight from the Mohanty et al. (2018) models of the inner disc is that in a steady state they predict gas surface densities that are considerably lower than those of the minimum mass solar nebula. This is not surprising as the minimum mass solar nebula simply extrapolated the surface density to small separations, whereas in reality the shrinking size of the dead-zone results in more efficient angular momentum transport, hence lowering the surface densities towards smaller separations.

The atmospheres of many of the close-in super-Earths and mini-Neptunes must be H/He dominated (e.g. Jontof-Hutter et al. 2016), and they typically make up 0.1–10 per cent of their total mass (Lopez & Fortney 2014; Wolfgang & Lopez 2015). Thus, they are considerably larger than the atmospheres of the planets in the inner Solar system. Outgassing of hydrogen from a rocky core is not sufficient to explain the majority of these atmospheres (Rogers et al. 2011). Thus, these atmospheres are most likely composed of gas accreted from the protoplanetary disc after the formation of a solid core. If so, these atmospheres are formed steadily through core accretion.

Lee, Chiang & Ormel (2014; see also Lee & Chiang 2015, 2016; Lee, Chiang & Ferguson 2017) argue that core accretion is so efficient that the key concern is how to stop the super-Earth cores from undergoing run-away accretion and becoming gas giants (Mizuno 1980). This led Lee & Chiang (2016) to suggest that super-Earth/mini-Neptune formation occurred in gas-poor 'transition discs', during the final short-lived phase of disc dispersal. The requirement for a gas-poor inner disc raises the question if planet formation in the gas-poor inner disc arising due to steady-state MRI accretion could be a desirable scenario and a possible alternative to the Lee & Chiang (2016) proposal of atmospheric accretion during disc dispersal.

In this paper, we examine the possibility of the *in situ* formation of the close-in planets in the inner disc structure arising from MRI-driven accretion, obtained using the self-consistent model of Mohanty et al. (2018). First, in Section 2, we examine the evolution of dust in the inner disc and discuss the possibility of planetesimal formation. In Section 3, we calculate the atmospheres that super-Earth and mini-Neptune cores can accrete in the gas-poor inner disc implied by the MRI, and then evolve them forward in time, in order to compare our calculations to the data.

An important caveat to our dust calculations is that we evolve the dust assuming the gas profile is fixed in time. Dust grains can act to suppress the MRI by lowering the coupling between the gas and the magnetic field (e.g. Sano et al. 2000; Ilgner & Nelson 2006), and could thus significantly alter the gas disc structure. Moreover, at high dust-to-gas ratios, which we show can be achieved in the inner disc, dust becomes dynamically important, and the dynamical back reaction on the gas should be taken into account. In Section 4, we qualitatively discuss these effects, and how they might influence our results. We shall address the self-consistent feedback of dust enhancement on the gas disc structure in subsequent studies.

2 DUST EVOLUTION

We consider the evolution of the dust, including growth, fragmentation, and radial drift, in a steady-state gas disc that is viscously accreting due to the MRI. The structure of the gas disc (gas surface density, temperature, pressure, and viscosity) is obtained from the steady-state models of the inner protoplanetary gas disc calculated by Mohanty et al. (2018). In these models, the viscosity (α parameter) is determined self-consistently with the disc structure (Shakura & Sunyaev 1973), thermal ionization, and MRI criteria (Bai 2011; Bai & Stone 2011). The parameters of the model are the stellar mass and radius (M_*, R_*), steady-state accretion rate ($\dot{M}_{\rm g}$), and minimum viscosity of the gas due to purely hydrodynamical effects ($\alpha_{\rm DZ}$). The minimum viscosity $\alpha_{\rm DZ}$ is the assumed value of α inside the MRI-dead zones; it is a minimum in the sense that such hydrodynamical effects are assumed to dominate over the MRI-induced turbulence if the MRI implies a viscosity lower than $\alpha_{\rm DZ}$. In this work, we primarily use $M_* = 1\,{\rm M}_\odot$, $R_* = 2.33\,{\rm R}_\odot$, $\dot{M}_{\rm g} = 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$, and $\alpha_{\rm DZ} = 10^{-4}$. The disc structure for these parameters is shown in Fig. 1. The local gas pressure maximum is at an orbital distance of ~0.7 au, and temperature and surface density at that location are ~1000 K and ~5000 g cm^{-2}, respectively. Outwards from the pressure (and the surface density) maximum, the MRI is suppressed and $\alpha = \alpha_{\rm DZ}$.

Throughout this paper, we use the above fiducial values for the disc parameters. Here, we briefly describe how the disc structure depends on these parameters. For a higher gas accretion rate $\dot{M}_{\rm g}$, the radial α profile is roughly an outward-translated version of the one shown in the top panel of Fig. 1, and inward-translated

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

for a smaller $\dot{M}_{\rm g}$. The radial location of the local gas pressure maximum scales with the accretion rate approximately as $\dot{M}_{\rm g}^{1/2}$. For a higher $\alpha_{\rm DZ}$, the α falls to this value closer to the star, and vice versa, but the value of α as a function of orbital distance remains almost the same otherwise; as a result, the radial location of the local pressure maximum scales with the minimum dead-zone viscosity as $\alpha_{\rm DZ}^{-1/4}$. Furthermore, inwards of the pressure maximum the temperature has to be sufficiently high for thermal ionization of potassium to support the MRI, and so it is always larger than 1000 K, regardless of the exact choice of $\dot{M}_{\rm g}$ and $\alpha_{\rm DZ}$. For a steady-state vertically-isothermal α-disc, the accretion rate is $\dot{M}_{\rm g} = 3\pi c_{\rm s}^2 \alpha \Sigma_{\rm g}/\Omega$ (ignoring an additional factor that depends on the boundary condition at the inner disc edge, and that becomes unimportant far away from the edge). Further assuming (as is approximately the case) that the temperature at the location of the pressure maximum is constant regardless of the disc parameters, it follows that the maximum gas surface density approximately depends on the disc parameters[1] as $\dot{M}_{\rm g}^{1/4}$ and $\alpha_{\rm DZ}^{-5/8}$.

As we neglect the dynamical back reaction of the dust on the gas and the effect of the dust on the MRI, the local evolutionary time-scales are considerably shorter than the Myr time-scale on which the accretion rate will evolve. Thus, the structure of the gas disc is held fixed, i.e. not evolved in time. Note also that our gas disc model considers MRI-driven accretion only, and magnetic winds, if present, can also affect the inner gas disc structure (e.g. reduce the gas surface density compared to the minimum mass solar nebula even in the absence of the MRI, Suzuki et al. 2016).

2.1 Methods

The dust particle size distribution is evolved using the two-population model of Birnstiel et al. (2012). The dust surface density $\Sigma_{\rm d}$ is evolved using the advection–diffusion equation

$$\frac{\partial \Sigma_{\rm d}}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(\Sigma_{\rm d}\bar{u} - D_{\rm gas}\Sigma_{\rm g}\frac{\partial}{\partial r}\left(\frac{\Sigma_{\rm d}}{\Sigma_{\rm g}}\right)\right)\right] = 0\,, \tag{1}$$

where r is the cylindrical radius, $\bar{u}$ is the dust advection velocity, $D_{\rm gas}$ is the gas diffusivity, and $\Sigma_{\rm g}$ is the gas surface density.

The dust advection velocity is a sum of the velocities due to advection with the accreting gas and radial drift. For particles with Stokes number $\mathrm{St}_i = \pi\rho_{\rm s}a_i/2\Sigma_{\rm g}$ (with $\rho_{\rm s}$ the internal density of the dust and a_i the particle size), and adopting the terminal velocity approximation (e.g. Takeuchi & Lin 2002), the dust velocity is given by

$$u_i = \frac{1}{1+\mathrm{St}_i^2}u_{\rm gas} + \frac{2}{\mathrm{St}_i + \mathrm{St}_i^{-1}}u_{\rm drift}\,, \tag{2}$$

where

$$u_{\rm drift} = \frac{c_{\rm s}^2}{2v_{\rm K}}\frac{{\rm d}\ \ln P}{{\rm d}\ \ln r}\,, \tag{3}$$

with $c_{\rm s}$ the speed of sound, $v_{\rm K}$ the Keplerian velocity, and P the mid-plane gas pressure. Small particles ($\mathrm{St} \ll 1$) move with the gas, and larger particles can move faster or slower than the gas, depending on the sign of the pressure gradient.

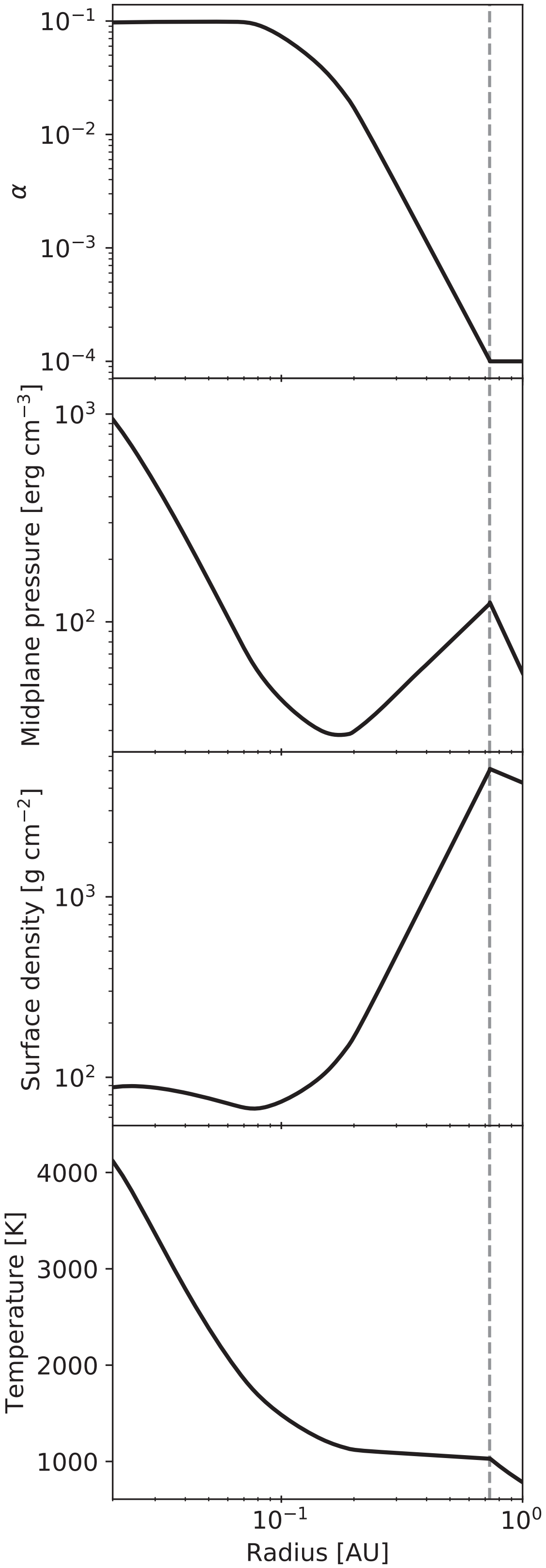


Figure 1 Gas disc structure from the steady-state model of Mohanty et al. (2018) for $M_* = 1\,{\rm M}_\odot$, $R_* = 2.34\,{\rm R}_\odot$, $\dot{M}_{\rm g} = 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$, and $\alpha_{\rm DZ} = 10^{-4}$. From top to bottom: α parameter, mid-plane pressure, surface density, and temperature, as functions of radius. Location of the local gas pressure maximum due to the MRI is indicated by the vertical dashed line.

---

[1] Derivation of the scalings from the Shakura–Sunyaev equations takes into account the small correction due to the dependence of the temperature on the disc parameters and yields that the maximum surface density depends on the disc parameters as $\dot{M}_{\rm g}^{3/10}$ and $\alpha_{\rm DZ}^{-13/20}$. These small corrections are omitted here for simplicity.

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

In the Birnstiel et al. (2012) model, the dust surface density $\Sigma_{\rm d}$ and dust advection velocity $\bar{u}$ are the sum and the mass-weighted average, respectively, of the surface density and velocity of two populations of particles: small monomer-sized particles ($a_0 = 1\ \mu$m) and large particles (a_1). The size of the large particles evolves in time and space. At first, small dust grains grow. Then, at each radius the size of the large particles is set by whichever process yields the smallest grain size: radial drift ($a_{\rm drift}$), where grains larger than $a_{\rm drift}$ radially drift more quickly than they can grow; drift-fragmentation ($a_{\rm df}$), where grains larger than $a_{\rm df}$ fragment due to relative radial drift velocities; or turbulent fragmentation ($a_{\rm frag}$), where grains larger than $a_{\rm frag}$ fragment due to relative velocities induced by turbulence.

As we are interested in the innermost protoplanetary disc, inside the ice line, we set the bulk density of particles to $\rho_{\rm s} = 3\,{\rm g\,cm^{-3}}$, and the critical fragmentation velocity to $u_{\rm f} = 1\,{\rm m\,s^{-1}}$, based on experiments on collisions of silicate grains (Blum & Münch 1993; Beitz et al. 2011; Schräpler et al. 2012; Bukhari Syed et al. 2017) of similar size (the regime applicable in the Birnstiel et al. 2012 model used here).[2]

The viscosity parameter α due to the MRI turbulence from our gas disc model is determined as a vertical average at each radius. We assume that this vertically averaged α signifies the strength of turbulence that particles feel, which in turn determines the particle size due to turbulent fragmentation ($a_{\rm frag}$) and the radial turbulent mixing (diffusivity $D_{\rm gas}$). However, the viscosity (and the level of turbulence) can be different at the disc mid-plane compared to the upper layers of the disc, depending on where the non-ideal magnetohydrodynamic effects suppress the MRI. Dust tends to settle towards mid-plane, its scale height being determined by the balance between gravitational settling and turbulent stirring (e.g. Youdin & Lithwick 2007). The use of a vertically averaged α could thus be invalid in weakly turbulent regions. Nevertheless, we proceed with this assumption for ease of computation; we do check the robustness of our results by swapping the vertically averaged α parameter for the mid-plane value in one run, and recover qualitatively the same results.

2.2 Numerical procedure

The advection–diffusion equation (1) is integrated using an explicit first order in time and second order in space finite element method. The advection term is integrated with an upwind scheme that adopts a van Leer flux limiter; the numerical scheme is described in detail by Owen (2014). The only modification we have made is the inclusion of the Birnstiel et al. (2012) dust evolution algorithm. Our simulations use 262 cells in the radial direction with a 0.002 au spacing inwards of 0.4 au and a 0.01 au spacing outwards. The inner boundary is set to 0.016 au, and the outer boundary is set to 1 au (i.e. outside the pressure maximum, but inside the ice line). The time-step is set with respect to the spatial resolution, the advection speed, and the diffusion coefficient, so that it obeys the Courant–Friedrichs–Lewy condition. Following Birnstiel et al. (2012), at each time-step the size of the large particles is updated to the smallest of the four size limits ($a_{\rm growth}$, $a_{\rm drift}$, $a_{\rm df}$, or $a_{\rm frag}$). At the outer boundary, we impose a constant dust accretion rate, which we vary in different runs. This is to mimic the fact that the dust flux is not a fixed quantity and can vary with time due to radial drift of dust from the outer disc (e.g. Birnstiel et al. 2012). Dust particle size at the outer boundary is calculated self-consistently.

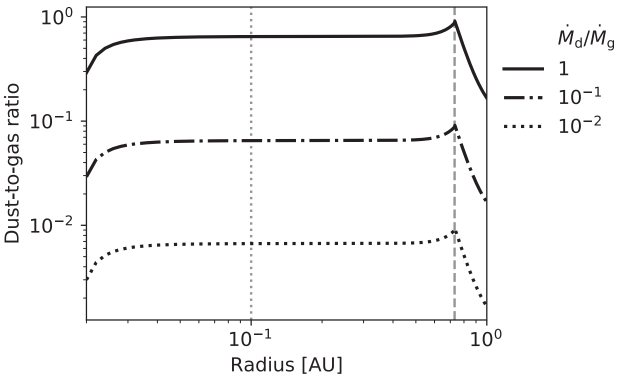


Figure 2 Dust-to-gas ratio $\Sigma_{\rm d}/\Sigma_{\rm g}$ as a function of radius after 0.2 Myr, for various dust accretion rates at the outer boundary, as indicated in plot legend. Location of the local gas pressure maximum due to the MRI is indicated by the vertical dashed line, and location of the dust sublimation line by the vertical dotted line.

Furthermore, we neglect the effects of dust sublimation, and note that the temperature in our gas disc model exceeds the dust sublimation value (∼1500 K) only inwards of 0.1 au.

2.3 Results

Our simulations proceed as follows: initially, the dust-to-gas ratio is 0.01 at all radii and all dust grains are monomers ($a = 1\ \mu$m). We evolve the dust for 0.2 Myr, by which time it has reached steady state. The steady-state dust-to-gas ratio is shown in Fig. 2 as a function of radius for three different values of dust accretion rates $\dot{M}_{\rm d}$ at the outer boundary condition. For any given dust accretion rate, the steady-state dust-to-gas ratio is roughly constant inwards of the pressure maximum (indicated by the vertical dashed line), and it decreases outwards from the pressure maximum. There is only a moderate accumulation of dust at the pressure maximum, compared to the rest of the inner disc. This implies that the pressure maximum does not efficiently trap dust particles.

Essentially, the particles do not feel significant gas drag, and thus do not significantly feel the effect of the change in the sign of the pressure gradient inwards of the pressure maximum. This happens because the MRI-induced turbulence causes fragmentation, resulting in small particles. Fig. 3 shows the three dust size limits (due to radial drift, drift-fragmentation, and turbulent fragmentation) as functions of radius, calculated in steady state. The smallest of the three ($a_{\rm frag}$), due to turbulent fragmentation, sets the size of the population of large particles in these simulations, which dictates the evolution of dust overall. Particle size is thus limited to only

---

[2] We note that simulations of grain collisions (Meru et al. 2013) indicate that the critical fragmentation velocity could be significantly higher for porous grains than for compact ones, for a range of porosities that is not robustly covered by the above experiments. Now, in the fragmentation-limited regime particle size depends quadratically on $u_{\rm f}$. Thus, *if* porosity is important, and hence $u_{\rm f}$ is set to, e.g. $10\,{\rm m\,s^{-1}}$ instead, particle sizes (Stokes number) would be larger by a factor of 100, strongly affecting how coupled a particle is to the gas flow and how susceptible to radial drift. However, since particles also become less porous (to the point of becoming compact, and fragile) in a wide variety of conditions – e.g. in collisions that result in coagulation (simulations by Meru et al. 2013, experiments by Kothe, Güttler & Blum 2010), collisions that result in bouncing (Weidling et al. 2009), and collisions of larger grains with monomers (Schräpler & Blum 2011) – we opt to use the compact grain value of $u_{\rm f} = 1\,{\rm m\,s^{-1}}$.

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

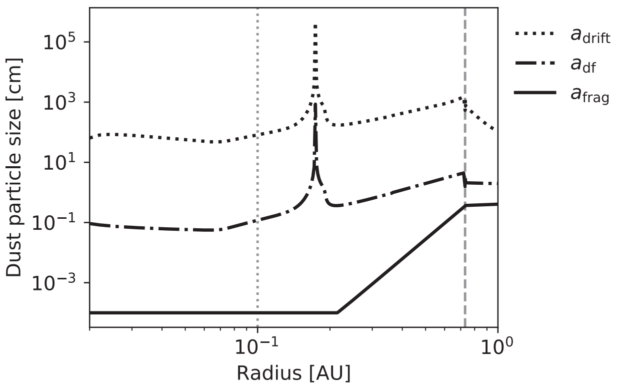


Figure 3 Dust particle size limits due to radial drift ($a_{\rm drift}$), drift-fragmentation ($a_{\rm df}$), and turbulent fragmentation ($a_{\rm frag}$) as functions of radius after 0.2 Myr. Location of the local gas pressure maximum due to the MRI is indicated by the vertical dashed line, and location of the dust sublimation line by the vertical dotted line. Spikes in $a_{\rm drift}$ and $a_{\rm df}$ correspond to gas pressure extrema.

a few millimetres near the pressure maximum (indicated by the vertical dashed line), and the particles are monomer sized in the innermost disc.

The particle size determines, through the Stokes number, how coupled the dust is to the gas. Thus, the particle size determines to what extent the particles move with the accreting gas towards the star and also by how much they are slowed down or sped up by the gas drag. In this case, inwards of the pressure maximum (and outwards from the pressure minimum at ~ 0.2 au) the gas drag acts outwards ($u_{\rm drift} > 0$). However, as dust particles are small inside the pressure maximum ($\mathrm{St} \sim 4 \times 10^{-4}$ even for the large particles) and their size further decreases inwards, the dust advection velocity is outwards only in a very narrow region. Consequently, after accounting for the diffusivity (i.e. the radial turbulent mixing of dust, which limits the radial gradient of the dust-to-gas ratio), the mass build-up inside the pressure maximum is moderate compared to the rest of the inner disc.

The dust advection velocity used in our method is a mass-weighted average of the velocity of the monomer-sized particles and the large particles (of size $a_{\rm frag}$, Fig. 3). The monomer-sized particles are just advected by the gas through the pressure maximum, but it can be shown that an individual large particle will also not be trapped. This is due to dust particles being in the fragmentation limit, in which the particles are fragmented faster than they drift. The drift time-scale for the large particles inside the pressure trap (the region inwards of the pressure maximum where their advection velocity is outwards) can be estimated by

$$t_{\rm drift} \approx \frac{1}{2} \frac{d_{\rm trap}}{u_1(r_{P_{\rm max}})}, \tag{4}$$

where $d_{\rm trap} \approx 0.06$ au is the radial width of the trap, and the particle velocity is $u_1 \approx 2\,\mathrm{cm\,s^{-1}}$ (see equation 2), which is evaluated just inwards of the pressure maximum. The velocity u_1 decreases inwards, and so this estimate, $t_{\rm drift} \approx 7400$ yr, is a lower limit. The collisional (i.e. fragmentation) time-scale for the large particles, $t_{\rm col} = (n\sigma\Delta v)^{-1}$, is much shorter in comparison. Here, $n = f_m \rho_{\rm d}/m_{\rm d}$ is the number density of large particles, f_m is the mass fraction of the large particles ($f_m = 0.75$ in the fragmentation limit, Birnstiel et al. 2012), $\rho_{\rm d} \approx \Sigma_{\rm d}\Omega/(\sqrt{2\pi}c_{\rm s})$ is the mid-plane mass density of particles, $m_{\rm d}$ is mass of a single particle, σ is the collisional cross-section, and $\Delta v \approx \sqrt{3\alpha \mathrm{St}}c_{\rm s}$ is the typical relative velocity between the particles due to turbulence (Ormel & Cuzzi 2007). This yields the collisional time-scale of

$$\begin{aligned} t_{\rm col} &= \sqrt{\frac{8}{27\pi}} \frac{\Sigma_{\rm g}}{f_m \Sigma_{\rm d}} \sqrt{\frac{\mathrm{St}}{\alpha}} \frac{1}{\Omega}, \\ &\approx 4 \left(\frac{f_m}{0.75}\right)^{-1} \left(\frac{\Sigma_{\rm d}}{0.01\Sigma_{\rm g}}\right)^{-1} \left(\frac{\mathrm{St}}{10^{-4}}\right)^{1/2} \left(\frac{\alpha}{10^{-4}}\right)^{-1/2} \\ &\quad \times \left(\frac{\Omega}{10\ \mathrm{yr}^{-1}}\right)^{-1} \mathrm{yr}, \end{aligned} \tag{5}$$

where we have expressed the particle size $a_{\rm frag}$ in terms of the Stokes number St. At the pressure maximum $\alpha = 10^{-4}$, and so $t_{\rm col} \approx 0.08\Sigma_{\rm g}/\Sigma_{\rm d}$ yr $\ll t_{\rm drift}$ for dust-to-gas ratios $\Sigma_{\rm d}/\Sigma_{\rm g} \gtrsim 0.01$. Thus, instead of getting trapped in the pressure maximum, dust particles fragment and flow inwards.

In the innermost disc, turbulent fragmentation yields monomer-sized particles that are entrained with the gas ($\bar{u} \sim u_{\rm gas}$). And so the radial drift inwards is slowed by particles becoming well coupled to the gas. This is also why, in steady state, the dust-to-gas ratio is roughly constant inwards of the pressure maximum.

Finally, despite the pressure maximum not trapping the inflowing particles, the dust-to-gas ratio is enhanced. Because the dust moves with the gas in the innermost disc, the steady-state dust-to-gas ratio is directly proportional to the ratio of dust-to-gas accretion rates at the outer boundary, $\dot{M}_{\rm d}/\dot{M}_{\rm g}$. Now, the initial dust-to-gas ratio everywhere is $\Sigma_{\rm d}/\Sigma_{\rm g} = 10^{-2}$. Preserving this ratio overtime in the inner disc would thus require $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 10^{-2}$ at the outer boundary. However, the growth of dust grains in the outer disc, and the attendant inwards radial drift of grains there, means that $\dot{M}_{\rm d}/\dot{M}_{\rm g} > 10^{-2}$ at the outer boundary (i.e. dust accretes inwards preferentially compared to gas). In this case, the inner disc in steady state will also have $\Sigma_{\rm d}/\Sigma_{\rm g} > 10^{-2}$. In other words, as Fig. 2 shows, radial drift of grains from the outer disc leads to an enrichment of solids in the inner disc.

What level of the enrichment is attainable depends on the ratio of dust and gas accretion rates $\dot{M}_{\rm d}/\dot{M}_{\rm g}$, i.e. how quickly the grains drift from the outer disc relative to gas accretion. Since the grain growth in the outer disc is limited by radial drift rather than fragmentation (Birnstiel et al. 2012), high grain drift rates are possible. For example, assuming that $\Sigma_{\rm d}/\Sigma_{\rm g} = 10^{-2}$ in the outer disc, achieving $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 1$ requires the radial drift velocity of grains ($\approx 2\,\mathrm{St}\,u_{\rm drift}$) to reach $10^2 u_{\rm gas}$. For the standard α-disc model and $\alpha = 10^{-4}$ (extrapolation of the disc model shown in Fig. 1), this is satisfied if grains grow to $\mathrm{St} \sim 10^{-2}$. This corresponds to a particle size an order of magnitude below the radial drift limit ($a_{\rm drift}$) throughout the outer disc. Therefore, the grains easily grow large enough to achieve accretion rates of $\dot{M}_{\rm d}/\dot{M}_{\rm g} \gtrsim 1$. These grains in the outer disc will contain ices that will evaporate as grains drift across the ice lines, towards the inner disc. However, even in the outermost disc silicates account for a considerable portion of the total solid mass (e.g. adopting the abundances of oxygen and carbon in their main molecular carriers from Öberg, Murray-Clay & Bergin 2011, 23 per cent of the total oxygen and carbon mass in the outer disc is in silicates and other refractories). Hence, Fig. 2 features dust accretion rates up to $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 1$ (which, as argued above, are likely), in which case the dust-to-gas ratio in the inner disc also approaches unity.

2.4 Implications for planetesimal formation

The above results show that the MRI yields a dust-enhanced inner disc, although at the expense of the dust particle size. Further-

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

more, as there is no trap for the dust particles, the accumulation of dust is limited by the dust inflow rate from the outer disc, and does not increase indefinitely. In this section, we explore if further concentration of particles via the streaming instability (SI; Youdin & Goodman 2005) and subsequent gravitational collapse into planetesimals could be the next step towards forming the close-in super-Earths and mini-Neptunes.

The SI can greatly concentrate dust particles if the ratio of dust-to-gas bulk densities is $\rho_{\rm d}/\rho_{\rm g} \gtrsim 1$ (Youdin & Goodman 2005; Johansen & Youdin 2007). This is most likely to be attained in the disc mid-plane, as dust particles gravitationally settle. The settling is balanced by turbulent stirring. One source of turbulence is the MRI. To reach $\rho_{\rm d}/\rho_{\rm g} \geq 1$ in the mid-plane in the presence of such turbulence, the dust-to-gas surface density ratio $\Sigma_{\rm d}/\Sigma_{\rm g}$ needs to be greater than or equal to $Z_{\rm cr1} = \sqrt{\alpha/({\rm St}+\alpha)}$ (Carrera et al. 2017).

Even in discs that are weakly turbulent or completely laminar, as dust settles the dust–gas interaction leads to turbulence (self-stirring) that can prevent clumping by the SI. In this case, SI can only successfully concentrate dust particles if the dust-to-gas ratio $\Sigma_{\rm d}/\Sigma_{\rm g}$ is greater than a critical value $Z_{\rm cr2}$ that depends on the particle Stokes number (Johansen, Youdin & Mac Low 2009; Carrera, Johansen & Davies 2015). For St < 0.1 (relevant to our simulations), this critical value has been most recently revised by Yang, Johansen & Carrera (2017), who find

$$\log Z_{\rm cr2} = 0.1\log^2 {\rm St} + 0.2\log {\rm St} - 1.76\,. \qquad (6)$$

Small dust grains that are entrained with the gas do not participate in the SI. Hence, to compare our results against the SI criteria, we use the dust-to-gas ratio $f_m\Sigma_{\rm d}/\Sigma_{\rm g}$ of the large grain population only (where $f_m = 0.75$ is the mass fraction of large particles; Birnstiel et al. 2012).

In the top panel of Fig. 4, we compare this dust-to-gas ratio, for the outer boundary condition $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 1$, with the above two criteria for the onset of the SI. We find that in the inner disc, turbulence due to the MRI is generally more prohibitive to dust settling than the turbulence due to dust–gas interactions. Both conditions are fulfilled only near the pressure (and density) maximum.

Provided that the SI successfully concentrates dust particles in the disc mid-plane, the dust bulk density there may reach up to 100–1000 times the local gas density (Johansen & Youdin 2007). Gravitational collapse of such particle concentrations will occur if the dust density exceeds the local Roche density (below which the star can tidally disrupt the fragment), $\rho_{\rm Roche} = 9M_*/4\pi r^3$. Comparing $\rho_{\rm Roche}$ with the mid-plane gas densities in our steady-state MRI disc (Fig. 4, bottom panel), we find that the condition for gravitational collapse is only satisfied near the pressure (and density) maximum, or at larger separations. Importantly, the bottom panel of Fig. 4 shows that the possibility of gravitational collapse of solids in the inner disc is severely limited by the steep gradient of the Roche density.

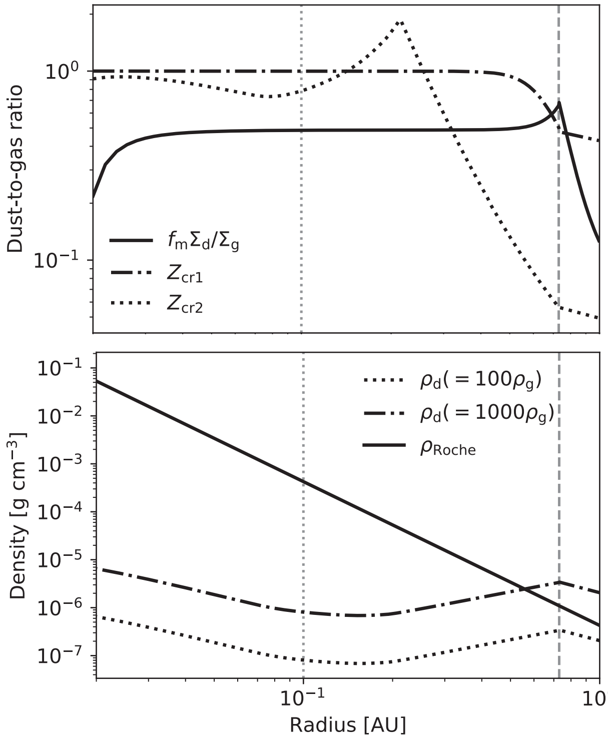


Figure 4 *Top panel:* Dust-to-gas ratio of large dust grains $f_{\rm m}\Sigma_{\rm d}/\Sigma_{\rm g}$ as a function of radius (when $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 1$ at the outer boundary; $f_m = 0.75$), compared to the two criteria for the onset of the streaming instabilities. *Bottom panel:* Expected peak local dust densities $\rho_{\rm d}$ if streaming instabilities successfully concentrate particles in the disc mid-plane as functions of radius, compared to the Roche density $\rho_{\rm Roche}$. In both panels, the vertical dashed line indicates location of the local pressure maximum, and the dotted line indicates the location of the dust sublimation line.

Overall, given the stellar and disc parameters used here, we find that the SI and the gravitational collapse pathway to planetesimals are viable in the inner disc only in a very narrow region near the pressure maximum. For gas accretion rates larger than the one used here ($\dot{M}_{\rm g} > 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$, which could be expected in the early phase of disc evolution, e.g. Manara et al. 2012), this conclusion will hold, while for sufficiently smaller accretion rates planetesimals would not form in this way anywhere near or inwards of the pressure maximum. First, in both cases dust evolution is expected to yield roughly the same steady-state dust-to-gas ratios given the same ratio of dust and gas accretion rates at the outer boundary. To confirm this conclusion, we repeat the dust evolution calculations for the gas accretion rate of $\dot{M}_{\rm g} = 10^{-9}\,{\rm M}_\odot\,{\rm yr}^{-1}$, obtaining results very similar to those above. In addition to similar steady-state dust-to-gas ratios, the particle Stokes number sharply drops inwards of the pressure maximum, and the SI is similarly triggered only around the pressure maximum. This is because the slope of the increase in α inwards of the pressure maximum is roughly the same for different $\dot{M}_{\rm g}$. Hence, the SI criteria is expected to be fulfilled only near the pressure maximum for higher gas accretion rates as well.

Secondly, even if SI is successfully triggered, to form planetesimals the peak dust density needs to be above the Roche density. The peak dust density scales with the mid-plane gas density at the pressure maximum, so it scales with the gas accretion rate approximately as $\dot{M}_{\rm g}^{-1/2}$, and with the radial location of the pressure (and density) maximum as $r_{P_{\rm max}}^{-1}$. Since the peak dust density is larger than the Roche density at the pressure maximum for $\dot{M}_{\rm g} = 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$, and the Roche density $\rho_{\rm Roche} \propto r^{-3}$, for $\dot{M}_{\rm g} > 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$ the gravitational collapse criterion will also be fulfilled near the pressure maximum, while for a sufficiently smaller accretion rate (including $\dot{M}_{\rm g} = 10^{-9}\,{\rm M}_\odot\,{\rm yr}^{-1}$) peak dust density will be too low.

Moreover, it is important to note that, in order for the SI to operate, there must be a relative azimuthal velocity between the dust and gas (e.g. Squire & Hopkins 2018), in addition to the density criteria invoked above. Such a relative velocity necessarily disappears at the pressure maximum itself, further constricting the region where the SI is viable in our disc.

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

However, we reiterate that our calculations do not include the effects of dust on the gas dynamics and on the MRI. The latter effect in particular may relax some of the constraints on the SI, and we discuss this possibility in Section 4. Overall, if the close-in sub-Neptunes form *in situ*, in an MRI-accreting inner disc (as suggested by Chatterjee & Tan 2014), any theory needs to explain how the dust grains can grow larger to either start the SI in a wider vicinity of the pressure maximum, or for the pressure maximum to become an efficient dust trap that can concentrate grains to densities needed for the gravitational collapse, or find an alternative pathway to planetesimal and core formation.

3 ACCRETION AND EVOLUTION OF PLANETARY ATMOSPHERES

We showed above that planetesimal formation through streaming instabilities and gravitational collapse is challenging despite the dust enhancement in the inner disc. Although these conclusions could change (see Section 4), it is not presently clear how exactly planetesimals or cores would arise in the inner disc, and we are unable to predict properties of solid cores formed *in situ*. Nevertheless, orbital distances, radii, and masses of many close-in planets have been well determined observationally. Furthermore, their radius distribution has been shown to be consistent with Earth-like composition and thus formation inside the ice line (Owen & Wu 2017), possibly in the inner disc. Thus, in this section we use observational results to perform a separate test of *in situ* formation by considering accretion of planetary atmospheres in the inner disc and their subsequent evolution.

To follow the accretion of planetary atmospheres, we assume that solid super-Earth-sized cores are present in the inner regions of our gas-poor inner disc structure taken from Mohanty et al. (2018). Furthermore, following the arguments of Lee & Chiang (2016), we ignore additional heating arising from further accretion of solids, as any amount of accretion providing significant heating typically results in the core rapidly reaching run-away masses.

After disc dispersal (which we do not model), we also account for atmospheric mass lost due to photoevaporation. This is an important addition, as several theoretical (e.g. Lopez & Fortney 2013; Owen & Wu 2013) and observational (e.g. Lundkvist et al. 2016; Fulton et al. 2017; Fulton & Petigura 2018) studies have shown that photoevaporation significantly sculpts the exoplanet population after formation.

3.1 Methods

3.1.1 *Accretion of planetary atmospheres*

We assume that solid super-Earth-sized planet cores accrete their gaseous envelopes in a gas-poor inner disc that is viscously accreting due to the MRI; this disc structure does not evolve in time. For a quasi-steady-state envelope, the accreted envelope mass fraction after time t can be estimated by the scaling relations from Lee & Chiang (2015; with an additional factor that accounts for varying gas surface density from Lee et al. 2017, Fung & Lee 2018):

$$X(t) = 0.07\left(\frac{t}{1\ \mathrm{Myr}}\right)^{0.4}\left(\frac{0.02}{Z}\right)^{0.4}\left(\frac{\mu}{2.37}\right)^{3.4} \times\left(\frac{M_{\mathrm{core}}}{5\mathrm{M}_{\oplus}}\right)^{1.7}\left(\frac{f_{\Sigma}}{0.1}\right)^{0.12} \tag{7}$$

for dusty atmospheres, and

$$X(t) = 0.18\left(\frac{t}{1\ \mathrm{Myr}}\right)^{0.4}\left(\frac{0.02}{Z}\right)^{0.4}\left(\frac{\mu}{2.37}\right)^{3.3} \times\left(\frac{M_{\mathrm{core}}}{5\mathrm{M}_{\oplus}}\right)^{1.6}\left(\frac{1600\ \ \mathrm{K}}{T_{\mathrm{rcb}}}\right)^{1.9}\left(\frac{f_{\Sigma}}{0.1}\right)^{0.12} \tag{8}$$

for dust-free atmospheres. Here, Z is the metallicity of the atmosphere; $\mu = 1/(0.5W + 0.25Y + 0.06Z)$ is the mean molecular weight, with $W = (1 - Z)/1.4$, $Y = 0.4(1 - Z)/1.4$; T_{rcb} is the temperature at the radiative–convective boundary inside the atmosphere; $f_{\Sigma} = \Sigma_{\mathrm{g}}/\Sigma_{\mathrm{MMSN}}$ is the ratio of the gas surface density (Σ_{g} from our inner disc model, Fig. 1) and the gas surface density profile of the minimum mass solar nebula ($\Sigma_{\mathrm{MMSN}} = 1700(d/1\ \ \mathrm{au})^{-3/2}\,\mathrm{g\,cm^{-2}}$, where d is the orbital radius, Hayashi 1981). Furthermore, we have assumed the gas adiabatic index is $\gamma = 1.2$, and that in dusty atmospheres $T_{\mathrm{rcb}} = 2500$ K, which arises from the disassociation of Hydrogen (see Lee & Chiang 2015, their section 2.1).

We use the expressions (7,8) to calculate how much gas a planet accretes in 1 Myr as a function of core mass, for various metallicities Z and gas surface density factors f_{Σ} in the case of dusty atmospheres (with $T_{\mathrm{rcb}} = 2500$ K), and various Z, f_{Σ}, and T_{rcb} in the case of dust-free atmospheres.

3.1.2 *Photoevaporation of planetary atmospheres*

These accreted atmospheres are then subjected to photoevaporation following disc dispersal. We use a simplified estimate of the photoevaporative mass-loss. First, for a given planet core mass M_{core} and (accreted) envelope mass fraction X we find the photospheric radius of the planet R_{p}. For this, we use a simple model of an atmosphere at hydrostatic equilibrium (Owen & Wu 2017), in which the solid core is surrounded by an adiabatic convective envelope, topped by an isothermal radiative photosphere. Next, the mass-loss time-scale due to high-energy stellar irradiation is (Owen & Wu 2017)

$$t_{\dot{X}} = \frac{4\pi d^2 G M_{\mathrm{core}}^2 X(1+X)}{\eta\pi L_{\mathrm{HE}}}\frac{1}{R_{\mathrm{p}}^3}, \tag{9}$$

where d is the orbital radius of the planet, and L_{HE} is the stellar high-energy flux. We consider a Sun-like star, as in our disc model above.

To determine the final envelope mass fractions, we do not explicitly evolve the atmospheres in time. Instead, we use the fact that most of the mass-loss happens during the first ∼ 100 Myr after disc dispersal since during this period the stellar high-energy flux L_{HE} is saturated ($L_{\mathrm{HE}} = L_{\mathrm{sat}} \sim 10^{-3.5}L_{\odot}$ for a Sun-like star), while after this time it quickly decays (Jackson, Davis & Wheatley 2012; Tu et al. 2015).

Thus, if a planet's mass-loss time-scale $t_{\dot{X}}$ is initially (i.e. at the time of disc dispersal) longer than 100 Myr, the planet will not suffer significant mass-loss. Here, we assume that such a planet remains unchanged by the photoevaporation.

On the other hand, a planet with initial $t_{\dot{X}} < 100$ Myr will lose mass. Now, for a given core mass M_{core} and orbital distance d, the mass-loss time-scale as a function of the envelope mass fraction, $t_{\dot{X}}(X)$, peaks at $X \equiv X_{\mathrm{peak}}$ of a few per cent, decreasing for both smaller and larger X (Owen & Wu 2017). Thus, for a planet with an initially small accreted envelope mass fraction ($X < X_{\mathrm{peak}}$), the mass-loss further shortens the loss time-scale. Hence, if such a planet's initial mass-loss time-scale is less than 100 Myr it is subject to run-away mass-loss, and we assume it is completely stripped

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

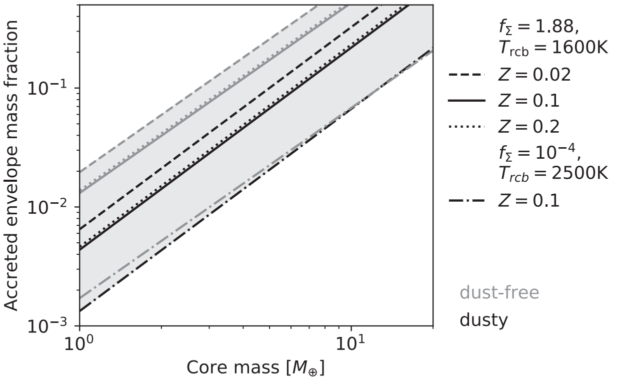


Figure 5 Envelope mass fraction of atmospheres accreted in 1 Myr as a function planet core mass for dusty (black lines) and dust-free (grey lines) atmospheres and a variety of metallicities *Z*, gas surface density factors f_Σ and (in the case of dust-free atmospheres) radiative–convective boundary temperatures $T_{\rm rcb}$, as indicated in plot legend. The grey region indicates the total range of expected envelope mass fractions (except for those that would reach an envelope mass fraction of $X = 0.5$ within 1 Myr and thereby expected to undergo run-away accretion to form gas giants; these are not shown).

of its atmosphere. Conversely, the mass-loss time-scale of a large atmosphere ($X > X_{\rm peak}$) increases as it loses mass, tending towards the peak value of $t_{\dot{X}}(X_{\rm peak})$. If $M_{\rm core}$ and *d* are such that the latter peak time-scale is ≥100 Myr, we assume that a planet with such a large initial atmosphere will stall at an envelope mass fraction *X* corresponding to a loss time-scale of $t_{\dot{X}} = 100$ Myr. However, if the peak time-scale is <100 Myr, the mass-loss time-scale can increase to this peak value and then descend into the run-away regime on the other side while the stellar activity is still saturated; thus, such a planet will lose its entire atmosphere regardless of the accreted *X*. This simple prescription adequately captures the basic physics of atmospheric photoevaporation (see far left-hand panel of fig. 6 in Owen & Wu 2017).

3.2 Results

3.2.1 Accretion of planetary atmospheres

Using the scaling relations (7,8), we calculate the envelope mass fractions that planetary cores of various masses accrete from the gas disc in 1 Myr. Results are shown in Fig. 5 for both dusty and dust-free atmospheres of various metallicities, ranging from solar ($Z = 0.02$) to the metallicity of Neptune's atmosphere ($Z = 0.2$; Karkoschka & Tomasko 2011), gas surface density factors $f_\Sigma = 10^{-4}$–1.88, and radiative–convective boundary temperatures $T_{\rm rcb} = 1600$–2500 K.

The dependence on metallicity *Z* is non-monotonous, as the accreted envelope mass fraction depends separately on the metallicity and on the mean molecular weight (which is set by the metallicity). The smallest accreted atmospheres, with the rest of the parameters fixed, have $Z \sim 0.1$.

The radiative–convective boundary temperature $T_{\rm rcb}$ is expected to be roughly constant in dusty atmospheres, so we only explore the effect of this parameter in dust-free atmospheres. In the latter, $T_{\rm rcb}$ is related to the temperature of the environment. Additionally, accretion of both dust-free and dusty atmospheres depends on the density of the environment. Here, we are interested in atmospheres that are accreted in the inner disc, near or inwards of the pressure maximum. The location of the pressure maximum is determined by the extent of thermal ionization of potassium in our disc model, and so this corresponds to disc temperatures of $T \gtrsim 1000$ K, regardless of the exact disc parameters (e.g. gas accretion rate). For a disc temperature of $T \sim 1000$ K, numerical models of the accreting atmospheres give $T_{\rm rcb} \sim 1600$ K (Lee & Chiang 2015), which thus sets a lower bound on $T_{\rm rcb}$ for our calculations. Moreover, the location of the pressure maximum is also where the gas surface density is highest (see Fig. 1). Colder atmospheres in a more dense environment accrete more. So, to show the maximum accreted atmospheres in an MRI-accreting disc, we plot a set of dusty and dust-free atmospheres (of various metallicities *Z*) for the maximum $f_\Sigma = 1.88$, and the minimum $T_{\rm rcb} = 1600$ K (the latter refers only to the dust-free atmospheres). Conversely, the maximum temperature at which equation (8) is valid (due to the limitations of the opacity tables used by Lee et al. 2014) is $T_{\rm rcb} = 2500$ K,[3] and the minimum gas surface density in our inner disc model with respect to the minimum mass solar nebula is $f_\Sigma \approx 10^{-4}$ (corresponding to the inner disc edge in Fig. 1). Hotter atmospheres in lower density environments accrete less, and so to show the smallest accreted dust-free atmospheres, we plot the $f_\Sigma = 10^{-4}$ atmospheres, with $T_{\rm rcb} = 2500$ K for the dust-free atmospheres, and with metallicity $Z = 0.1$ (since, as noted above, $Z \sim 0.1$ yields the smallest atmosphere for any given set of other parameters).

Finally, atmospheres that grow above a threshold of $X = 0.5$ undergo run-away accretion and end up as gas giants (Rafikov 2006). The scaling relations (7,8) are not applicable in this case, neither are we interested in the much rarer close-in Jupiters. Therefore, Fig. 5 is cut-off at $X = 0.5$, and the grey region indicates how small or large super-Earth/mini-Neptune atmospheres may be at the time of disc dispersal. Overall we see that, if the cores are formed 1 Myr before the dispersal, run-away accretion is avoided for the majority of relevant core masses, but they do accrete significant gaseous envelopes of up to a few × 10 per cent of core mass.

The envelopes shown in Fig. 5 have been calculated assuming that the accretion lasts for 1 Myr. Since disc lifetimes can be longer (Mamajek 2009), these envelopes could be conservative estimates if planets form sooner than 1 Myr before disc dispersal. If, for example, the envelopes are accreted for 5 Myr, the envelope mass will double. We do not expect the results to be very sensitive to the exact disc parameters, as long as the cores accrete their atmospheres in a thermally ionized MRI-active inner disc. We note, however, that our disc model implies that the extent of such inner disc does not encompass all observed sub-Neptunes for all relevant accretion rates; e.g. the gas pressure maximum is at an orbital period longer than 100 d only for gas accretion rates of $\dot{M}_{\rm g} \gtrsim 3 \times 10^{-9}\,{\rm M}_\odot\,{\rm yr}^{-1}$. Thus, a planet with a longer orbital period might spend at least some time in a colder MRI-dead zone, which we do not take into account.

3.2.2 Photoevaporation of planetary atmospheres

To further check the consistency of core accretion in the MRI-accreting inner disc with observations, we need to consider whether these accreted atmospheres survive photoevaporation. We calculate the final (remaining) envelope mass fraction of the minimum and

---

[3] Note that assuming $T_{\rm rcb}$ is directly proportional to the disc temperature *T*, and scaling from the numerical models' result that $T_{\rm rcb} \sim 1600$ K corresponds to $T \sim 1000$ K, yields $T \sim 1500$ K for $T_{\rm rcb} \sim 2500$ K. The disc temperature in our model only exceeds 1500 K at radii <0.1 au, so a maximum $T_{\rm rcb}$ of 2500 K is indeed roughly valid over most of our inner disc.

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

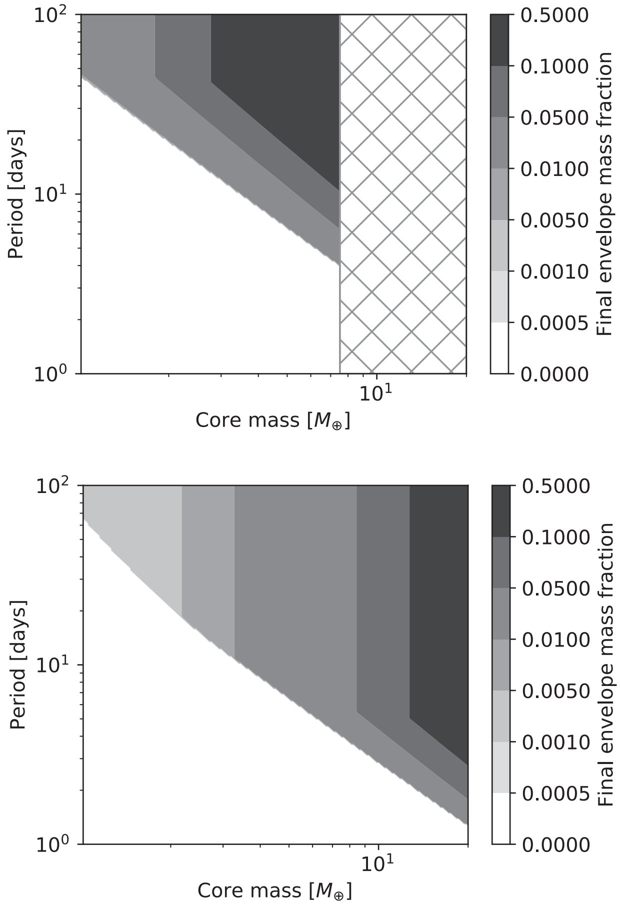


Figure 6 Maximum (*top*) and minimum (*bottom*) envelope mass fraction of the atmospheres after accounting for photoevaporation, as functions of planet core mass and orbital period. In the top figure, the hatched region indicates the core masses for which the planets would undergo run-away accretion and are thus excluded from here.

maximum possible accreted atmospheres (corresponding, respectively, to dusty atmospheres with $f_\Sigma = 10^{-4}$ and $Z = 0.1$, and dust-free atmospheres with $Z = 0.02$, $f_\Sigma = 1.88$, and $T_{\rm rcb} = 1600\,{\rm K}$) for each core mass and as a function of orbital period. Results are shown in Fig. 6. In the case of the maximum accreted atmospheres (top panel), the atmospheres would undergo run-away accretion for core masses $\gtrsim 8\,{\rm M}_\oplus$ (indicated by the hatched region), which are thus excluded here.

The figures show that the orbital period at which the atmosphere can be completely evaporated decreases with increasing core mass, and cores that retain their atmospheres generally evolve towards a 1 per cent envelope mass fraction as expected from theory. At 100 d, the atmospheres are unaffected by photoevaporation, and at periods shorter than 1 d all planets are predicted to end up as bare cores. Massive cores are predicted to keep their 1–50 per cent atmospheres at the majority of orbital periods, and planets with Earth-mass cores are safe from complete mass-loss at periods larger than 50 d.

Note that here the orbital period determines the level of high-energy flux that planet experiences and planet equilibrium temperature (and thus planet radius), but does not directly reflect variations in temperature and density of the protoplanetary disc inside which the atmospheres were accreted. As discussed above, the effect of the disc temperature on the accreted envelope mass fraction is negligible for dusty atmospheres. For dust-free atmospheres, the dependence is monotonous and the extent of the effect is explored by considering the minimum $T_{\rm rcb}$ we expect in the inner disc, and the maximum $T_{\rm rcb}$ for which the scaling relations (7,8) are valid. Similarly, the dependence on the disc density is explored by considering the smallest and largest values of the ratio of the MRI-disc model and the minimum mass solar nebula surface densities. Thus, by calculating the effect of photoevaporation on both the minimum and maximum accreted atmospheres shown in Fig. 5 for each core mass, we also encompass the possible range of disc densities and temperatures.

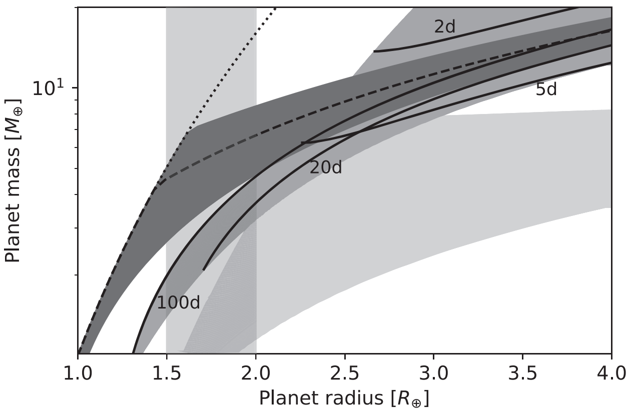


Figure 7 Mass–radius relationship for sub-Neptune planets: Earth-like composition solid cores (dotted line), probabilistic fit to observations mean value (dashed line) and scatter (solid dark grey region; Wolfgang et al. 2016), region of low planet occurrence rates from the observed radius distribution of planets (sheer grey region) (Fulton et al. 2017), and predictions from the minimum and maximum accreted atmospheres and photoevaporation (medium grey and light grey region, respectively) with orbital period contours for the minimum accreted atmospheres (solid-line contours).

3.3 Comparison to observations

Overall, Fig. 6 shows that the envelopes formed in a gas-poor inner disc due to the MRI survive photoevaporation for a large range of orbital periods, and the low gas surface densities are not a hindrance to the formation of mini-Neptunes. On the contrary, the final envelope mass fractions of the planets that do keep their atmospheres are typically overestimated. The planets with core mass larger than $2\,{\rm M}_\oplus$ are predicted to either have a > 1 per cent atmosphere or to be completely evaporated. On the other hand, from the observations, the typical envelope mass fraction of mini-Neptunes that hold on to their atmospheres is 1 per cent (Wolfgang & Lopez 2015). To look into this further, we compare the predictions of our calculations against the observed mass–radius relationship for sub-Neptune planets in Fig. 7, and against measured masses and radii of individual sub-Neptune planets in Fig. 8.

For the observations in Fig. 7, we show the probabilistic best-fitting mass–radius relationship of Wolfgang, Rogers & Ford (2016): a power law $M/{\rm M}_\oplus = 2.7(R/{\rm R}_\oplus)^{1.3}$ (indicated by the dashed line) with a standard deviation of $\pm 1.9\,{\rm M}_\oplus$ due to an intrinsic scatter in planet mass (the dark grey region), and an upper limit constraint on the planet density corresponding to a mass–radius relationship for solid cores of Earth-like composition $M/{\rm M}_\oplus = (R/{\rm R}_\oplus)^4$ (dotted line; Valencia et al. 2010). Additionally, the above mass–radius relationship does not capture a significant feature of the observed radius distribution of sub-Neptunes, a decrease in occurrence rates

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

of planets with radii of 1.5–2 $R_\oplus$ (indicated here by the sheer grey region; Fulton et al. 2017).

To show the predictions of our atmospheric calculations in the mass–radius plane, we take the calculated envelope mass fraction as a function of core mass and period and re-calculate the planet radius at the planet age of 5 Gyr as a function of core mass and period, using the same simple atmospheric evolution model of Owen & Wu (2017). We show the results for the minimum and maximum accreted atmospheres (Fig. 6, and excluding the completely evaporated planets) in Fig. 7 (the medium and the light grey region, respectively). Note that the light grey region has a cut-off at about 8 $M_\oplus$ because we exclude massive cores that, given the parameters of the maximum accreted atmospheres, would be subject to run-away accretion. The solid line contours show how the planet mass and radii change as a function of period for the minimum accreted atmospheres. At the orbital period of 100 d, the planets are largely unaffected by the atmospheric loss, and closer to the star the photoevaporation removes atmospheres of the lower mass planets entirely. For the planets that keep their atmospheres at large periods, a decrease in period means little to no change in planet mass. Consequently, for these planets a decrease in period results in an increase in planet radius as atmospheres are hotter and more expanded closer to the star due to stronger stellar irradiation. At small periods, the atmospheric loss is significant for all planets, and the trend is reversed.

It is clear from Fig. 7 that for planets with radii $R \lesssim 2.3\,R_\oplus$ the core accretion of atmospheres in the inner disc predicts larger planet radii than those observed, due to the overestimated envelope mass fractions. The predicted atmospheres are massive enough to populate the region corresponding to planet radii of 1.5–2 $R_\oplus$, which is inconsistent with the observed decrease in planet occurrence rates at those radii (sheer grey region). For planets with $R \gtrsim 2.3\,R_\oplus$, there is a region in which the observed (dark grey) and the predicted (medium and light grey) mass–radius relationships overlap. This overlap corresponds to the (minimum accreted) predicted atmospheres for orbital periods between 20 and 100 d, and a narrow range of short orbital periods (2–5 d). Notably, even for the minimum accreted atmospheres, the planet radii, at fixed planet mass, are smaller than those observed only for significant high-energy fluxes at orbital periods of less than about 2 d. Taking into account the full range of accreted atmospheres (up to the maximum accreted atmospheres shown in light grey) further suggests that the predicted atmospheres are typically larger than the atmospheres of the observed sub-Neptunes.

We further compare the predictions of our calculations to sub-Neptune planets with measured masses and radii (taken from Wolfgang et al. 2016, excluding the planets where only the upper limit on the mass was known). The observed and the predicted radii and masses are shown in Fig. 8 in four panels corresponding to four orbital period bins. As in Fig. 7, the medium and light grey regions correspond to the predictions from the minimum and maximum atmosphere mass models, respectively. To facilitate comparison against the planets that are bare solid cores in each period bin, here we also show the core masses that are predicted to lose their entire atmospheres in a given period bin (the grey lines shown below the dotted lines that represent the Earth-like composition mass–radius relationship). Fig. 8 shows that the masses of the predicted bare cores and the period at which photoevaporation can strip them are largely consistent with those observed. That is, there are no observed planets consistent with the Earth-like composition that are (significantly) more massive than the largest core that the photoevaporation can strip (the upper limit of the

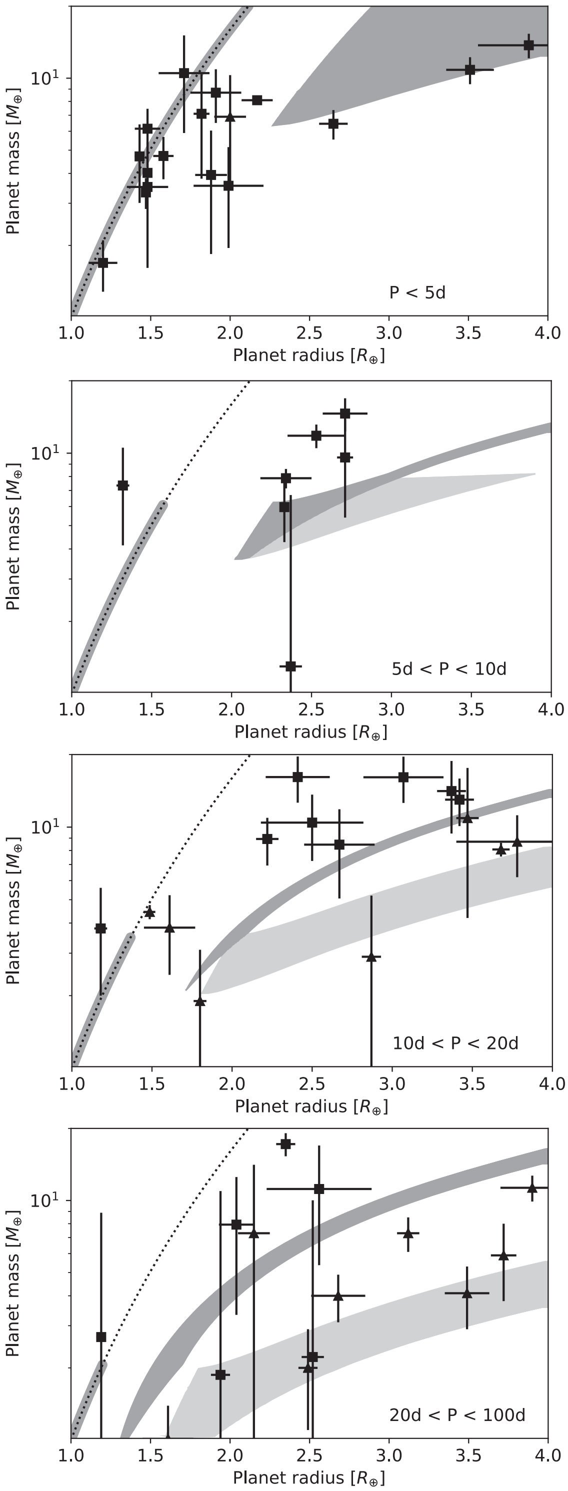


Figure 8 Sub-Neptune planets with measured masses and radii (square markers if the mass was determined using the radial velocity method, and triangles if the mass was determined using transit timing variations), with uncertainties as listed in Wolfgang et al. (2016), Earth-like composition solid cores (dotted line), and predictions from the minimum and maximum atmosphere models (medium grey and light grey region respectively), in period bins as indicated in plot labels. This figure indicates that while the period range at which planets can be stripped by photoevaporation is consistent with the data, our planets typically have larger H/He envelopes that expected.

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

grey line, the predicted bare cores) in each period bin. Fig. 8 also explicitly demonstrates that for the planets that maintain their atmospheres against the photoevaporation, the predicted planet radii are consistent with or larger than those observed for the majority of the planets. At long orbital periods (20–100 d), all planets except one are consistent, within the observational uncertainties, with the predictions (the minimum accreted atmospheres in medium grey, the maximum accreted atmospheres in light grey, and the region in between). At intermediate periods (5–20 d), about a third of planets that are not bare cores have radii smaller than the predicted radii at the same mass. Finally, at short periods of less than 5 d, there are noticeably five planets with radii of ∼1.8–2 $R_\oplus$ that are neither consistent with the mass–radius relationship of rocky cores, nor with the presence of H/He envelopes. This suggests, potentially, that the cores of these planets could contain significant amounts of ice. Still, majority of the short-period planets are consistent with the predictions. Additionally, we reiterate that, while there might be exceptions, the radius distribution of sub-Neptunes is consistent with cores being largely rocky (Owen & Wu 2017). Therefore, these results confirm our inference that typically our planets accrete too much gas. In Section 4, we suggest possible explanations for this result, such as incorrect assumptions of quasi-hydrostatic accretion and negligible heating from planetesimal accretion, or missing mass-loss mechanisms that might act during, or after, disc dispersal.

Overall, the atmospheres accreted in the inner disc are typically in agreement with or larger than those observed, with the exception of planets with significant high-energy fluxes within a very narrow range. This is because core accretion is so efficient that considerable atmospheres can be accreted in the hot and low-density MRI-accreting inner disc and also maintained against photoevaporation.

4 DISCUSSION AND CONCLUSIONS

In this paper, we have investigated two aspects of planet formation in the inner disc that is viscously accreting due to the MRI. We considered the earliest phase of planet formation that is evolution of dust particles, and the final phase that is the shaping of the planetary atmospheres. We present simple calculations that include both the accretion and subsequent photoevaporation of close-in super-Earth/mini-Neptune planets. By coupling these processes to perform end-to-end calculations, we are able to assess the viability of the *in situ* formation model for close-in planets.

In Section 2, we consider the evolution of dust grains in the inner disc that is viscously accreting due to the MRI (Mohanty et al. 2018), which features a local gas pressure maximum at the orbital distance of a few tenths of AU. Taking into account the effect of the MRI-induced turbulence on the dust grain size, we find that fragmentation of particles due to turbulent relative velocities limits the particle size to below few millimetres. As a result, the particles are not efficiently accumulated inside the pressure maximum as hypothesized by Chatterjee & Tan (2014) and Hu et al. (2018). Regardless of that, as the particles become well coupled to the gas, the radial drift is negated in the inner disc, and the dust-to-gas ratio is enhanced throughout the inner disc. Thus, the local gas pressure maximum might play a lesser role in the *in situ* planet formation than previously thought.

The pressure maximum is, however, still the location of a local density maximum in both gas and dust. We explored if the resulting inner disc structure that is enriched in dust could be susceptible to the onset of the streaming instabilities. This pathway to planetesimals seems to be viable only in a narrow region near the pressure (and density) maximum, for the chosen disc parameters.

The gas is not, however, evolved in this work and effects of the growing amounts of dust on to the MRI have not been taken into account. Dust grains lower the gas ionization levels by absorbing free charges and enhancing recombination rates, as ions recombine on the grains (Draine & Sutin 1987; Ilgner & Nelson 2006). Charged grains are not themselves well coupled to the magnetic field as they are too massive, and so their presence promotes the non-ideal MHD effects, which can suppress the MRI (Sano et al. 2000; Ilgner & Nelson 2006; Wardle 2007; Salmeron & Wardle 2008; Bai & Goodman 2009). The likely result of taking dust effects into account is thus weakened turbulence, and the change of the disc structure in the longer term. The consequences can only be investigated by modelling both the gas and the dust self-consistently. Here, we sketch out a potential scenario by considering the relevant time-scales.

Assuming that the steady-state solution of the gas structure (Fig. 1) is reached before dust starts affecting the MRI, we expect the dust enhancement of the inner disc to ensue. At a certain dust-to-gas ratio the dust will suppress the MRI, and we expect the levels of turbulence to adapt almost instantly, as the time-scale of the magnetic field regeneration is the orbital time-scale $t_{\rm orb}$ (e.g. Balbus & Hawley 1991). With the decreasing levels of turbulence, the dust particle size will rapidly grow due to particle coagulation. The growth due to coagulation happens on the time-scales of $\Sigma_{\rm g}/\Sigma_{\rm d} t_{\rm orb}$ (e.g. Brauer, Dullemond & Henning 2008), so faster than $10^2 t_{\rm orb}$ if the inner disc is indeed enriched in dust.

The gas disc structure would not change over such short time-scales. But, in the absence of the viscous heating due to the MRI, the disc will cool at a time-scale of 10^2–$10^4 t_{\rm orb}$ (the equilibrium thermal time-scale from roughly the pressure minimum to the pressure maximum in the model considered here; Mohanty et al. 2018). The pressure profile would also follow this time-scale, as the vertical hydrostatic equilibrium is established quickly on the orbital time-scale ($t_{\rm orb}$). This would likely result in the pressure maximum moving radially inwards.

Concurrently, due to larger particle size and lower turbulent stirring the particles would vertically settle towards the mid-plane and radially towards the pressure maximum, increasing both the surface density and the mid-plane bulk density dust-to-gas ratio there. Such formation of a ring of solids could potentially trigger formation of larger bodies, such as planetesimals (as hypothesized by e.g. Chatterjee & Tan 2014). The larger particle size and the settling towards the mid-plane would likely trigger the SI (inwards of the pressure maximum; see Fig. 4). However, it is unclear if this could lead to the formation of planetesimals, as a gravitational collapse is unlikely due to the low bulk dust densities and high Roche density in the inner disc. Moreover, as discussed above, the pressure maximum is expected to move radially inwards, and so will the accumulated dust, whereas the Roche density steeply increases inwards. Thus, overtime the pressure maximum would need to accumulate significantly more dust to cross the gravitational collapse threshold.

The gas accretion rate and the gas surface density will change slowly in comparison to the above processes, on the long viscous time-scale, $\sim 10^3$–$10^5 t_{\rm orb}$ from the pressure minimum to the pressure maximum in the model considered here (Mohanty et al. 2018). Suppression of the MRI would lead to increased amounts of gas in the inner disc on this time-scale. However, if planetesimals are formed, this would clear the inner disc of the dust grains and the MRI could be induced again, decreasing the gas surface density and moving the pressure maximum outwards. At this stage, it is unclear whether these competing processes are balanced in a steady state, or

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

the behaviour of the inner disc is dynamic and quasi-periodic. Such a determination can only be investigated through self-consistent modelling of dust, gas, and the MRI.

At high dust-to-gas ratios, the dust also becomes dynamically important, and affects the gas disc structure through the drag back reaction (Nakagawa, Sekiya & Hayashi 1986). The gas rotation profile is then driven towards Keplerian, and as a result the radial gas pressure profile flattens. This, in turn, slows down the radial drift of dust particles. If dust already pile ups in the inner disc due to radial drift being slower than in the outer disc, the dust back reaction amplifies the effect (Drążkowska et al. 2016). In this work, dust enhancement is driven by the dust grains already being completely coupled to the gas in the innermost disc, and thus the effect of dust back reaction would be limited. However, the back reaction would become important if the dust grains grow (e.g. due to the suppression of the MRI-induced turbulence discussed above), especially near the pressure maximum. If the dust grains grow in the innermost disc where the pressure gradient is negative, the backreaction would slow down the loss of dust to the star. However, the back reaction would also limit the concentration of dust that can be achieved at the pressure maximum since it acts to flatten the overall gas pressure profile (Taki, Fujimoto & Ida 2016).

Furthermore, if super-Earth and mini-Neptune cores indeed form *in situ*, would the inferred low gas surface densities due to the MRI allow them to acquire the observed $0.1-10$ per cent envelope mass fractions? We find that they would. In fact, even after accounting for atmospheric evaporation, the calculated atmospheres tend to overestimate the observed ones.

Could the atmospheric accretion in the MRI-implied disc and the observations be brought into agreement, without invoking an assumption that cores form just before the beginning of disc dispersal (e.g. Ikoma & Hori 2012; Lee & Chiang 2016)? The calculations shown here do not include several effects that could contribute.

First of all, for core masses smaller than $10\,\mathrm{M}_{\oplus}$, the discrepancy could be explained by the 'boil-off' or core-powered mass-loss (Owen & Wu 2016; Ginzburg, Schlichting & Sari 2018; see also Ikoma & Hori 2012), a process in which a planet atmosphere that had not cooled and contracted before the disc dispersal loses its mass. Upon the dispersal, the stellar continuum radiation illuminates the planet and launches a Parker wind. The mass-loss causes rapid contraction of the atmosphere, and the contraction in turn shuts off the mass-loss. Planets that start out with few tens of per cent atmospheres, may be left with 1 per cent after the boil-off. This process precedes the mass-loss caused by the stellar high-energy flux considered above, and can operate at larger distances from the star.

Secondly, the scaling relations we use to calculate the accreted atmospheres are derived assuming no sources of heating due to planetesimal accretion, or due to heat deposited in the hypothesized final stage of giant mergers of planetary embryos. The latter could be released for several kyr (e.g. Inamdar & Schlichting 2015), lowering the cooling rate of the atmosphere, and thus allowing less gas to be accreted. Furthermore, the scaling relations assume that the gas inside the planet's Hill sphere is bound and static. 3D simulations suggest this may not be true and that high-entropy disc material is recycled between the envelope and disc (e.g. Fung, Artymowicz & Wu 2015; Ormel, Shi & Kuiper 2015; Cimerman, Kuiper & Ormel 2017) potentially modifying the atmospheric cooling rate.

Finally, if the giant mergers happen between planets, after the disc has fully dispersed, they would likely result in significant atmospheric mass-loss. Head-on collisions between Earth/super-Earth-sized planets with few per cent atmospheres can remove tens of per cent of the total atmospheric mass (Liu et al. 2015; Inamdar & Schlichting 2016).

Nevertheless, to avoid the run-away accretion for more massive cores, the low gas surface densities the MRI provides are favourable compared to the MMSN environment. Furthermore, gas-poor conditions in this case are provided in a long-lived state, and not in a transient phase (e.g. a transition disc, as proposed by Lee & Chiang 2016).

In summary, our results support the hypothesis that the MRI-driven accretion in the inner protoplanetary disc could lead to *in situ* planet formation. However, there are several avenues that need to be explored in more detail until we can make quantitative predictions. In particular, the feedback between the enhancement of the dust in the inner disc and the suppression of the MRI, the feedback between the dust enhancement and the gas dynamics, and the role of disc dispersal and boil-off/core-powered mass-loss in shaping the final envelope masses of super-Earths/mini-Neptunes are issues that deserve closer study.

ACKNOWLEDGEMENTS

We thank the reviewer for helpful suggestions that improved the manuscript. We thank Jonathan Tan, Lauren Weiss, and Eve Lee for helpful discussions. MRJ is funded by the President's PhD scholarship of the Imperial College London and the 'Dositeja' stipend from the Fund for Young Talents of the Serbian Ministry for Youth and Sport. JEO is supported by a Royal Society University Research Fellowship.

REFERENCES

Andrews S. M. et al., 2012, ApJ, 744, 162
Armitage P. J. , 2018, in Deeg H., Belmonte J., eds, Handbook of Exoplanets, Springer, Cham
Bai X.-N., 2011, ApJ, 739, 50
Bai X.-N., Goodman J., 2009, ApJ, 701, 737
Bai X.-N., Stone J. M., 2011, ApJ, 736, 144
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Baruteau C. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. Univ. Arizona Press, Tucson, AZ, p. 667
Batalha N. M. et al., 2013, ApJS, 204, 24
Beitz E., Güttler C., Blum J., Meisner T., Teiser J., Wurm G., 2011, ApJ, 736, 34
Birnstiel T., Dullemond C. P., Brauer F., 2010, A&A, 513, A79
Birnstiel T., Klahr H., Ercolano B., 2012, A&A, 539, A148
Blum J., Münch M., 1993, Icarus, 106, 151
Borucki W. J. et al., 2011, ApJ, 736, 19
Brauer F., Dullemond C. P., Henning T., 2008, A&A, 480, 859
Bukhari Syed M., Blum J., Wahlberg Jansson K., Johansen A., 2017, ApJ, 834, 145
Carrera D., Johansen A., Davies M. B., 2015, A&A, 579, A43
Carrera D., Gorti U., Johansen A., Davies M. B., 2017, ApJ, 839, 16
Chatterjee S., Tan J. C., 2014, ApJ, 780, 53
Chiang E., Laughlin G., 2013, MNRAS, 431, 3444
Cimerman N. P., Kuiper R., Ormel C. W., 2017, MNRAS, 471, 4662
Cossou C., Raymond S. N., Pierens A., 2013, A&A, 553, L2
Cossou C., Raymond S. N., Hersant F., Pierens A., 2014, A&A, 569, A56
Draine B. T., Sutin B., 1987, ApJ, 320, 803
Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45
Drążkowska J., Windmark F., Dullemond C. P., 2013, A&A, 556, A37
Drążkowska J., Alibert Y., Moore B., 2016, A&A, 594, A105
Dzyurkevich N., Flock M., Turner N. J., Klahr H., Henning T., 2010, A&A, 515, A70

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Fressin F. et al., 2013, ApJ, 766, 81
Fulton B. J. et al., 2017, AJ, 154, 109
Fulton B. J., Petigura E. A., 2018, AJ, 156, 264
Fung J., Lee E. J., 2018, ApJ, 859, 126
Fung J., Artymowicz P., Wu Y., 2015, ApJ, 811, 101
Gammie C. F., 1996, ApJ, 457, 355
Ginzburg S., Schlichting H. E., Sari R., 2018, MNRAS, 476, 759
Goldreich P., Schlichting H. E., 2014, AJ, 147, 32
Hansen B. M. S., Murray N., 2012, ApJ, 751, 158
Hansen B. M. S., Murray N., 2013, ApJ, 775, 53
Hayashi C., 1981, Prog. Theor. Phys. Suppl., 70, 35
Hu X., Zhu Z., Tan J. C., Chatterjee S., 2016, ApJ, 816, 19
Hu X., Tan J. C., Zhu Z., Chatterjee S., Birnstiel T., Youdin A. N., Mohanty S., 2018, ApJ, 857, 20
Ida S., Lin D. N. C., 2008, ApJ, 685, 584
Ikoma M., Hori Y., 2012, ApJ, 753, 66
Ilgner M., Nelson R. P., 2006, A&A, 445, 205
Inamdar N. K., Schlichting H. E., 2015, MNRAS, 448, 1751
Inamdar N. K., Schlichting H. E., 2016, ApJ, 817, L13
Isella A., Pérez L. M., Carpenter J. M., 2012, ApJ, 747, 136
Izidoro A., Ogihara M., Raymond S. N., Morbidelli A., Pierens A., Bitsch B., Cossou C., Hersant F., 2017, MNRAS, 470, 1750
Jackson A. P., Davis T. A., Wheatley P. J., 2012, MNRAS, 422, 2024
Johansen A., Youdin A., 2007, ApJ, 662, 627
Johansen A., Youdin A., Mac Low M.-M., 2009, ApJ, 704, L75
Jontof-Hutter D. et al., 2016, ApJ, 820, 39
Karkoschka E., Tomasko M. G., 2011, Icarus, 211, 780
Kley W., Nelson R. P., 2012, ARA&A, 50, 211
Kothe S., Güttler C., Blum J., 2010, ApJ, 725, 1242
Kretke K. A., Lin D. N. C., Garaud P., Turner N. J., 2009, ApJ, 690, 407
Lee E. J., Chiang E., 2015, ApJ, 811, 41
Lee E. J., Chiang E., 2016, ApJ, 817, 90
Lee E. J., Chiang E., 2017, ApJ, 842, 40
Lee E. J., Chiang E., Ormel C. W., 2014, ApJ, 797, 95
Lee E. J., Chiang E., Ferguson J. W., 2018, MNRAS, 476, 2199
Liu B., Ormel C. W., Lin D. N. C., 2017, A&A, 601, A15
Liu S.-F., Hori Y., Lin D. N. C., Asphaug E., 2015, ApJ, 812, 164
Lopez E. D., Fortney J. J., 2013, ApJ, 776, 2
Lopez E. D., Fortney J. J., 2014, ApJ, 792, 1
Lundkvist M. S. et al., 2016, Nat. Commun., 7, 11201
Mamajek E. E., 2009, in Usuda T., Tamura M., Ishii M., eds, AIP Conf. Proc. Vol. 1158, Exoplanets and Disks: Their Formation and Diversity. Am. Inst. Phys., New York. p. 3
Manara C. F., Robberto M., Da Rio N., Lodato G., Hillenbrand L. A., Stassun K. G., Soderblom D. R., 2012, ApJ, 755, 154
Masset F. S., Morbidelli A., Crida A., Ferreira J., 2006, ApJ, 642, 478
Meru F., Geretshauser R. J., Schäfer C., Speith R., Kley W., 2013, MNRAS, 435, 2371
Mizuno H., 1980, Prog. Theor. Phys., 64, 544
Mohanty S., Jankovic M. R., Tan J. C., Owen J. E., 2018, ApJ, 861, 144
Mulders G. D., Pascucci I., Apai D., 2015, ApJ, 798, 112
Mulders G. D., Pascucci I., Apai D., Ciesla F. J., 2018, AJ, 156, 24
Nakagawa Y., Sekiya M., Hayashi C., 1986, Icarus, 67, 375
Ogihara M., Morbidelli A., Guillot T., 2015, A&A, 578, A36
Ogihara M., Kokubo E., Suzuki T. K., Morbidelli A., 2018, A&A, 615, A63
Ormel C. W., Cuzzi J. N., 2007, A&A, 466, 413
Ormel C. W., Shi J.-M., Kuiper R., 2015, MNRAS, 447, 3512
Owen J. E., 2014, ApJ, 789, 59
Owen J. E., Wu Y., 2013, ApJ, 775, 105
Owen J. E., Wu Y., 2016, ApJ, 817, 107
Owen J. E., Wu Y., 2017, ApJ, 847, 29
Panić O., Hogerheijde M. R., Wilner D., Qi C., 2009, A&A, 501, 269
Pinilla P., Benisty M., Birnstiel T., 2012, A&A, 545, A81
Powell D., Murray-Clay R., Schlichting H. E., 2017, ApJ, 840, 93
Rafikov R. R., 2006, ApJ, 648, 666
Rogers L. A., Bodenheimer P., Lissauer J. J., Seager S., 2011, ApJ, 738, 59
Rosenfeld K. A., Andrews S. M., Wilner D. J., Kastner J. H., McClure M. K., 2013, ApJ, 775, 136
Salmeron R., Wardle M., 2008, MNRAS, 388, 1223
Sano T., Miyama S. M., Umebayashi T., Nakano T., 2000, ApJ, 543, 486
Schräpler R., Blum J., 2011, ApJ, 734, 108
Schräpler R., Blum J., Seizinger A., Kley W., 2012, ApJ, 758, 35
Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33
Squire J., Hopkins P. F., 2018, ApJ, 856, L15
Suzuki T. K., Ogihara M., Morbidelli A., Crida A., Guillot T., 2016, A&A, 596, A74
Takeuchi T., Lin D. N. C., 2002, ApJ, 581, 1344
Taki T., Fujimoto M., Ida S., 2016, A&A, 591, A86
Tu L., Johnstone C. P., Güdel M., Lammer H., 2015, A&A, 577, L3
Valencia D., Ikoma M., Guillot T., Nettelmann N., 2010, A&A, 516, A20
Wardle M., 2007, Ap&SS, 311, 35
Weidenschilling S. J., 1977, MNRAS, 180, 57
Weidling R., Güttler C., Blum J., Brauer F., 2009, ApJ, 696, 2036
Winn J. N., Fabrycky D. C., 2015, ARA&A, 53, 409
Wolfgang A., Lopez E., 2015, ApJ, 806, 183
Wolfgang A., Rogers L. A., Ford E. B., 2016, ApJ, 825, 19
Wu Y., 2018, preprint (arXiv:1806.04693)
Yang C.-C., Johansen A., Carrera D., 2017, A&A, 606, A80
Youdin A. N., Chiang E. I., 2004, ApJ, 601, 1109
Youdin A. N., Goodman J., 2005, ApJ, 620, 459
Youdin A. N., Lithwick Y., 2007, Icarus, 192, 588
Youdin A. N., Shu F. H., 2002, ApJ, 580, 494
Öberg K. I., Murray-Clay R., Bergin E. A., 2011, ApJ, 743, L16

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/484/2/2296/5274141 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **482,** 4673–4686 (2019)
Advance Access publication 2018 November 10

doi:10.1093/mnras/sty3038

Observing substructure in circumstellar discs around massive young stellar objects

M. R. Jankovic,[1]★ T. J. Haworth,[1] J. D. Ilee,[2,3] D. H. Forgan,[4] C. J. Cyganowski,[4] C. Walsh,[3] C. L. Brogan,[5] T. R. Hunter[5] and S. Mohanty[1]

[1] *Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*
[2] *Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK*
[3] *School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK*
[4] *SUPA, School of Physics, Astronomy, University of St Andrews, North Haugh, St Andrews, Scotland KY16 9SS, UK*
[5] *NRAO, 520 Edgemont Rd, Charlottesville, VA 22903, USA*

Accepted 2018 November 6. Received 2018 November 6; in original form 2018 March 27

ABSTRACT

Simulations of massive star formation predict the formation of discs with significant substructure, such as spiral arms and clumps due to fragmentation. Here, we present a semi-analytic framework for producing synthetic observations of discs with substructure, in order to determine their observability in interferometric observations. Unlike post-processing of hydrodynamical models, the speed inherent to our approach permits us to explore a large parameter space of star and disc parameters, and thus constrain properties for real observations. We compute synthetic dust continuum and molecular line observations probing different disc masses, distances, inclinations, thermal structures, dust distributions, and number and orientation of spirals and fragments. With appropriate spatial and kinematic filtering applied, our models predict that Atacama Large Millimetre Array observations of massive young stellar objects at <5 kpc distances should detect spirals in both gas and dust in strongly self-gravitating discs (i.e. discs with up to two spiral arms and strong kinematic perturbations). Detecting spirals will be possible in discs of arbitrary inclination, either by directly spatially resolving them for more face-on discs (inclinations up to $\sim$50 deg), or through a kinematic signature otherwise. Clumps resulting from disc fragmentation should be detectable in the continuum, if the clump is sufficiently hotter than the surrounding disc material.

Key words: accretion, accretion discs – radiative transfer – circumstellar matter – stars: formation – stars: massive.

1 INTRODUCTION

Circumstellar discs have long been detected around low- to intermediate-mass ($M_* \lesssim 8\,\mathrm{M}_\odot$) young stars. These discs are known to accrete material on to the star, and are of increasing interest as the sites of planet formation (e.g. Williams & Cieza 2011; Armitage 2015; Casassus 2016; Morbidelli & Raymond 2016). In the higher mass regime ($M_* \gtrsim 8\,\mathrm{M}_\odot$), the very short pre-main-sequence lifetime (<10 Myr), coupled with the embedded, distant ($\geq$kpc) and relatively scarce nature of such stars, makes the detection of discs more difficult. Nevertheless, there is growing observational evidence for the presence of discs around massive young stellar objects (MYSOs; Hunter et al. 2014a; Johnston et al. 2015; Zapata et al. 2015; Chen et al. 2016; Ilee et al. 2016).

Investigating the nature and properties of discs around massive stars is important for a number of reasons. First, as a star gains mass, radiative and kinematic (i.e. stellar wind) feedback drives ambient material away and could effectively shut off accretion and hence limit the potential stellar mass (Kahn 1974; Nakano 1989; Edgar & Clarke 2004). The fact that we observe very massive stars (e.g. Massey, DeGioia-Eastwood & Waterhouse 2001) requires some mechanism that permits the star to continue to accrete. Numerical simulations have demonstrated that feedback can be overcome by accretion through a disc (e.g. Kuiper et al. 2010, 2011; Klassen et al. 2016; Rosen et al. 2016; Harries, Douglas & Ali 2017), with the feedback energy escaping through the poles. Therefore, circumstellar discs are pivotal in allowing massive stars to grow, and so characterizing them observationally is key.

Secondly, if gravitationally unstable, a circumstellar disc can produce substructure that can affect the accretion on to the star, and even fragment into gravitationally bound objects (discussed further

---

★ E-mail: m.jankovic16@imperial.ac.uk

© 2018 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

below). A Keplerian disc will be gravitationally unstable when the Toomre (1964) Q parameter satisfies

$$Q \equiv \frac{c_s \Omega}{\pi G \Sigma} < 1 \quad (1)$$

where c_s, Ω, and Σ are the sound speed, angular frequency, $(GM_*/R^3)^{1/2}$, and surface density at a distance R from a star of mass M_*. If we assume that the disc mass scales linearly with the stellar mass and the temperature in the deeper layers of the outer disc is roughly constant at $\sim$10 K (i.e. is independent of stellar mass), we would hence expect from equation (1) that $Q \propto M_*^{-1/2}$. Under this simple argument, discs around more massive YSOs are more susceptible to becoming gravitationally unstable (though instability in discs around lower mass stars certainly can arise; Pérez et al. 2016). Gravitational instability is therefore something that is important to consider in the study of discs around MYSOs.

Numerical models of gravitationally unstable discs predict spiral overdensities and fragmentation of the disc to form clumps (Kratter & Matzner 2006; Hall, Forgan & Rice 2017). This instability of the disc may drive up the accretion rate compared to regular viscous evolution (Laughlin & Bodenheimer 1994). Such an effect may provide an explanation for recent outbursting behaviour detected towards several MYSOs (Caratti o Garatti et al. 2017; Hunter et al. 2017). Furthermore, the resulting clumps may be the seeds of additional gravitationally bound objects such as stars or planets (Forgan et al. 2018b). If bound to the massive primary, future interaction (i.e. accretion) of such secondary objects can have consequences for the later stages of massive stellar evolution (for a review, see De Marco & Izzard 2017).

Observationally, however, it is extremely difficult to detect the immediate circumstellar environments of massive young stars due to their embedded nature, their relative scarcity, and their correspondingly large distances. Infrared (IR) interferometry and high-resolution near-infrared spectroscopy have revealed discs on scales of less than 1000 au around MYSOs (Bik & Thi 2004; Kraus et al. 2010; Wheelwright et al. 2010; Boley et al. 2013; Ilee et al. 2013; Ramírez-Tannus et al. 2017), but these techniques are limited to tracing only the inner regions of discs, and provide little information on the bulk of the circumstellar environment. Longer-wavelength interferometric observations allow access to the circum(proto)stellar environments of less evolved, more embedded MYSOs, but often probe larger spatial scales. In many cases, velocity gradients detected in millimetre and centimetre-wavelength molecular line observations trace large-scale (1000 s to $\gtrsim$10 000 au), massive ($M_{\mathrm{toroid}} \geq M_*$), non-equilibrium rotating structures known as 'toroids' (e.g. Cesaroni 2005; Cesaroni et al. 2006, 2007; Beuther & Walsh 2008; Beltrán et al. 2011; Cesaroni et al. 2011; Johnston et al. 2014; and references therein). Nevertheless, high angular resolution observations are beginning to provide tantalizing evidence for Keplerian discs around massive protostars (e.g. Hunter et al. 2014a; Johnston et al. 2015; Zapata et al. 2015; Chen et al. 2016; Ilee et al. 2016).

It is therefore imperative to develop a robust framework within which we are able to extrapolate a robust physical interpretation from such observations. Synthetic observations, where the appearance of a model to an observer is computed (ideally including the instrumentation response), are particularly valuable in this regard (for a review, see Haworth et al. 2018). To date, some synthetic observations have been computed from dynamical simulations of massive star formation/disc evolution, and are generally very optimistic about the detection of substructure in discs (e.g. Krumholz, Klein & McKee 2007; Harries et al. 2017; Meyer et al. 2018). However, these studies typically place objects at nearby distances ($\leq$1 kpc), do not always account for the full instrumental (e.g. interferometric) effects, and are typically concerned with predicting the dust continuum emission from these objects, rather than observations of molecular lines that can trace the gas emission and kinematics. In addition, approaches involving full two-dimensional (2D) or three-dimensional (3D) (radiation-)hydrodynamics are extremely computationally expensive, and thus only permit a small region of parameter space to be explored. The main caveat of existing synthetic observations of discs around MYSOs is that they are only able to focus on a small number of scenarios, owing to the cost of the dynamical simulations. Given the huge range of star-disc parameters, such as the stellar and disc mass, thermal and chemical structure, and number of spirals/clumps, a less expensive means of producing models from which to generate synthetic observations is extremely valuable.

In this paper, we combine semi-analytic models of self-gravitating discs with radiative transfer models to provide forward modelling predictions for upcoming Atacama Large Millimetre Array (ALMA) observations, in order to gauge the impact of different physical conditions on the observability of massive discs and their substructure. We note that the framework we present can be easily adapted to perform retrieval modelling of observations. This paper is organized as follows. In Section 2, we describe how we construct the semi-analytic models of self-gravitating discs, our radiative transfer calculations and how we account for the instrumentational interferometric effects. In Section 3, we present an overview of the challenges that arise in the spatial and kinematic detection of substructure in our fiducial disc model, and how we use fast and efficient filtering techniques in order to enhance this substructure. In Section 4, we explore and discuss critical effects on substructure observability and make recommendations for future observational campaigns. Finally, in Section 5, we summarize our main conclusions.

2 METHODOLOGY

We begin by describing how we construct our discs and process them to produce synthetic ALMA observations. There are three phases to this process. First, we set up the semi-analytic disc structure (as detailed in Section 2.1). Next, we compute the properties of the dust and the abundance of the molecule, and a radiative transfer calculation is used to compute the flux received by an observer (Section 2.2). Finally, we account for instrumentational effects inherent to interferometry and produce simulated observations of the models in both continuum and molecular line emissions (Section 2.3).

2.1 Disc construction

The circumstellar matter surrounding our MYSOs is constructed using a simple semi-analytic approach, following Clarke (2009), Rice & Armitage (2009), and Forgan et al. (2016). In this framework, we set up a radial profile of an axisymmetric quasi-steady self-gravitating disc. Upon this we then impose a spiral structure as a perturbation (following the procedure of Hall et al. 2016), and in some models include the presence of a fragment within the disc.

2.1.1 Underlying axisymmetric radial structure of the disc

From the inner ($r_{\mathrm{in}} = 50$ au) to the outer ($r_{\mathrm{out}} = 1000$ au) boundary of the disc, at each radius we iteratively solve a system of equations

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

that determines the disc structure. Solutions are determined by the stellar mass M_* and the accretion rate $\dot{M}$. We fix the stellar mass to $M_* = 20\,\mathrm{M}_\odot$ in this paper, but the accretion rate $\dot{M}$ is varied to produce a set of corresponding quasi-steady state disc solutions (i.e. $\dot{M}$ is constant at all radii).

Starting with an initial estimate for the surface density Σ, we calculate the angular velocity Ω at radius r

$$\Omega = \sqrt{\frac{G(M_* + \int_{r_\mathrm{in}}^{r} 2\pi r\,\Sigma\,\mathrm{d}r)}{r^3}}. \tag{2}$$

Numerical simulations predict that self-gravitating discs settle into a marginally stable state where the Toomre Q parameter (Toomre 1964) is roughly constant throughout the disc and close to ~1.7 (Durisen et al. 2007). Thus, following Rice & Armitage (2009), Forgan et al. (2016), and Hall et al. (2016), we impose a constant value of the Toomre parameter

$$Q = \frac{c_\mathrm{s}\kappa_\mathrm{epi}}{\pi G \Sigma} = 2, \tag{3}$$

where κ_epi is the epicyclic frequency and c_s is the local sound speed. For Keplerian discs $\kappa_\mathrm{epi} = \Omega$, and we assume the same here. From the last expression, we then find the sound speed c_s and from it the disc scale height H, where

$$H = \frac{c_\mathrm{s}}{\Omega}, \tag{4}$$

and mid-plane density

$$\rho_0 = \frac{\Sigma}{2H}. \tag{5}$$

Given the sound speed c_s and mid-plane density ρ_0, we use the equations of state table (Black & Bodenheimer 1975; Stamatellos et al. 2007) to find the temperature. We use the temperature T and the density ρ_0 to find the mass-mean opacity κ (Stamatellos et al. 2007), and the optical depth from the surface of the disc to the disc mid-plane, at the given radius,

$$\tau = \frac{1}{2}\kappa\Sigma. \tag{6}$$

The mass-mean opacity κ is obtained by averaging the Rosseland-mean opacity from Bell & Lin (1994) over a spherically symmetric polytropic pseudo-cloud. This takes into account effects of the surrounding warmer/colder matter.

Following Rice & Armitage (2009) and Forgan et al. (2016), we assume that angular momentum transport can be approximated as local and pseudo-viscous. The accretion rate $\dot{M}$ is then given by

$$\dot{M} = 3\pi\nu\Sigma = \frac{3\pi\alpha c_\mathrm{s}^2\Sigma}{\Omega}, \tag{7}$$

where ν is the viscosity and α is the pseudo-viscous parameter (Shakura & Sunyaev 1973). The two are related by $\nu = \alpha c_\mathrm{s} H$. In quasi-steady state $\dot{M}$ is constant at all radii.

The unknown pseudo-viscous parameter α is found by assuming that the strength of the angular momentum transport is set by the cooling rate. The cooling rate is given by (Rice & Armitage 2009)

$$\Lambda = \frac{16\sigma_\mathrm{SB}}{3}\frac{(T^4 - T_\mathrm{irr}^4)}{\tau + \tau^{-1}}, \tag{8}$$

where σ_SB is the Stefan–Boltzmann constant and T_irr is the temperature in the disc due to irradiation from the star, and we assume that gas and dust temperatures are the same. We find that using the standard Stefan–Boltzmann relationship between the stellar luminosity ($L_* \sim 7.5 \times 10^4\,\mathrm{L}_\odot$ for the stellar mass $M_* = 20\,\mathrm{M}_\odot$, Hosokawa & Omukai 2009), distance from the star, and the equilibrium irradiation temperature, T_irr, yields unrealistically high disc temperatures compared to results from hydrodynamical simulations with detailed radiative transfer (T. Douglas, T. Harries, private communication). This is possibly due to the stellar irradiation being heavily reprocessed in the outer disc. Thus, we follow Rice & Armitage (2009) and assume a constant $T_\mathrm{irr} = 10\,\mathrm{K}$ throughout the disc, bringing our results into agreement with the detailed simulations. The stellar irradiation should become much more important in the innermost disc ($r \lesssim 50$ au), where the line-of-sight optical depth from the star to the disc mid-plane is much smaller than in the outer disc. However, given the observational parameters that we probe in this work, the innermost disc is not resolved in our synthetic observations. Thus, we do not model the highly irradiated inner disc and artificially set the inner boundary of the disc to $r_\mathrm{in} = 50$ au. Finally, if the irradiation dominates over the pseudo-viscous heating ($T_\mathrm{irr} > T$), we set $T = T_\mathrm{irr}$. For our fiducial accretion rate $\dot{M} = 10^{-3}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, $T > T_\mathrm{irr}$ throughout the disc.

The cooling rate, given by equation (8), is balanced against (equal to) the viscous dissipation rate

$$D = \frac{9}{4}\nu\Sigma\Omega^2 = \frac{9}{4}\alpha c_\mathrm{s}^2\Sigma\Omega, \tag{9}$$

which we use to find the pseudo-viscous parameter α. Next, we can find the accretion rate $\dot{M}$ from equation (7) and compare it to the value for which we want to solve the structure of the steady-state disc, based on which we iteratively improve the value of the surface density Σ.

2.1.2 Spiral substructure

The radial model described above gives us an axisymmetric disc. We impose a spiral structure on that disc by adding perturbations to it. The surface density is perturbed as in Hall et al. (2016), using results of Cossins, Lodato & Clarke (2009)

$$\Sigma' = \Sigma - \alpha^{1/2}\Sigma\cos(m(\Theta(r) - \phi + \theta)), \tag{10}$$

where $\phi = \tan(y/x)$, m is the number of spiral arms, and $\Theta(r)$ determines how tightly the spirals are wrapped. We set $\Theta(r) = \frac{1}{b}\log(\frac{r}{a})$, $a = 13.5$ and $b = 0.38$, the same as Hall et al. (2016). The angle θ is the phase offset of the spiral, which we introduce in order to investigate how the angular orientation of the spirals with respect to the observer, at the time of the observation, might influence the ability to resolve the spiral arms.

In this work, we also vary the number of spiral arms m. Numerical simulations show that lower-mass ($M_\mathrm{d}/M_* < 0.1$) self-gravitating discs develop a large number of spiral arms ($m > 10$; Cossins et al. 2009), while higher-mass discs ($M_\mathrm{d}/M_* > 0.25$) develop fewer ($m \sim 2$; Lodato & Rice 2004, 2005; Forgan et al. 2011). For our fiducial disc accretion rate $\dot{M} = 10^{-3}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ the semi-analytic one-dimensional (1D) disc model yields a disc-to-star mass ratio of ~0.3, placing it in the regime of massive discs with few spiral arms. We thus set $m = 2$ in our fiducial disc model, and also run a model with $m = 4$ spiral arms to probe how the spatial separation between the spirals affects the possibility of their detection.

Finally, if α obtained from the 1D model is above a certain value $\alpha_\mathrm{sat} = 0.1$, at which gravitational torque saturates (Gammie 2001; Rice, Lodato & Armitage 2005; Deng, Mayer & Meru 2017), here we set it to $\alpha = \alpha_\mathrm{sat}$.

Here, we are also interested in kinematic effects, so we also perturb the angular velocity of matter in the disc. If we assume that the radial velocity of matter is negligible compared to the angular

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

component, we can use the angular projection of the continuity equation to show that the velocity perturbation is proportional to the density perturbation. We neglect the radial velocity $\dot{M}/(2\pi r\Sigma)$ in our models, although we note that it can be up to a few tenths of the Keplerian velocity in the outermost disc. The sign of the angular velocity perturbation then depends on the velocity of the spiral pattern compared to the local Keplerian velocity – it is positive if the spiral pattern is rotating faster than the unperturbed velocity, and negative in the opposite case. Therefore, we perturb the angular velocity via

$$\Omega' = \Omega - \eta\Omega\cos(m(\Theta(r) - \phi + \theta)), \tag{11}$$

where η is a free parameter that controls the magnitude of the perturbation. The η parameter is determined by the rotation velocity of the spiral pattern that is roughly proportional to the disc-to-star mass ratio (Cossins et al. 2009; Forgan et al. 2011). We set $\eta = 0.2$ in our fiducial disc model, appropriate given the disc-to-star mass ratio in the model, but also vary η to probe the observational effect of weaker perturbations.

Furthermore, we set the perturbed sound speed to the value of the local fluid velocity relative to the unperturbed speed $c_\mathrm{s}' = |\Omega' - \Omega| r$ (with a lower limit of the unperturbed sound speed, c_s). This accounts for the higher temperatures expected in the spirals due to heating from shocks (Cossins et al. 2009).

An illustration of our semi-analytic discs is given in Fig. 1, which compares the surface density, sound speed and local bulk speed of axisymmetric, two-arm spiral and four-arm spiral discs.

2.1.3 *Prescription for fragments*

In addition to discs with spiral structure, we also explore the observational characteristics of discs undergoing fragmentation. To that end, we insert clumps within the disc to represent the presence of a fragment within our disc models. The clump is modelled as an isothermal Gaussian within a sphere of radius R_C

$$\rho = \rho_\mathrm{C}\exp\left(-\frac{|\boldsymbol{r} - \boldsymbol{r}_\mathrm{C}|^2}{2R_\mathrm{C}^2}\right), \tag{12}$$

where $\vec{r}$ is the vector relative to the clump centre. We assign the clump temperature T_C, central density ρ_C, and radius R_C guided by the analysis of fragments in smoothed particle hydrodynamic simulations by Hall et al. (2017). The centre of the clump $\vec{r_\mathrm{C}}$ in our models is always set so that the clump is located within a spiral arm (as is often the case in simulations of fragmenting discs, see e.g. Boley 2009).

2.2 Radiative transfer modelling

We produce synthetic images and molecular line data cubes from the above disc models using the TORUS radiative transfer and hydrodynamic code (e.g. Harries 2000; Kurosawa et al. 2004; Haworth & Harries 2012), with the molecular-line transfer ray-tracing scheme presented by Rundle et al. (2010). We now summarize the relevant details of these calculations.

2.2.1 *Constructing a 3D disc*

We map the 2D models, described in Section 2.1, on to the midplane of a TORUS grid using bilinear interpolation. We then construct the vertical structure of the disc assuming hydrostatic equilibrium. We give ourselves the freedom to explore the effect of a disc that is not isothermal using a parametric temperature profile of the form:

$$T(R,z) = \begin{cases} T_\mathrm{mid} + (T_\mathrm{atm} - T_\mathrm{mid})\left[\sin\left(\frac{\pi z}{2z_\mathrm{q}}\right)\right] & z < z_\mathrm{q} \\ T_\mathrm{atm} & z \geq z_\mathrm{q} \end{cases} \tag{13}$$

following Williams & Best (2014), where we set $T_\mathrm{atm} = f_\mathrm{atm}T_\mathrm{mid}$ and z_q is four times the disc scale height. This allows us to compare an isothermal disc ($f_\mathrm{atm} = 1$) with increasing amounts of vertical heating ($f_\mathrm{atm} > 1$) that could have important observational impact (for example sublimating dust, or destroying or exciting molecules).

2.2.2 *Dust continuum emission*

Dust is included in our radiative transfer models, contributing as a continuum emission source and an opacity source to the line emission. We assume a single distribution of grains at all radii that is well mixed with the gas with a dust-to-gas mass ratio of 10^{-2}, though we also follow Vaidya, Fendt & Beuther (2009) and sublimate dust in regions where the temperature is above 1500 K. The dust distribution is a Mathis, Rumpl & Nordsieck (1977) power law ($n(a) \propto a^{-q}$) specified in terms of the minimum and maximum size and a power law for the grain size distribution. For most of the models in this paper, we use $a_\mathrm{min} = 1$ nm, $a_\mathrm{max} = 1$ cm and $q = 3.3$, though we probe the impact of grain growth in Section 4.4.

2.2.3 *Gas emission and molecular abundances*

We have chosen to focus our observational predictions for the gas on the CH_3CN $J = 13$–12 ladder of transitions. Emission from CH_3CN is widely detected in the vicinity of massive protostars (see e.g. Hunter et al. 2014b; Ilee et al. 2016; Beuther et al. 2017; Cesaroni et al. 2017). The ladder is an incredibly useful diagnostic of the physical conditions of the gas, due to the large number of K transitions that span a range of excitation energies, and the fact that they are sufficiently closely spaced to enable simultaneous observation. Within the $J = 13$–12 ladder, transitions range from the $K = 0$ transition at 239.138 GHz ($E_\mathrm{up} = 80$ K) to the $K = 12$ transition at 238.478 GHz ($E_\mathrm{up} = 1106$ K). We have chosen to concentrate on the $K = 3$ transition (239.096 GHz, $E_\mathrm{up} = 145$ K) as this line traces the entire radial extent of the disc in our models.

We assume local thermodynamic equilibrium (LTE), which significantly reduces the computational cost of the models compared to non-LTE line transfer. The CH_3CN abundance is canonically assumed to be 10^{-7} relative to molecular hydrogen (Herbst & van Dishoeck 2009), but we assume that the molecular abundance can be depleted by freeze out at low temperatures or destroyed through reactions. Assuming that freeze out and thermal desorption are in equilibrium, the threshold density at a given temperature T above which CH_3CN will be frozen out is

$$\rho_\mathrm{thr} = \frac{m_\mathrm{d}}{\delta\pi\sigma_\mathrm{d}}\sqrt{\frac{N_\mathrm{s}E_\mathrm{b}}{k_\mathrm{B}T}}\exp\left(-\frac{E_\mathrm{b}}{k_\mathrm{B}T}\right), \tag{14}$$

where δ, m_d, σ_d, N_s, and E_b are the dust-to-gas mass ratio, mean dust grain mass, dust cross section, the number density of surface binding sites ($\approx 1.5 \times 10^{15}$ cm^{-2}), and the species-dependent binding energy $E_\mathrm{b}/k_\mathrm{B} = 4680$ K (Collings et al. 2004). We assume a mean grain size of 0.1 µm, and grain density of 3.5 g cm^{-3} (from which the cross-section and grain mass can be computed assuming spherical grains).

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

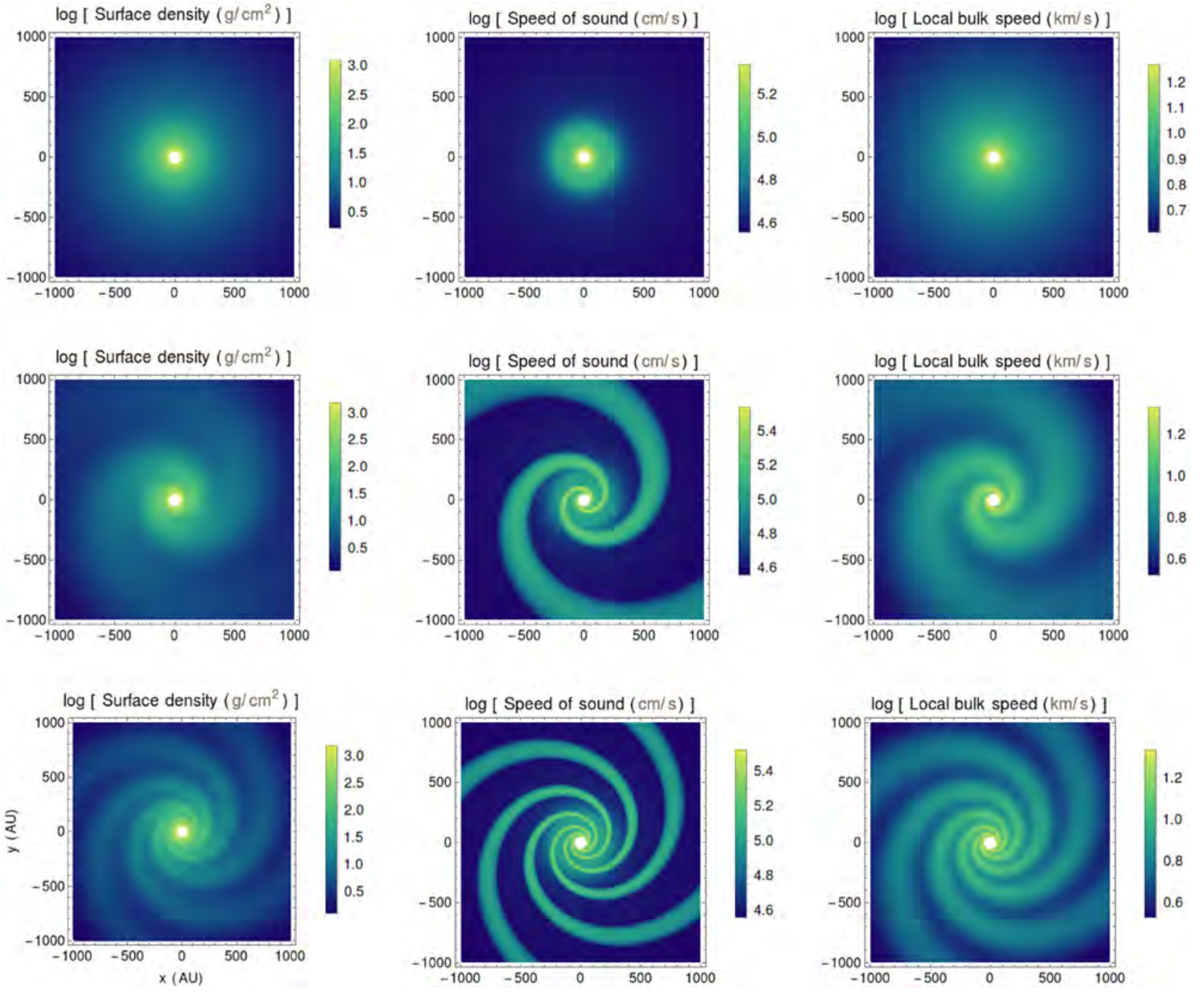


Figure 1. Self-gravitating discs constructed using the 2D semi-analytic model: axisymmetric (top row), two spiral arms (middle row), four spiral arms (bottom row). The panels show the surface density, sound speed, and bulk velocity from left to right.

We also assume a dust-to-gas mass ratio of 10^{-2}. Where CH_3CN is frozen out, we assume a negligible abundance of 10^{-20}. We do not account for destruction of CH_3CN by two-body reactions/cosmic rays (despite the time-scale for destruction by these processes being short) because doing so requires a sophisticated chemical network including CH_3CN production pathways. We do however destroy CH_3CN in regions where the temperature is above 2300 K, which is the regime in which the chemical models of Walsh et al. (2014); Walsh, Nomura & van Dishoeck (2015) find its abundance sharply drops and becomes negligible.

With the abundances and level populations known, a position–position–velocity data cube is constructed using a ray tracing scheme through the disc (Rundle et al. 2010). Both the continuum and line emission are calculated in the frequency range $\nu \pm \nu u_{\max}/c$, where ν is the rest frequency of the molecular line considered in a given synthetic observation, $u_{\max} = 40\,\mathrm{km\,s^{-1}}$ (which probes right down to our inner radius of 50 au for the $20\,\mathrm{M}_\odot$ star we consider) and c is the speed of light.

2.3 Accounting for the instrumentational response of ALMA using CASA

Given that we are interested in interpreting observations from ALMA, it is essential that we account for the instrumentational response. For example, limited time on source, finite resolution and the lack of sensitivity to large-scale structure inherent to high-resolution interferometry can all have significant effects on the resulting data. We therefore postprocess the TORUS molecular line data cubes described in Section 2.2 using the CASA[1] software (McMullin et al. 2007) to account for the above.

In our fiducial model, we consider band 6 observations in ALMA configuration 40.7, which has a minimum baseline of 81 m, a maximum baseline of 3.7 km, an angular resolution of 0.09 arcsec and a maximum recoverable scale of 0.8 arcsec. Given that a 1000 au disc at ~1 kpc has an angular size of ~1 arcsec, and given ALMA

---

[1] https://casa.nrao.edu/

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

capabilities in cycle 4, the cycle at which ALMA reached close to full operations for the 12-m array, this configuration was a pragmatic choice for attempting to better detect and resolve discs around MYSOs and perhaps also to search for substructure within them. We use a spectral resolution of 0.4 km s^{-1}, representative of the practical spectral resolution achieved in cycle 4 observations of candidate high-mass star–disc systems (Ilee et al., in prep.); however higher spectral resolution is possible with ALMA.

To account for the instrument response of ALMA, we use the CASA SIMOBSERVE and CLEAN routines. We set the zenith precipitable water vapour to 1.796 mm (an estimate from the ALMA sensitivity calculator for the fifth octile). The total time on source is set to 2 h (with an integration time of 20 s), with the source transiting in the middle of the observation. We obtain an image of the continuum emission from the 40 line-free channels at either end of the cube. Thus, the continuum bandwidth used here is $\sim$0.025 GHz. Even in line-rich sources such as NGC6334I, line-free continuum bandwidths $\gtrsim$10× larger are readily obtainable with ALMA (see e.g. table 1 of Hunter et al. 2017), so our results represent a conservative lower limit for continuum signal-to-noise ratios. The line-free channels are used to subtract the continuum in the *uv* plane, using the UVCONTSUB routine. To produce the continuum and line images, we use the CASA CLEAN routine with a threshold of 3× the RMS noise (RMS $\approx$ 2.5 mJy beam^{-1} per channel for the $K = 3$ line, and $\approx$0.3 mJy beam^{-1} for the continuum), and using a circular mask around the source. Integrated intensity (moment 0) and intensity-weighted velocity (moment 1) maps are made using the IMMOMENTS routine, and the position–velocity (PV) diagrams are made using the IMPV routine. In making moment 1 maps, only pixels above a cut-off of 5× the RMS are included.

We adopt a distance of 3 kpc for the majority of our models in this paper, but also explore 1–5 kpc. We are hence using substantially larger distances than previous synthetic observations of discs about MYSOs, which typically consider 0.5–1 kpc and/or note that at larger distances detecting discs and structure is difficult (e.g. the 2 kpc distant models of Krumholz et al. 2007). In all observations considered here, we assume our object lies in the direction of the centre of the protocluster G11.92−0.61, with on sky co-ordinates of RA = $18^{\mathrm{h}}13^{\mathrm{m}}58.1$, Dec. = $-18^{\mathrm{d}}54^{\mathrm{m}}16\overset{\mathrm{s}}{.}7$ (Cyganowski et al. 2017).

3 RESULTS

We use our framework for quickly producing synthetic observations of self-gravitating discs to explore whether or not self-gravitational substructure, spiral density waves, and fragmentation are detectable in discs around massive stars with upcoming ALMA observations. We find that the spiral features are not always obvious neither spatially, in continuum images and integrated intensity (moment 0) maps of line emission, nor spectrally, in intensity-weighted velocity (moment 1) maps and PV diagrams. This is in contrast, but not in conflict, with the findings of Douglas et al. (2013), who modelled signatures of self-gravitating discs in *nearby, lower mass* systems (their model is of a 1 $\mathrm{M}_{\odot}$ star and a 0.39 $\mathrm{M}_{\odot}$ disc) in lines from more abundant species such as CO and HCO^{+}. Significant enhancement of substructure can, however, be obtained, using filtering techniques presented in this section.

Our fiducial model is a disc around a $M_{*} = 20\,\mathrm{M}_{\odot}$ star at a distance of 3 kpc with an inclination of 30 deg, an accretion rate of $\dot{M} = 10^{-3}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, a ratio of atmospheric to mid-plane temperature $f_{\mathrm{atm}} = 4$, an intensity of velocity and temperature perturbations $\eta = 0.2$, and two spiral arms (see the second row of Fig. 1). In Section 4, we will explore how the detection of self-gravitational signatures is sensitive to the disc and observational (e.g. distance, line, inclination) parameters. In this section, we focus on enhancing substructure using filtering techniques.

3.1 Enhancing spatial detection of substructure in discs

As mentioned above, spirals are not always obvious in our synthetic continuum images and moment 0 maps. To improve the spatial detectability of spirals we subtract the continuous background disc emission to highlight the more spatially confined substructure. We define the base synthetic image (i.e. that from TORUS and CASA) as the ‘primary image’. We consider subtraction of 2D (elliptical) Gaussian fits to this primary image using the CASA IMFIT routine. The IMFIT routine is a well-tested, easy-to-use component of the publicly available CASA software.

Fig. 2 compares the primary image with the residuals of fitting a single Gaussian and of simultaneously fitting two Gaussian components to the 239 GHz continuum image of our fiducial disc model with two spiral arms. Fig. 3 presents the same comparison for moment 0 maps of the line emission. We find that in the moment 0 maps the two-component Gaussian fit is better at filtering the more continuous disc emission, as one component fits the bright inner disc (that is optically thick, see Section 4.4) and the other covers a larger extent of the disc. On the other hand, subtraction of a single Gaussian fit from the moment 0 map of our fiducial disc model (the middle panel of Fig. 3) is inadequate, as it does not fit the continuous/axisymmetric emission from the disc well. It does not subtract the unresolved bright inner disc, but subtracts the inner regions of the spiral arms that are clearly revealed in the image filtered using the two-component Gaussian fit. On the other hand, we find that in the continuum image it is sufficient to fit and subtract a single elliptical Gaussian profile to clearly highlight the spiral substructure. Note that the continuum image filtered using the two-component Gaussian fit (the right-hand panel of Fig. 2) features two compact bright spots near the centre of the disc that are not real structures. The artefacts also appear in the image filtered using a single Gaussian profile, but they are much less significant.

Throughout the rest of this paper, we apply the technique of subtracting the one-component and the two-component Gaussian fits when considering continuum and moment 0 line emission maps, respectively. We found this technique to have a good balance of ease of implementation and ability to detect substructure.

We note two points of caution when using these techniques. First, our results show that even in discs that do not possess real clump-like features (such as our models with completely smooth spirals), the interferometric imaging can induce artefacts that could be mistaken for clumps and/or substructure in the disc. We therefore suggest that caution be taken when attributing such features to the presence of a real feature, such as disc fragmentation due to gravitational instability, particularly when the feature is of low significance and close in spatial scale to the interferometric beam. Secondly, the CASA IMFIT routine requires an initial estimate of the parameters for multiple component Gaussian fits as input, and the results can be sensitive to this input. We find it useful to run the fitting routine several times with varying input estimates to check for convergence.

Finally, we note that fitting more tailored parametric axisymmetric disc models of the intensity profile (e.g. those presented here) could do an even better job at revealing substructure, particularly since the brightness profile will deviate from Gaussian in practice. However, this requires more detailed modelling.

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

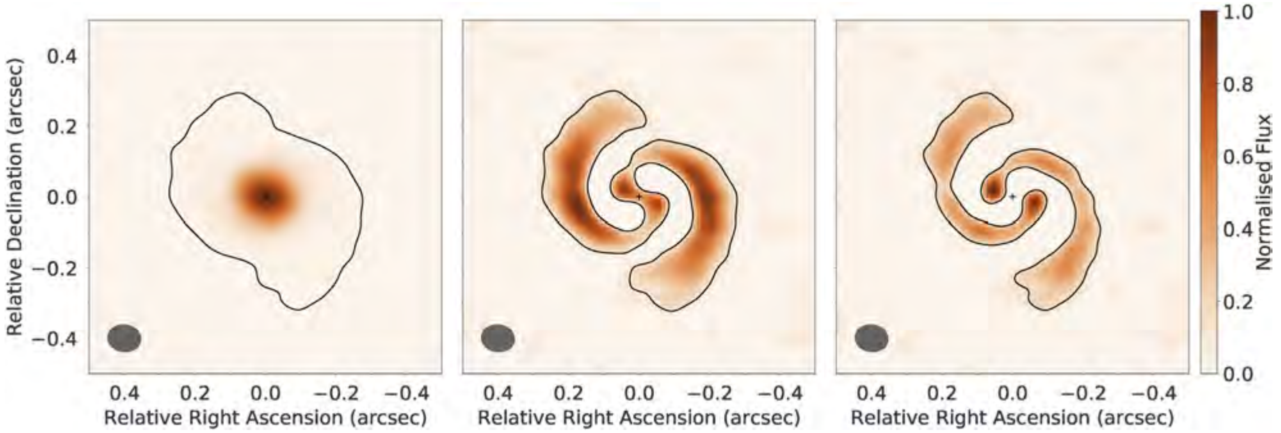


Figure 2. Synthetic continuum images for a two arm spiral disc model with two filtering techniques applied, with 3σ (σ = off-source RMS noise) contours. Flux is normalized with respect to the peak value in each image. Left is the primary (unfiltered) image of the 239 GHz continuum, with a signal-to-noise ratio of 500. Middle subtracts a single Gaussian fit to the primary image. Right subtracts a two-component Gaussian fit. The signal-to-noise ratio in the filtered images is 25 and 20, for the single and double Gaussian fits respectively.

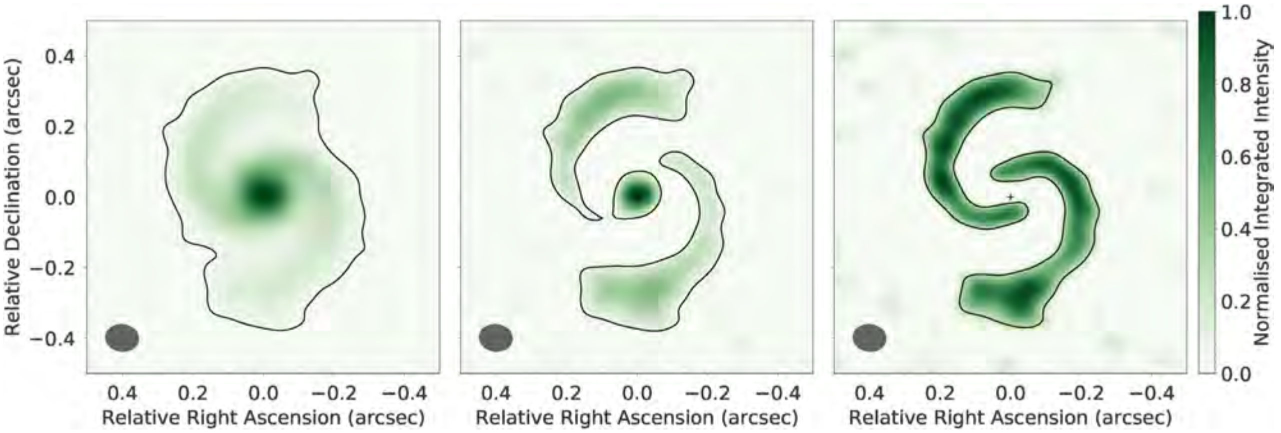


Figure 3. Synthetic moment 0 maps for a two arm spiral disc model with two filtering techniques applied, with 3σ (σ = off-source RMS noise) contours. Integrated intensity is normalized with respect to the peak value in each map. Left is the primary (unfiltered) image, with a signal-to-noise ratio of 82. Middle subtracts a single Gaussian fit to the primary image. Right subtracts a two-component Gaussian fit. Signal-to-noise ratio in the filtered images is 33 and 11, for the single and double Gaussian fits respectively.

3.2 Enhancing spectral detection of substructure in discs

To improve the kinematic detectability of spirals we subtract line-of-sight projected Keplerian velocities, convolved with the synthesized beam of our synthetic observations from our model moment 1 maps. This method was used by Walsh et al. (2017) to constrain the inclination and the position angle of the HD 100546 disc, as well as to infer the presence of either a warp or radial flow residing <100 au from the central star (distinguishing between the two is difficult, e.g. Rosenfeld, Chiang & Andrews 2014).

By using the same fitting metrics as Walsh et al. (2017) we find that the best-fitting inclination and position angle do not always correspond to the known (i.e. input) inclination and position angle of our disc models, with typical variation in both of ±10 deg. We attribute this to deviations from a simple flat thin disc model. Irrespective of this, we find that an incorrect best-fitting inclination (that is within ±5–15 deg from the correct value) does neither affect the ability to detect the spiral signatures in our synthetic observations nor induce the appearance of any artificial substructure. We thus proceed with subtraction of Keplerian profiles corresponding to the actual disc model inclinations and position angles, but noting that the automatically fitted value would also suffice.

The moment 1 map of the synthetic line emission observation, the moment 1 map of equally inclined flat Keplerian emission, and the result of the subtraction of the latter from the former map (the moment 1 map residuals) are shown in Fig. 4 for our fiducial disc model with two spiral arms. The velocity residuals are given in units of the spectral resolution $\delta v = 0.4\,\mathrm{km\,s^{-1}}$. The warp-like features resulting from spiral structure are much more prominent following the kinematic filtering technique we apply.

Throughout the rest of this paper, we apply the described subtraction technique to enhance the kinematic detectability of disc substructure in moment 1 maps and present only the residuals of the subtraction.

4 DISCUSSION

The speed of our disc model construction and synthetic observations (under an hour per model using four cores for the radiative transfer calculations, and a single core for the rest, on a desktop) allows us to

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

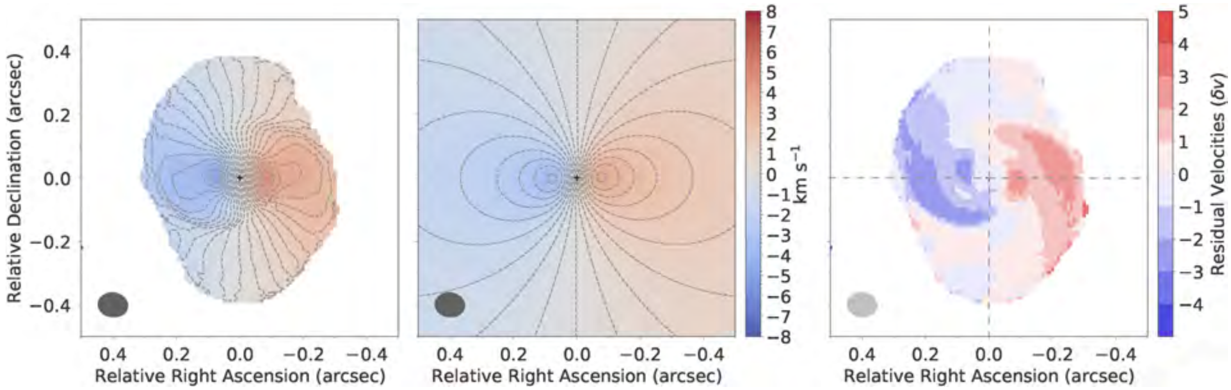


Figure 4. An illustration of kinematic filtering applied to a moment 1 map of a two spiral disc. The left-hand panel is the initial synthetic observation, the middle a Keplerian profile, and the right-hand panel the residual. The velocity residuals are given in units of the spectral resolution $\delta v = 0.4\,\mathrm{km\,s^{-1}}$. Dashed contours in the left-hand and the middle panels also show the velocity in units of δv. In the right-hand panel, dashed horizontal and vertical lines indicate disc position angle and the line perpendicular to it, respectively. The warp-like features resulting from spirals are much more evident in the residuals map.

explore a parameter space of disc (e.g. mass and temperature) and observing (e.g. distance and line) parameters. We detail the specific parameter space that we explore in Table 1. Starting from our fiducial disc model, we vary the parameters one-by-one, producing sets of models that are different from each other in the value of a single parameter. This is with the exception of disc inclination that is also separately set for each set of models. All of the models are available for download.[2] In this section, we discuss what the limiting factors are for the detection of self-gravitational substructure both spatially and spectrally, and gauge the expectations for upcoming ALMA observations of MYSOs. We focus first on self-gravitating disc models with spiral density waves, for which we also discuss the impact of disc thermal and chemical structure (Section 4.3) and optical depth (Section 4.4). We then explore detectability of disc fragmentation into clumps as a function of the fragment properties (Section 4.5).

4.1 Spatially resolving spiral substructure

Spatial detection of spiral substructure is sensitive to the angular resolution relative to the angular separation between spirals, as viewed by the observer. Manifestly, the spatial detectability of spirals depends on the disc distance, as well as on the disc inclination and the nature of spirals, as illustrated in Fig. 5. This section focuses on the line moment 0 maps from our synthetic observations, but in most cases, the same conclusions can be drawn from analysis of the continuum images.

Spiral substructure in strongly self-gravitating discs, i.e. those with a lower number of spiral arms and less tightly wound spirals, is likely to be detected. For 1000-au-radius discs with an inclination of 30 deg, we find that two-arm spirals are resolvable up to approximately 5 kpc distances with ALMA configuration 40.7. Such a distance encompasses approximately 70 per cent of the MYSO population in the Red MSX Source (RMS) Survey[3] for which distances are available (Lumsden et al. 2013). At close distances (<1 kpc) the moment 0 maps of line emission might even yield hints of vertical stratification in the outer disc (the top left-hand panel of Fig. 5), since the beam size is smaller than the disc scale height inside the spiral arms. Indeed, the vertical stratification of molecular emission in more evolved, axisymmetric discs has already been observed (e.g. HD 163296, Rosenfeld et al. 2013; de Gregorio-Monsalvo et al. 2013), but our results suggest that similar structures will be observable in non-axisymmetric discs.

As the number of spiral arms increases, however, and spirals become more tightly wrapped, it becomes increasingly more difficult to resolve the spirals. We find that four-arm spirals are barely discernible at a moderate distance of 3 kpc. This principally happens because the size of the synthesized beam is large compared to the angular separation between spiral arms.

Spiral substructure is more easily resolved spatially for more face-on discs, but spirals in our models are detected for the majority of disc inclinations, from 0 (face-on discs) up to ∼50 deg. Furthermore, over the inclinations that the spirals are not spatially detected, below we will show that they are likely to be detected spectrally.

4.2 Spectrally resolving spiral substructure

The usefulness of spectral diagnostics in probing the kinematic perturbation caused by the self-gravity-induced spiral substructure depends on how the spectral resolution of the observation relates to the line-of-sight-projected velocities, and is thus sensitive to disc inclination: an edge-on disc will be easier to resolve spectrally than a face-on disc (if otherwise identical). We thus expect that deviations from Keplerian motion will be much easier to detect in PV diagrams for more edge-on discs. Indeed, this is seen in the bottom panels of Fig. 6, which show PV diagrams of our fiducial disc model inclined at 30, 60, and 80 deg. In each case, the red line denotes the Keplerian profile. The signature of spirals is a twin-lobe feature that cuts diagonally across the profile, which if sufficiently resolved should appear as a figure-of-eight like structure, similar to that suggested by Douglas et al. (2013). The signature becomes more prominent with increasing inclination.

In addition to the spiral signatures, there are other deviations from the Keplerian profile in the PV diagrams, i.e. deviations from the curved red lines. These deviations are seen in the PV diagrams of both spiral (Fig. 6) and axisymmetric disc models (Fig. 7). For example, in discs inclined at 30 and 60 deg, in the lower left and middle panels in Fig. 6 (and Fig. 7), there is emission at low velocities from the inner disc. This is simply due to convolution of the emission with the beam in each velocity channel, that is, an observational effect due to the disc not being perfectly resolved.

---

[2] DOI 10.5281/zenodo.1408072.

[3] http://rms.leeds.ac.uk

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Table 1. A summary of the models with spiral arms presented in this paper, except for the models with clumps. The first column lists the model parameters that were varied, second the fiducial (default) value of each parameter, third the values that were explored for each parameter, and fourth the disc inclination used for the set of models in each row. In the first column η is the magnitude of spiral velocity and temperature perturbation, and f_{atm} the ratio of atmospheric to mid-plane temperature.

Parameter	Fiducial	Values explored	Inclination (deg)
Line choice	$K = 3, J = 13 \rightarrow 12$	$K = 7, J = 13 \rightarrow 12$	30
Inclination (deg)	30, 60	0−90	
Number of spirals	2	0, 4	30
Accretion rate ($\mathrm{M_{\odot}\ yr^{-1}}$)	10^{-3}	10^{-5}, 10^{-4}, 10^{-2}	60
η	0.2	−0.2, 0.1	60
Rotational offset (rad)	0	$\pi/4$, $\pi/2$, $3\pi/4$	60
Distance (kpc)	3	1, 5	30
f_{atm}	4	1, 2, 8	60
Dust	$a_{\max} = 1$ cm, $q = 3.3$	$a_{\max} = 0.1\ \mu$m −1 cm; $q = 3.3$–3.5	60

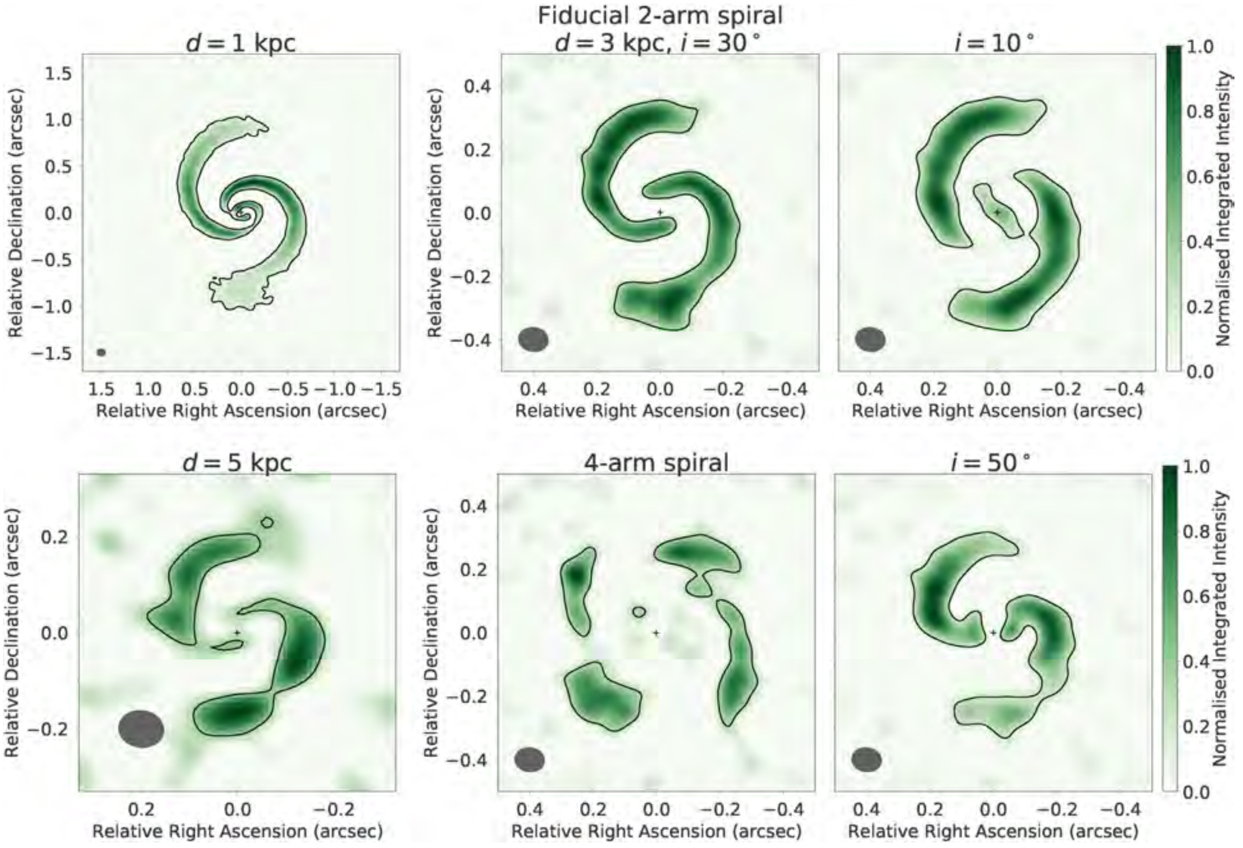


Figure 5. A summary of the parameters to which the spatial detection of spirals is sensitive. All of the panels show line moment 0 maps from synthetic observations, with 3σ (σ = off-source RMS noise) contours, and in which two-component Gaussian fits are subtracted to enhance substructure (see Section 3.1). For the six moment 0 maps shown here, the signal-to-noise ratio in the unfiltered images varies roughly between 55 and 110, and in the filtered images between 8 and 25; the lower boundaries of these ranges correspond to the disc model at a distance of $d = 5$ kpc, and the upper boundaries to the disc model at $d = 1$ kpc. Integrated intensity is normalized with respect to the peak value in each map. The top central panel is our fiducial two-arm spiral model. The lower central panel has four spirals, but is otherwise identical to the two-arm fiducial. The left-hand panels illustrate the sensitivity to distance, and the right-hand panels the sensitivity to inclination.

Noticeably, there is no such emission in the disc inclined at 80 deg (the bottom right-hand panels). This, on the other hand, is an optical depth effect. PV diagrams are produced based on cuts along the major axis of the disc image (i.e. based on the disc position angle; we use the input value from our models). In a fairly face-on disc, the cut is made across the ‘top’ disc surface, and in a sufficiently inclined disc the cut is along the outermost flared side of the disc. In the disc inclined at 80 deg the line of sight towards the inner disc cuts through a large column of gas with near-zero line-of-sight velocities. Consequently, the near-zero line-of-sight velocity emission from the inner disc region is obscured by the line self-absorption. We discuss the more general effects of optical depth in Section 4.4.

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

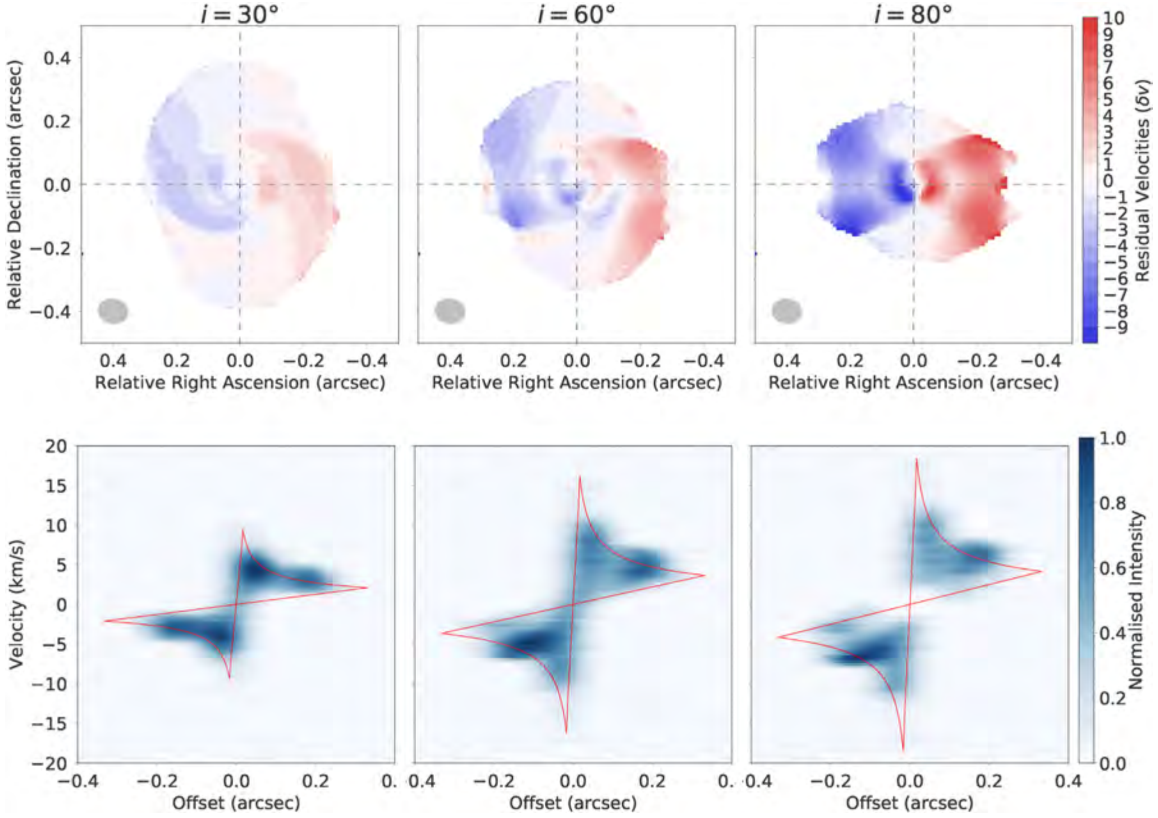


Figure 6. Sensitivity of spectral detection of spirals to disc inclination. The top panels are moment 1 map residuals of the kinematic filtering (given in the units of the spectral resolution $\delta v = 0.4\,\mathrm{km\,s^{-1}}$), and the bottom panels are PV diagrams of our fiducial disc model. Intensity is normalized with respect to the peak value in each diagram. Red lines indicate the expected position–velocity signature of a smooth, Keplerian disc with the same parameters as the model. For both top and bottom panels: left to right disc inclinations are 30, 60, and 80 deg.

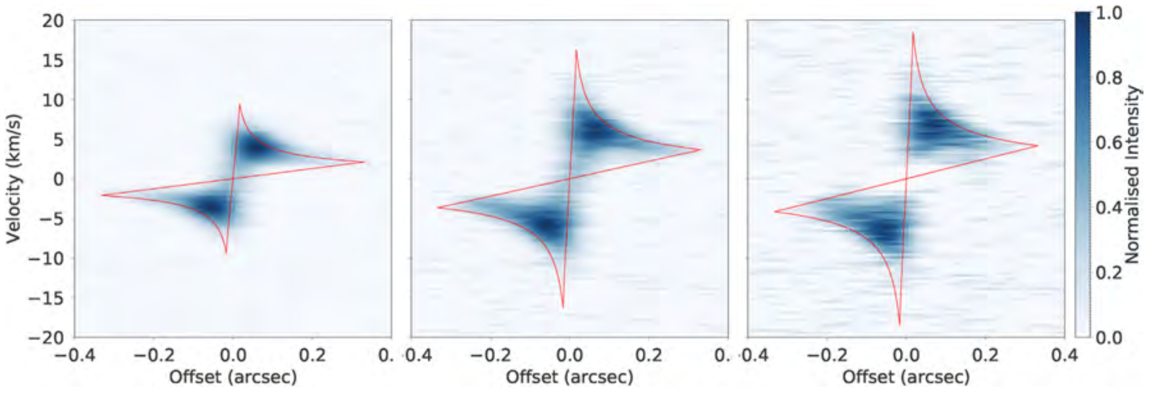


Figure 7. PV diagrams of an axisymmetric disc model, identical to our fiducial disc model except for the lack of spirals. Intensity is normalized with respect to the peak value in each diagram. Left to right disc inclinations are 30, 60, and 80 deg. Red lines indicate the expected position-velocity signature of a smooth, Keplerian disc with the same parameters as the model. Deviations from this Keplerian profile are discussed in Section 4.2.

Furthermore, there is also missing emission at Keplerian velocities from the outermost disc in the PV diagram of the disc inclined at 80 deg, unlike in the less inclined discs. The emission is missing because the PV cut probes the cold disc mid-plane of the outer disc in this case, where the molecule is frozen out. For a discussion of the observational effects of the molecule freeze out, see Section 4.3.

Fig. 6 also illustrates how the dependence on inclination is twofold for the moment 1 maps, as they are also sensitive to the spatial resolution (i.e. beam size). As discussed above, spirals are easier to detect spatially in more face-on discs. Interestingly, the moment 1 maps are also affected by the molecule freeze out. The disc inclined at 30 deg (the top left-hand panel of Fig. 6) has an apparent aspect ratio inconsistent with its inclination (and rotation profile). This is due to the molecule freeze out in the outer disc outside of spiral arms. Furthermore, the moment 1 map of the disc inclined at 80 deg (the top right-hand panel in Fig. 6) also reveals the absence of the molecule from the disc mid-plane: the radial extent of emission is smaller along the horizontal dashed line (that cuts the disc mid-plane in the outermost disc) than below and above this line (where emission from the hot disc atmosphere is seen).

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

We conclude from our results that around MYSOs at ∼3 kpc spirals will be readily resolved spatially in low-inclined discs ($\lesssim$50 deg), and readily detected spectrally in PV diagrams in more inclined discs ($\gtrsim$60 deg). For discs at intermediate inclinations (∼50–60 deg), spatial diagnostics aided by moment 1 maps and PV diagrams should unambiguously show spiral substructure. Similar dependencies in detecting substructure on inclination have been found by Douglas et al. (2013), albeit for lower mass discs that are much closer than a few kpc.

We note that MYSOs may be embedded in a natal envelope that may affect observations both by obscuration of emission, and via influencing the kinematic signature of lines (i.e. infall, see Ilee et al. 2016); we do not account for this effect in our models. Kinematic signatures could also be affected by larger-scale filamentary flows in the immediate star-forming environment (see e.g. Maud et al. 2017; Izquierdo et al. 2018).

We find that both spatial and spectral detectability of spiral substructure are only weakly dependent on the rotational offset (rotation angle relative to a fixed disc axis) of spirals in our models (see Section 2.1.2).

The kinematic perturbation introduced by the spirals in the above models is super-Keplerian, i.e. the spiral pattern is rotating faster than the disc and the matter in spiral overdensities is moving at a velocity that is higher than the local Keplerian velocity. The magnitude of the kinematic perturbation (parameter η in our models) is expected to roughly scale with disc-star mass ratio, i.e. with how self-gravitating the disc is (Cossins et al. 2009; Forgan et al. 2011). The above results correspond to a strongly self-gravitating disc ($\eta = 0.2$), and we find that it quickly becomes difficult to resolve the perturbation in less self-gravitating discs ($\eta = 0.1$). Obviously, this is a function of the spectral resolution of the (synthetic) observation, but the disc physics also plays a role: kinematic perturbation directly determines thermal perturbation in the spirals. The heating due to shocks in spiral arms is expected to be proportional to the kinematic perturbation (Cossins et al. 2009), as accounted for in our models.

We also explore the possibility of sub-Keplerian kinematic perturbation and find that the twin-lobe feature in the PV diagrams appears similarly as in the super-Keplerian case, but the moment 1 maps become much more difficult to interpret, and spiral signatures are not easily identified. Nevertheless, the pattern speed of spiral density waves in low-mass self-gravitating discs is found to be generally super-Keplerian in simulations (Forgan, Ramón-Fox & Bonnell 2018a). Therefore, spirals should be easy to spectrally resolve in sufficiently inclined and strongly self-gravitating discs.

4.3 Sensitivity to thermal and chemical structure

Molecular abundances can vary greatly within a circumstellar disc, with different molecular species surviving in the gas phase within different regions in the disc. For example, the CH_3CN in our fiducial disc model resides in the hot inner disc, the stellar-irradiated warm disc atmosphere and the shock-heated regions of the spiral arms (for the details of the chemical prescription, see Section 2.2.3). In the disc mid-plane, the molecule is mostly frozen out outside the spirals and outside the snowline at ∼300 au.

The molecular abundance in the atmosphere is regulated by the atmospheric heating, a free parameter in our models. In a disc colder than our fiducial one the molecule is frozen out higher into the atmosphere. For sufficiently weaker atmospheric heating, the molecule in the outer disc will practically only trace the spirals, as emission from the low-density upper layers will be too weak. Thus, it could be expected for a cooler disc to make spirals easier to detect as they would have a higher contrast with the rest of the disc. However, we find that varying the atmospheric heating in our fiducial disc model does not affect detection of spirals, since there is already a significant contrast between the shock-heated spiral arms and the rest of the disc. The effect might become important in less massive discs with weaker spiral perturbations, as discussed above.

4.4 The optical depth

The TORUS radiative transfer calculations give us quick access to the optical depth in both gas and dust in the models. For reference, the dust absorption coefficient at frequency ν in a medium of density ρ is

$$\alpha_\nu^{\mathrm{dust}} = \kappa_\nu \rho \tag{15}$$

for opacity κ_ν. Additionally, the absorption coefficient in the line is set by the balance between stimulated photoabsorption and induced emission

$$\alpha_\nu^{\mathrm{gas}} = \frac{h\nu}{4\pi}\left(n_l B_{lu} - n_u B_{ul}\right)\phi_\nu, \tag{16}$$

where n_u, n_l are the molecule populations in the upper and lower states of the transition, ϕ_ν is the line profile function which is sensitive to both the bulk and microturbulent velocities, and B_{ul} and B_{lu} are the Einstein coefficients of stimulated emission and absorption. We computed the optical depth at the line frequency due to dust and gas (the line self-absorption) from each grid cell on the mid-plane along the trajectory towards the observer in a given model (i.e. the optical depth is inclination dependent).

We first consider the opacity due to dust only in an axisymmetric disc, i.e. a disc without spirals. In this section we allow ourselves to vary the maximum grain size in the distribution to probe the impact of grain growth. We find that there is an inner zone that is optically thick due to dust alone (this extent, i.e. radius, is roughly equivalent for both the $K = 3$ and $K = 7$, $J = 13 \rightarrow 12$ lines, which only differ in frequency by ∼0.1 GHz). For a maximum grain size $a_{\mathrm{max}} < 100\ \mu\mathrm{m}$ (minimum grain size 1 nm, $q = 3.3$), this optically thick inner region due to dust alone is around 50 au in extent (or less in the case of lower accretion rates, though our inner radius is 50 au so we cannot constrain smaller optically thick extents). If a_{max} is increased to around 300–500 μm the optically thick region sharply jumps in size to around 100–250 au, and remains so for maximum grain sizes up to 1 cm. These extents are only mildly sensitive to the inclination, at least until the disc is almost edge on and the projected column becomes a lot higher.

Examples of the extent of the optically thick dust disc as a function of the maximum grain size for a series of axisymmetric discs of different mass accretion rates (and hence total masses) are shown in the upper panel of Fig. 8. The lower panel of Fig. 8 shows how the opacity varies as a function of wavelength, the behaviour of which at the transition wavelength (black line) explains the form of the upper panel. Unsurprisingly, a larger extent of the disc is optically thick for larger disc masses.

We hence expect that discs around MYSOs observed in these lines will have an optically thick inner region due to dust alone. If this is detected beyond 100 au, up to around 250 au, it may serve as evidence for widespread grain growth beyond the usual ISM maximum grain size (this is expected in discs, see e.g. Testi et al. 2014 for a review). Note that dense spirals can also induce localized

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

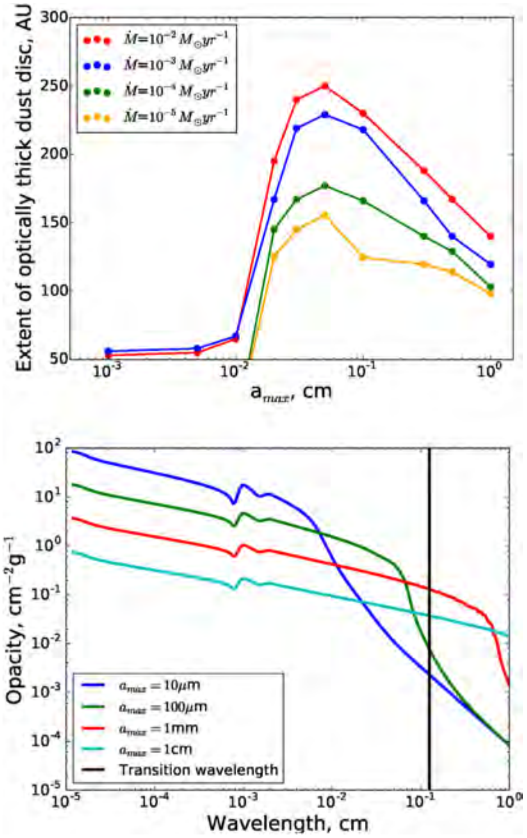


Figure 8. The upper panel shows the extent of axisymmetric regions of discs that are optically thick due to dust alone in the CH_3CN $K = 3$, $J = 13 \rightarrow 12$ line. The optical depth at a given radius is computed by moving integrating vertically through the disc. Different lines/colours correspond to different mass accretion rates and hence total disc masses. The lower panel shows the wavelength dependence of the opacity for different $a_{\max}$, the behaviour of which at the transition wavelength (black vertical line) determines the form of the upper panel.

regions that are optically thick in dust at larger radii. The expectation that massive discs around MYSOs will have such optically thick regions suggests that mass estimates from the continuum are likely to be underestimates (a similar conclusion was reached by Forgan et al. 2016).

The line opacity (equation 16) can also make the emission optically thick at radii larger than the inner region set by the dust opacity. However, we find that this is difficult to generalize as the line is marginally optically thick/thin. It depends on a plethora of factors including the chemical and thermal structure of the disc, the viewing angle (and hence the distribution of velocities along the line of sight), the disc mass, and so on. The inclination is particularly important to the line opacity, since self-absorption is sensitive to the kinematics along the line of sight. Our fiducial face-on disc is optically thick in the line centre out to around 600–800 au when the line opacity is included but at an inclination of 60 deg the same model is only optically thick in the line centre out to the extent set by the dust.

Overall, the $K = 3, 7,\ J = 13 \rightarrow 12$ CH_3CN lines cannot be assumed to be optically thick/thin at larger radii in the disc. However, the inner disc is expected to be optically thick due to the dust, with an extent that is sensitive to grain growth above an ISM size distribution.

4.5 Detecting disc fragmentation

In addition to discs with spiral substructure, we also consider fragmenting discs. Fragments contract into dense and hot clumps of gas and dust that ultimately might become stellar companions or planets.

Given their large distances, observations of MYSOs with ALMA will possess a typical spatial resolution of the order of a few 100s of au. For a distance of 3 kpc, the particular ALMA configuration we use for our synthetic observations yields spatial resolution of ~270 au. On the other hand, analysis of fragmentation in smoothed particle hydrodynamic simulations of self-gravitating discs around a solar mass star identified fragment radii of a few au (Hall et al. 2017). Even if fragment radii in massive systems are extrapolated to reach 10s of au, the fragments are not expected to be directly resolved by ALMA. However, they could still be detected appearing as beam-size clumps.

To explore the possibility of fragment detection, we insert an isothermal clump (see Section 2.1.3) into our fiducial two-arm spiral disc model with an inclination of 30 deg. The clump is placed inside one of the spiral arms, at disc mid-plane, and its velocity is equal to the local bulk velocity. We set the clump radius to 20 au. We do not consider smaller clumps because smaller clumps are challenging to resolve in our radiative transfer calculations of the 2000 au 3D disc model, even with an adaptive mesh,[4] and because a 40 au-diameter is still only a fraction of the beam, as discussed above. In six disc models with clumps, we vary the clump orbital distance (200–800 au), total mass (0.1–5 $M_\odot$), and temperature (150–1400 K).

Our main findings are that none of the clumps are detectable in the moment 0 nor the moment 1 maps and that the clumps are unambiguously detected in the continuum if they are sufficiently hotter than the surrounding matter, independently of the mass (over the range of masses we consider). Due to shock heating in the spiral arms, even at the orbital distance of 800 au the temperature in the spiral arms reaches ~300 K at the disc mid-plane. Consequently, in our synthetic observations, the clumps with a temperature of 1400 K are easily detectable, and those with a temperature of 300 K are not detectable at all. Again, we find that this is insensitive to the total clump mass in the range of 0.1–5 $M_\odot$. As an example, we show the continuum image of the model with a 0.1 $M_\odot$ and 1400 K clump in Fig. 9.

We therefore conclude that the next generation of ALMA observations towards MYSOs at ~3 kpc distances could expect to detect hotter fragments (i.e. those that have sufficiently contracted; Hall et al. 2017) in the dust continuum, but neither in the moment 0 map nor in the moment 1 map of line emission. We note, however, that we have not included kinematics of the matter accreting on to a fragment. We also infer that it would be more likely to detect clumps if they are outside of spiral arms due to increased contrast. As the spiral arms are expected to be the birthplace of fragments, this refers to the later stages of fragment evolution when the fragment decouples from the spiral density wave to fall on to a Keplerian orbit. Finally, we stress again that caution is required when interpreting substructure in terms of fragmentation, as we have shown in Section 3.1 that image filtering may result in a clumpy structure, which may be difficult to disentangle.

---

[4] The total processing time for a model with a clump is about 5 h using four cores for the radiative transfer calculations, and a single core for the rest, on a desktop, much longer than the 1 h processing for models without clumps due to a much higher spatial resolution of the data cubes necessary to resolve the clump.

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

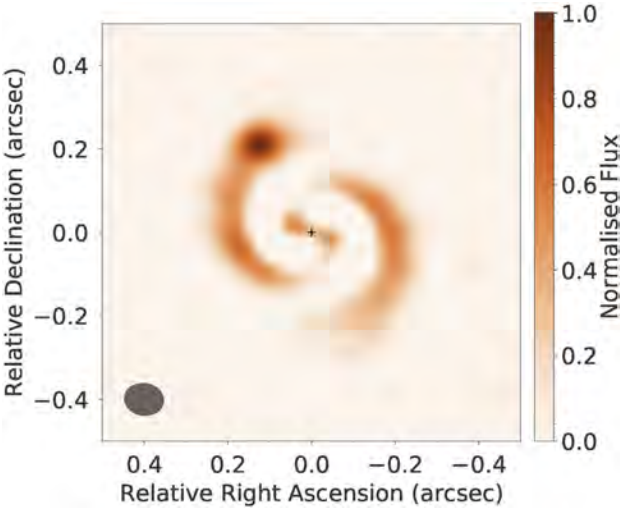


Figure 9. Simulated observed image of our fiducial disc model inclined at 30 deg, with a 0.1 $M_\odot$, 1400 K fragment inserted at the orbital distance of 800 au, in the 239 GHz continuum. Flux is normalized with respect to the peak value. The relative coordinates of the fragment in the image are (0.12 arcsec, 0.24 arcsec). A Gaussian fit has been subtracted from the original image to enhance substructure (see Section 3.1).

5 SUMMARY AND CONCLUSIONS

We have developed a means of quickly generating flexible synthetic observations of self-gravitating discs. Our approach is not entirely dissimilar to the parametric approach of Williams & Best (2014) that is applied to low mass, axisymmetric discs. Unlike the time-consuming process of performing synthetic imaging of hydrodynamical simulations, our approach permits us to explore a range of parameters and to understand what influences the detectability of substructure in gravitationally unstable discs around MYSOs. In our approach, we calculate semi-analytic models of such discs, including features such as spirals and clumps. These are then post-processed with a radiative transfer code to compute synthetic molecular line observations, which are then modified to account for interferometric/observing effects characteristic of ALMA. We draw the following main conclusions from this work:

(1) Spatial and kinematic filtering techniques, such as those described in Sections 3.1 and 3.2, are crucial for detection of substructure in discs around MYSOs at the kpc distances at which they are observed.

(2) We predict that ALMA will be able to resolve two-arm spirals at a majority of distances in the 1–5 kpc range explored here. Substructure in discs with a larger number of spiral arms, i.e. more weakly self-gravitating discs, will be much more difficult to infer.

(3) Spirals at ∼3 kpc will be easily resolved spatially in discs inclined up to ∼50 deg. At inclinations of ∼50–60 deg, it will be possible to identify spirals with the help of kinematic diagnostics (moment 1 maps). Above an inclination of 60 deg, we find that spirals can be detected only in PV diagrams.

(4) The molecular abundance distribution can, in principle, be a strong influencing factor for the detection of substructure. It is sensitive to the thermal conditions and molecules may be completely frozen out in the cold dense mid-plane, tracing only the hot disc atmosphere and shock-heated spirals. Weak atmospheric heating may lead to freeze out of the molecule higher into the atmosphere, but we find this not to be important for the detection of spirals, within the parameter space explored here.

(5) The optical depth of the CH_3CN lines considered here varies greatly with the disc inclination and other disc parameters, and the line is often marginally optically thick/thin in the outer disc. We find that the inner disc is always optically thick due to the dust. The extent of the optically thick inner disc may allow the amount of grain growth in the disc to be inferred.

(6) Our modelling suggests that disc fragmentation is unlikely to be detected in the line emission, but that fragments may be detectable in the continuum if they are sufficiently hotter than the surrounding disc material. Their detectability is not a strong function of their mass, but rather their temperature.

Our results suggest that upcoming observational campaigns will enable characterization of the immediate circumstellar environments of massive young stars. Such observations, under certain circumstances, will have the power to begin probing the physical and dynamical conditions of circumstellar discs around MYSOs, and therefore determine the formation mechanisms of massive stars themselves.

ACKNOWLEDGEMENTS

We thank the reviewer for helpful suggestions that improved the manuscript. MRJ is funded by the President’s PhD scholarship of the Imperial College London and the ‘Dositeja’ stipend from the Fund for Young Talents of the Serbian Ministry for Youth and Sport. TJH is funded by an Imperial College Junior Research Fellowship. JDI gratefully acknowledges support from the DISCSIM project, grant agreement 341137, funded by the European Research Council under ERC-2013-ADG and support from the STFC (grant number ST/R000549/1). DF gratefully acknowledges support from the ECOGAL project, grant agreement 291227, funded by the European Research Council under ERC-2011-ADG. CJC acknowledges support from the STFC (grant number ST/M001296/1). CW acknowledges financial support from the University of Leeds. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

REFERENCES

Armitage P. J., 2015, preprint (arXiv:1509.06382)
Bell K. R., Lin D. N. C., 1994, ApJ, 427, 987
Beltrán M. T., Cesaroni R., Neri R., Codella C., 2011, A&A, 525, A151
Beuther H., Walsh A. J., 2008, ApJ, 673, L55
Beuther H., Walsh A. J., Johnston K. G., Henning T., Kuiper R., Longmore S. N., Walmsley C. M., 2017, A&A, 603, A10
Bik A., Thi W. F., 2004, A&A, 427, L13
Black D. C., Bodenheimer P., 1975, ApJ, 199, 619
Boley A. C., 2009, ApJ, 695, L53
Boley P. A. et al., 2013, A&A, 558, A24
Caratti o Garatti A. et al., 2017, Nature Phys., 13, 276
Casassus S., 2016, PASA, 33, e013
Cesaroni R., 2005, Ap&SS, 295, 5
Cesaroni R., Galli D., Lodato G., Walmsley M., Zhang Q., 2006, Nature, 444, 703
Cesaroni R., Galli D., Lodato G., Walmsley C. M., Zhang Q., 2007, Protostars and Planets V. Univ. Arizona Press, Tucson, AZ, p. 197
Cesaroni R., Beltrán M. T., Zhang Q., Beuther H., Fallscheer C., 2011, A&A, 533, A73
Cesaroni R. et al., 2017, A&A, 602, A59
Chen H.-R. V., Keto E., Zhang Q., Sridharan T. K., Liu S.-Y., Su Y.-N., 2016, ApJ, 823, 125
Clarke C. J., 2009, MNRAS, 396, 1066

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

Collings M. P., Anderson M. A., Chen R., Dever J. W., Viti S., Williams D. A., McCoustra M. R. S., 2004, MNRAS, 354, 1133

Cossins P., Lodato G., Clarke C. J., 2009, MNRAS, 393, 1157

Cyganowski C. J., Brogan C. L., Hunter T. R., Smith R., Kruijssen J. M. D., Bonnell I. A., Zhang Q., 2017, MNRAS, 468, 3694

de Gregorio-Monsalvo I. et al., 2013, A&A, 557, A133

De Marco O., Izzard R. G., 2017, PASA, 34, e001

Deng H., Mayer L., Meru F., 2017, ApJ, 847, 43

Douglas T. A., Caselli P., Ilee J. D., Boley A. C., Hartquist T. W., Durisen R. H., Rawlings J. M. C., 2013, MNRAS, 433, 2064

Durisen R., Boss A., Mayer L., Nelson A., Quinn T., Rice W. K. M., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. Univ. Arizona Press, Tucson, AZ, p. 607

Edgar R., Clarke C., 2004, MNRAS, 349, 678

Forgan D., Rice K., Cossins P., Lodato G., 2011, MNRAS, 410, 994

Forgan D. H., Ilee J. D., Cyganowski C. J., Brogan C. L., Hunter T. R., 2016, MNRAS, 463, 957

Forgan D. H., Ramón-Fox F. G., Bonnell I. A., 2018a, MNRAS, 476, 2384

Forgan D. H., Hall C., Meru F., Rice W. K. M., 2018b, MNRAS, 474, 5036

Gammie C. F., 2001, ApJ, 553, 174

Hall C., Forgan D., Rice K., Harries T. J., Klaassen P. D., Biller B., 2016, MNRAS, 458, 306

Hall C., Forgan D., Rice K., 2017, MNRAS, 470, 2517

Harries T. J., 2000, MNRAS, 315, 722

Harries T. J., Douglas T. A., Ali A., 2017, MNRAS, 471, 4111

Haworth T. J., Harries T. J., 2012, MNRAS, 420, 562

Haworth T. J., Glover S. C. O., Koepferl C. M., Bisbas T. G., Dale J. E., 2018, New Astron. Rev., 82, 1

Herbst E., van Dishoeck E. F., 2009, ARA&A, 47, 427

Hosokawa T., Omukai K., 2009, ApJ, 691, 823

Hunter T. R., Brogan C. L., Cyganowski C. J., Young K. H., 2014a, ApJ, 788, 187

Hunter T. R., Brogan C. L., Cyganowski C. J., Young K. H., 2014b, ApJ, 788, 187

Hunter T. R. et al., 2017, ApJ, 837, L29

Ilee J. D. et al., 2013, MNRAS, 429, 2960

Ilee J. D., Cyganowski C. J., Nazari P., Hunter T. R., Brogan C. L., Forgan D. H., Zhang Q., 2016, MNRAS, 462, 4386

Izquierdo A. F., Galván-Madrid R., Maud L. T., Hoare M. G., Johnston K. G., Keto E. R., Zhang Q., de Wit W.-J., 2018, MNRAS, 478, 2505

Johnston K. G., Beuther H., Linz H., Boley P., Robitaille T. P., Keto E., Wood K., van Boekel R., 2014, Astrophys. Space Sci. Proc., 36, 413

Johnston K. G. et al., 2015, ApJ, 813, L19

Kahn F. D., 1974, A&A, 37, 149

Klassen M., Pudritz R. E., Kuiper R., Peters T., Banerjee R., 2016, ApJ, 823, 28

Kratter K. M., Matzner C. D., 2006, MNRAS, 373, 1563

Kraus S. et al., 2010, Nature, 466, 339

Krumholz M. R., Klein R. I., McKee C. F., 2007, ApJ, 656, 959

Kuiper R., Klahr H., Beuther H., Henning T., 2010, ApJ, 722, 1556

Kuiper R., Klahr H., Beuther H., Henning T., 2011, ApJ, 732, 20

Kurosawa R., Harries T. J., Bate M. R., Symington N. H., 2004, MNRAS, 351, 1134

Laughlin G., Bodenheimer P., 1994, ApJ, 436, 335

Lodato G., Rice W. K. M., 2004, MNRAS, 351, 630

Lodato G., Rice W. K. M., 2005, MNRAS, 358, 1489

Lumsden S. L., Hoare M. G., Urquhart J. S., Oudmaijer R. D., Davies B., Mottram J. C., Cooper H. D. B., Moore T. J. T., 2013, ApJS, 208, 11

Massey P., DeGioia-Eastwood K., Waterhouse E., 2001, AJ, 121, 1050

Mathis J. S., Rumpl W., Nordsieck K. H., 1977, ApJ, 217, 425

Maud L. T., Hoare M. G., Galván-Madrid R., Zhang Q., de Wit W. J., Keto E., Johnston K. G., Pineda J. E., 2017, MNRAS, 467, L120

McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Francisco, p. 127

Meyer D. M.-A., Kuiper R., Kley W., Johnston K. G., Vorobyov E., 2018, MNRAS, 473, 3615

Morbidelli A., Raymond S. N., 2016, J. Geophys. Res. (Planets), 121, 1962

Nakano T., 1989, ApJ, 345, 464

Pérez L. M. et al., 2016, Science, 353, 1519

Ramírez-Tannus M. C. et al., 2017, A&A, 604, A78

Rice W. K. M., Armitage P. J., 2009, MNRAS, 396, 2228

Rice W. K. M., Lodato G., Armitage P. J., 2005, MNRAS, 364, L56

Rosenfeld K. A., Andrews S. M., Hughes A. M., Wilner D. J., Qi C., 2013, ApJ, 774, 16

Rosenfeld K. A., Chiang E., Andrews S. M., 2014, ApJ, 782, 62

Rosen A. L., Krumholz M. R., McKee C. F., Klein R. I., 2016, MNRAS, 463, 2553

Rundle D., Harries T. J., Acreman D. M., Bate M. R., 2010, MNRAS, 407, 986

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Stamatellos D., Whitworth A. P., Bisbas T., Goodwin S., 2007, A&A, 475, 37

Testi L. et al., 2014, Protostars and Planets VI. Univ. Arizona Press, Tucson, AZ, p. 339

Toomre A., 1964, ApJ, 139, 1217

Vaidya B., Fendt C., Beuther H., 2009, ApJ, 702, 567

Walsh C., Millar T. J., Nomura H., Herbst E., Widicus Weaver S., Aikawa Y., Laas J. C., Vasyunin A. I., 2014, A&A, 563, A33

Walsh C., Nomura H., van Dishoeck E., 2015, A&A, 582, A88

Walsh C., Daley C., Facchini S., Juhász A., 2017, A&A, 607, A114

Wheelwright H. E., Oudmaijer R. D., de Wit W. J., Hoare M. G., Lumsden S. L., Urquhart J. S., 2010, MNRAS, 408, 1840

Williams J. P., Best W. M. J., 2014, ApJ, 788, 59

Williams J. P., Cieza L. A., 2011, ARA&A, 49, 67

Zapata L. A., Palau A., Galván-Madrid R., Rodríguez L. F., Garay G., Moran J. M., Franco-Hernández R., 2015, MNRAS, 447, 1826

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/482/4/4673/5173064 by University of Cambridge, mj577@cam.ac.uk on 30 June 2021

The Astrophysical Journal, 861:144 (27pp), 2018 July 10 https://doi.org/10.3847/1538-4357/aabcd0

© 2018. The American Astronomical Society. All rights reserved.

Inside-out Planet Formation. V. Structure of the Inner Disk as Implied by the MRI

Subhanjoy Mohanty[1], Marija R. Jankovic[1], Jonathan C. Tan[2], and James E. Owen[1]
[1] Department of Physics, Imperial College London, 1010 Blackett Lab., Prince Consort Rd., London SW7 2AZ, UK
s.mohanty@imperial.ac.uk, m.jankovic16@imperial.ac.uk
[2] Departments of Astronomy and Physics, University of Florida, Gainesville, FL 32611, USA

Received 2017 November 28; revised 2018 April 3; accepted 2018 April 5; published 2018 July 13

Abstract

The ubiquity of Earth- to super-Earth-sized planets found very close to their host stars has motivated in situ formation models. In particular, inside-out planet formation is a scenario in which planets coalesce sequentially in the disk, at the local gas pressure maximum near the inner boundary of the dead zone. The pressure maximum arises from a decline in viscosity, going from the active innermost disk (where thermal ionization yields high viscosities via the magnetorotational instability [MRI]) to the adjacent dead zone (where the MRI is quenched). Previous studies of the pressure maximum, based on α-disk models, have assumed ad hoc values for the viscosity parameter α in the active zone, ignoring the detailed MRI physics. Here we explicitly couple the MRI criteria to the α-disk equations, to find steady-state solutions for the disk structure. We consider both Ohmic and ambipolar resistivities, a range of disk accretion rates (10^{-10}–$10^{-8}\ M_\odot\ \mathrm{yr}^{-1}$), stellar masses (0.1–1 $M_\odot$), and fiducial values of the *non*-MRI α-viscosity in the dead zone ($\alpha_{\mathrm{DZ}} = 10^{-5}$ to 10^{-3}). We find that (1) a midplane pressure maximum forms radially *outside* the dead zone inner boundary; (2) Hall resistivity dominates near the inner disk midplane, perhaps explaining why close-in planets do *not* form in ~50% of systems; (3) X-ray ionization can compete with thermal ionization in the inner disk, because of the low steady-state surface density there; and (4) our inner disks are viscously unstable to surface density perturbations.

Key words: planets and satellites: formation – protoplanetary disks

1. Introduction

The *Kepler* mission has discovered more than 4000 exoplanet candidates from observations of their transits (e.g., Mullally et al. 2015; Coughlin et al. 2016). One of the great surprises from this data set is the ubiquity of Earth- and super-Earth-sized planets in very tight orbits, which have no solar system analogs. Specifically, more than 50% of Sun-like stars appear to harbor one or more planets of size 0.8–4 $R_\odot$ at orbital periods $P < 85$ days (i.e., shorter than Mercury's; Fressin et al. 2013). Similarly, nearly all M dwarfs seem to host one or more 0.5–4 $R_\odot$ sized planets at $P < 50$ days (Dressing & Charbonneau 2015). Note that the single-planet systems included in these statistics may have as-yet-undetected smaller planets as well. Moreover, a large fraction ($\gtrsim$30%) of the close-in multiplanet *Kepler* systems appear dynamically packed (i.e., cannot admit an additional planet without becoming unstable; Fang & Margot 2013). Thus, a major, and possibly the dominant, planet formation mechanism in our Galaxy produces small planets very close to the central star, with a large fraction of these in tightly packed multiplanet systems. Two main scenarios have been advanced to explain such planets: (1) formation in the outer disk followed by inward migration (e.g., Kley & Nelson 2012; Cossou et al. 2013, 2014), and (2) formation in situ (Hansen & Murray 2012, 2013; Chiang & Laughlin 2013; Chatterjee & Tan 2014, hereafter CT14).

The inward migration scenario tends to produce planets that are trapped in orbits of low-order mean motion resonances, which is not a particular feature of these *Kepler* systems (Baruteau et al. 2014; Fabrycky et al. 2014). Recently discovered trends in the atmospheric photoevaporation of these planets also indicate an Earth-like (rock/iron) core composition, implying formation inward of the ice line and thus arguing against significant migration (Owen & Wu 2017).

The inside-out planet formation (IOPF) scenario proposed by CT14 is a new type of in situ formation model. It is based on the fact that the effective viscosity in the disk is expected to decline, moving radially outward from the innermost disk—where efficient thermal ionization of alkali metals (Umebayashi & Nakano 1988) activates the magnetorotational instability (MRI; Balbus & Hawley 1991), leading to high viscosities—to the adjacent "dead zone," where decreasing thermal ionization leads to a suppression of the MRI by Ohmic resistivity, yielding low viscosities (Gammie 1996). In a steady-state disk, i.e., one with a constant disk accretion rate $\dot{M}$, this fall-off in viscosity produces a local maximum in the gas pressure in the vicinity of the dead zone inner boundary (DZIB). The IOPF mechanism proposes that dust grains that have grown to ~centimeter-sized "pebbles" in the outer disk (Hu et al. 2017) and are drifting radially inward are trapped in this pressure maximum, within which they rapidly coalesce into a protoplanet. The protoplanet itself is also expected to be trapped in this region (Hu et al. 2016) and thus able to continue growing (especially by pebble accretion), until it becomes massive enough to open a gap a few Hill radii wide in the disk.

Material interior to the inner rim of this gap will tend to drain rapidly (on a local viscous timescale) onto the star. While some replenishment of this interior region may continue owing to gas flowing across the gap, densities here are expected to decrease, potentially leaving the outer rim of the gap subject to direct stellar X-ray/UV irradiation. This can activate the MRI in disk gas close to the outer rim, over a thickness set by how far stellar ionizing photons penetrate radially into the rim (e.g., Chiang & Murray-Clay 2007). A new DZIB then forms at the outer edge of this MRI-active region, creating a new pressure trap where incoming pebbles can coagulate into another planet. The process continues until the pebble supply from the outer disk is exhausted, leaving behind a system of closely packed inner planets.

The formation of gas pressure maxima is thus central to the IOPF model. In particular, the location of the first maximum, controlled by thermal ionization of alkalis in the inner disk, sets the orbital radius of the innermost (so-called "Vulcan") planet in the system. The goal of this paper is to investigate the formation of this first pressure maximum.

There have been several previous works studying pressure traps in the disk created by changes in the viscosity (e.g., CT14; Kretke & Lin 2007; Kretke et al. 2009; Kretke & Lin 2010). All of these have been based on a steady-state Shakura–Sunyaev α-disk model, wherein the disk accretion rate is constant, viscous heating due to accretion is the main source of energy input, and the disk viscosity is parameterized in terms of the quantity α. Crucially, however, these studies have all adopted ad hoc prescriptions of α for computational ease, without accounting for the detailed physics of the MRI.

Conversely, several groups have investigated the behavior of active and dead zones in the disk, accounting for the detailed effects of non-ideal MHD and complex gas and dust chemistry on the MRI, either using direct numerical simulations (e.g., Bai 2011, 2017; Bai & Stone 2011; Turner et al. 2010) or based on the MRI criteria implied by such simulations (e.g., Perez-Becker & Chiang 2011a, 2011b; Mohanty et al. 2013). However, these studies all assume a passive disk (heated and ionized by stellar irradiation) and a predetermined temperature and surface density profile (usually minimum-mass solar nebula [MMSN]). Consequently, the results are generally neither in steady state ($\dot{M}$ varies with radius) nor applicable to the inner disk (where viscous heating dominates).

Our aim here is to marry the two approaches: we wish to solve for the structure of the inner disk assuming a steady-state, viscously heated α-disk, but with α determined self-consistently from detailed considerations of the MRI and non-ideal MHD effects. To the best of our knowledge, this is the first such unified disk model (Keith & Wardle [2014] present an elegant self-consistent α-disk model for circumplanetary disks, but their MRI-α is a more parameterized version than ours, with a saturation value set arbitrarily). As such, the results are germane not only to the IOPF mechanism and the specific purpose of locating a pressure maximum in the inner disk but also to the broader goal of understanding the structure of viscously heated steady-state disks with MRI-driven accretion and non-ideal MHD.

In Section 2, we provide an overview of our methodology and discuss some critical caveats to our assumption of MRI-driven accretion. In Section 3, we summarize the α-disk model, and in Section 4, we describe our treatment of the MRI. Our technique for calculating α and $\dot{M}$ is detailed in Sections 5 and 6, and our method of determining equilibrium solutions is outlined in Section 7. We present our results in Section 8 and discuss their implications in Section 9.

2. Overview of Methodology and Caveats

2.1. Methodology

We wish to investigate the location of the pressure maximum in the inner disk, by solving for the inner disk structure in steady state (i.e., with constant $\dot{M}$) and assuming that the MRI is the dominant magnetically controlled mechanism for local mass and angular momentum transport. We further wish to do this in the context of the Shakura–Sunyaev α-disk model. Consequently, we must solve the *coupled* set of equations for the MRI and disk structure: coupled because the effective viscosity parameter α from the MRI and the attendant $\dot{M}$ both depend on the underlying disk structure (as well as on the magnetic field strength B), while the disk structure itself, in the Shakura–Sunyaev model, is determined by α and $\dot{M}$ (and stellar parameters).

Briefly, we use a grid-based method of solution. A grid of disk structures is calculated for a desired $\dot{M}$ and a range of input values for α and field strength B; the MRI-induced output α and corresponding accretion rate $\dot{M}$ are derived for each of these disk structures, and the chosen solution structure is the one in which the output values of α and $\dot{M}$ match the input ones. We find that a range of such solutions are possible differing in B; a unique solution is chosen under the assumption that the MRI is maximally efficient, i.e., generates the largest field it can support (but see "Caveats" below).

Pressure Maximum: How do our solutions produce a pressure maximum? In the α-disk model, the gas pressure is a decreasing function of both α and radius. This leads to a turnover in pressure at the radial location where our derived α falls to its minimum value. What defines this minimum in our methodology? In previous work as well as in this paper, a lower limit ("floor") on α is set by its value $\alpha_{\rm DZ}$ in the dead zone, where the MRI is quenched but various (nonmagnetic) hydrodynamic/gravitational instabilities may still generate viscous stresses. Fiducial values for this floor are chosen based on theory and numerical simulations; we explore the plausible range $\alpha_{\rm DZ} = 10^{-5}$ to 10^{-3} (discussed in more detail later). The pressure maximum then occurs where the α in the MRI-active zone decreases to this dead zone limiting value. This floor will always be reached if heating due to viscous accretion (a decreasing function of radius for constant $\dot{M}$) is the only source of the ionization required to kindle the MRI (as assumed here; but see also X-rays/UV below).

Simplifications: In this pilot study, we adopt a number of simplifications: no ionization by stellar photons (X-ray or UV; we only consider thermal ionization due to accretion heating), ionization of a single alkali species (i.e., no complex chemical network), no dust, and a fixed opacity of 10 cm^2 g^{-1}. Relaxing these assumptions presents no conceptual difficulties, and we shall do so in a subsequent paper (M. R. Jankovic et al. 2018, in preparation); the inclusion of more physics will certainly change the precise location of the pressure maximum (e.g., dust grains will reduce the MRI efficiency, and X-rays may change the limiting value of α; these effects and others are discussed at appropriate junctures). Nevertheless, as an initial step, the mathematical ease afforded by these simplifications allows us to clearly present our methodology and identify important general trends in the solutions.

2.2. Caveats

Finally, there are crucial caveats, applicable to all work so far on pressure maxima in the inner disk (including this paper), concerning the basic assumption that mass and angular momentum transport are controlled by the MRI. In the innermost disk, where the inductive term in the field evolution equation greatly exceeds the resistive terms, the MRI is indeed likely to be dominant and maximally efficient (e.g., Bai 2013). Farther out, however, where the resistivities become non-negligible, the situation is much more complicated.

Specifically, first, when Ohmic and ambipolar resistivities are both important, vertically stratified 3D simulations (Bai 2013;

Bai & Stone 2013; Gressel et al. 2015) imply that (a) in the absence of any net vertical magnetic flux, the MRI is extremely weak, with an effective viscosity orders of magnitude lower than required to power the observed accretion rates in classical T Tauri stars; and (b) with even a small net vertical field, MRI turbulence is completely smothered (because, while the MRI is initially present, the field is subsequently amplified to strengths *greater* than that at which the MRI can operate under ambipolar diffusion; i.e., the assumption of maximally efficient MRI is no longer valid). The flow over the entire vertical extent of the disk now becomes fully laminar, and a magnetized disk wind develops instead, which efficiently carries angular momentum away from the disk and drives accretion at rates consistent with observations. In other words, where Ohmic and ambipolar effects are both important, mass accretion seems driven primarily by vertical angular momentum transport by magnetized winds, and not radial transport by the MRI.

Second, introducing the Hall effect into the above situation complicates matters further, depending on whether the net vertical magnetic field is aligned or anti-aligned with the spin axis of the disk (Bai 2014, 2015, 2017; Lesur et al. 2014; Simon et al. 2015). When the two are aligned (i.e., $\boldsymbol{\Omega} \cdot \boldsymbol{B} > 0$), the Hall shear instability (HSI) generates laminar viscous stresses via the amplification of horizontal components of the field (Kunz 2008), leading to strong radial angular momentum transport and hence significant mass accretion (in addition to the magnetized-wind-driven accretion at comparable rates). Conversely, when the field and disk spin axis point in opposite directions ($\boldsymbol{\Omega} \cdot \boldsymbol{B} < 0$), the horizontal field is considerably suppressed, and mass and angular momentum transport are predominantly wind driven.

At face value, these results suggest that using the Shakura–Sunyaev viscous disk model to search for a pressure maximum, with the expectation that α declines sharply across the interface between the MRI-active innermost disk and the adjacent dead-zone-dominated region, might not be a valid exercise for two reasons. First, in the region usually characterized as "dead zone" dominated, angular momentum in the aforementioned simulations is mainly transported vertically *out* of the disk by wind-related torques, instead of being radially redistributed *within* the disk by standard viscous torques (either hydrodynamic/gravitational within the dead zone, or MRI in an overlying active layer). Thus, the Shakura–Sunyaev viscous model is invalid here. Second, when the field and disk spin axis are aligned, the HSI activates efficient mass and angular momentum transport all the way down to the midplane here (in addition to wind-related transport higher up); i.e., there is no dead zone in any sense.

Nevertheless, it is premature to write off an inner disk pressure maximum in the standard viscous disk context. All the above simulations are restricted to radii $\gtrsim$1 au, an order of magnitude farther out than the presumed location of the pressure maximum at $\lesssim$few tenths of an au (the simulation domain of Bai [2017] formally extends into 0.6 au, but they deem the results at <2 au to be vitiated by boundary effects). Thus, it remains to be seen whether the above conclusions apply to our region of interest in the inner disk. Concurrently, if the close-in planets we address here are indeed formed in situ from inward-migrating solids, then some sort of pressure trap seems inescapable in this region, in order to corral these solids and prevent their falling into the star. As such, continuing this line of inquiry currently appears justified.

Finally, even if the wind/Hall results from the simulations extend to much smaller radii, a pressure maximum is still plausible (and, in general, a significant change in disk structure is expected) at the interface between the innermost MRI-active turbulent disk and the adjacent wind-dominated laminar disk, because of the qualitative difference in physical conditions between the two regions. The Shakura–Sunyaev α-disk model will not apply across the interface, and the controlling factor for any change in disk structure may be the radial distribution of magnetic flux (since the field ultimately determines the strength of the MRI, the wind, and the Hall effect; X. N. Bai 2017, private communication), rather than the radial behavior of α as assumed here. Nonetheless, the α-disk model will still apply to the MRI region, and insights into the latter gleaned from the present work will remain useful.

3. Disk Model

A detailed derivation of the steady-state (temporally constant) disk structure within the Shakura–Sunyaev viscous α-disk model is given by Hu et al. (2016, hereafter H16). We summarize the main results here. The viscosity parameter α is defined by the relation

$$\nu = \alpha \frac{c_s^2}{\Omega}, \tag{1}$$

where ν is the viscosity, c_s the sound speed, and Ω the Keplerian angular velocity at any given disk radius. Now, the α-disk model is fundamentally derived from vertically integrated quantities (surface density and accretion rate; see H16); as such, the "α" that enters into it is more precisely a vertical average. This issue is often elided (e.g., H16 do not discuss it) under the implicit assumption that α is vertically constant or slowly varying. However, in a vertically stratified disk (such as we will find), with MRI-active zones sandwiched between inactive ones, the nature of the viscosity changes with height, and the latter assumption is invalid. In this case, the relevant quantity is the *effective* viscosity parameter $\bar{\alpha}$, defined as the pressure-weighted vertical average of α:

$$\bar{\alpha} \equiv \frac{\int_{-\infty}^{+\infty} \alpha P_{\rm gas}\, dz}{\int_{-\infty}^{+\infty} P_{\rm gas}\, dz} = \frac{\int_{-\infty}^{+\infty} \alpha \rho\, dz}{\int_{-\infty}^{+\infty} \rho\, dz}, \tag{2}$$

where the second equality (derived using $P_{\rm gas} = \rho c_s^2$ for density ρ) holds only for a vertically isothermal disk (so that c_s^2 is constant with height; we shall assume such isothermality further below). We show how to calculate $\bar{\alpha}$ in Section 5. We explicitly append the subscript "gas" to pressure P to differentiate the gas pressure from the magnetic pressure ($P_{\rm B}$, encountered later); for all other quantities (density, temperature, etc.) we drop this subscript, since they always refer to gas alone in this paper.

With this definition of $\bar{\alpha}$, the steady-state gas surface density (summing both above and below the midplane) at any orbital radius r in the disk is given by

$$\Sigma(r) = 139.4\, \gamma_{1.4}^{-4/5}\, \kappa_{10}^{-1/5}\, \bar{\alpha}_{-3}^{-4/5}\, M_{*,1}^{1/5} \times (f_r \dot{M}_{-9})^{3/5}\, r_{\rm AU}^{-3/5}\ \ {\rm g\ cm^{-2}} \tag{3}$$

for a normalized stellar mass $M_{*,1} \equiv M_*/1\,M_\odot$, accretion rate $\dot{M}_{-9} \equiv \dot{M}/10^{-9}\,M_\odot\,\mathrm{yr}^{-1}$, radial distance $r_{\rm AU} \equiv r/1$ au, opacity $\kappa_{10} \equiv \kappa/10\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, adiabatic index $\gamma_{1.4} \equiv \gamma/1.4$, and effective viscosity parameter $\bar{\alpha}_{-3} \equiv \bar{\alpha}/10^{-3}$, with $f_r \equiv (1 - \sqrt{R_{\rm in}/r})$ for a disk inner edge located at $R_{\rm in}$ (if the disk extends to the stellar surface, then $R_{\rm in} = R_*$, the stellar radius). The associated midplane temperature is given by

$$T_0(r) = 192.6\,\gamma_{1.4}^{-1/5}\,\kappa_{10}^{1/5}\,\bar{\alpha}_{-3}^{-1/5}\,M_{*,1}^{3/10} \times (f_r \dot{M}_{-9})^{2/5}\,r_{\rm AU}^{-9/10}\ \ \mathrm{K} \tag{4}$$

if viscous heating is the main source of energy input (i.e., heating by stellar irradiation is ignored). The midplane pressure is then

$$P_{0,\rm gas}(r) = 0.773\,\gamma_{1.4}^{-7/5}\,\kappa_{10}^{-1/10}\,\bar{\alpha}_{-3}^{-9/10}\,M_{*,1}^{17/20} \times (f_r \dot{M}_{-9})^{4/5}\,r_{\rm AU}^{-51/20}\ \ \mathrm{erg\ \ cm^{-3}}, \tag{5}$$

and the midplane density (which follows from the ideal gas law $P_{\rm gas} = \rho\ k_B T/\mu$, for particles with mean molecular mass $\mu \approx 2.34\,m_{\rm H}$, where $m_{\rm H}$ is the atomic mass of hydrogen) is

$$\rho_0(r) = (1.133 \times 10^{-10})\,\gamma_{1.4}^{-6/5}\,\kappa_{-10}^{-3/10}\,\bar{\alpha}_{-3}^{-7/10}\,M_{*,1}^{11/20} \times (f_r \dot{M}_{-9})^{2/5}\,r_{\rm AU}^{-33/20}\ \ \mathrm{g\ \ cm^{-3}}. \tag{6}$$

Crucially, Equations (5) and (6) show that the *midplane* pressure and density do not depend on the *local* α, but rather on its vertically averaged value $\bar{\alpha}$. In other words, the midplane pressure (and thus density) is sensitive to conditions in the entire column pressing down from above (as intuitively expected), not simply local ones. This has the following important consequence. As we will show, the midplane pressure maximum does not form where the dead zone first develops in the midplane (i.e., at the DZIB, which is where the midplane α reaches its minimum), as often assumed. Instead, it forms farther out radially, where the *effective* parameter $\bar{\alpha}$ reaches its minimum (because the MRI-active zone continues outward for some distance above the dead zone). Thus, we will find that the midplane pressure maximum is actually located *within* the dead zone.

Unlike H16, we assume for simplicity that the disk is vertically isothermal (i.e., $\gamma = 1$). Strictly speaking, this is slightly inconsistent with the derivation of the midplane temperature (Equation (4) above) by H16, following the formalism of Hubeny (1990), wherein the temperature depends on the vertical optical depth in the disk. However, implementing this dependence couples together the vertical temperature and density profiles in a complicated fashion (H16 avoid this because they are concerned with just midplane values). Moreover, at small optical depths ($\tau \ll 1$), the temperature is also highly sensitive to the details of the appropriate radiative processes (a simplistic treatment of which leads to an infinitely hot disk surface; see discussion by Hubeny 1990); addressing these is beyond the scope of this paper. On the other hand, at large optical depths ($\tau \gg 1$), T only varies very slowly with depth, as $\tau^{1/4}$ (see Hubeny 1990). Therefore, since we expect the inner disk to only be active in optically thick regions close to the midplane, we approximate the vertical temperature profile in the region of interest by the midplane values: $T(z,\ r) \sim T_0\ (r)$.

The (isothermal) sound speed is then $c_s = \sqrt{k_B T_0/\mu}$, and the vertical pressure profile in hydrostatic equilibrium becomes

$$P_{\rm gas}(z,\ r) = P_{0,\rm gas}(r)\,\exp\left(-\frac{z^2}{z_H^2}\right), \tag{7}$$

where the pressure scale height is defined as $z_H \equiv \sqrt{2}\ c_s/\Omega$. Finally, we assume a constant opacity of $\kappa = 10\ \mathrm{cm}^2\,\mathrm{g}^{-1}$, approximately the expected value in protoplanetary disks (e.g., Wood et al. 2002). H16 use the detailed opacity tables of Zhu et al. (2012), where the values depend on the pressure and temperature structure of the disk, and solve for the equilibrium opacities and structure iteratively. In our case, however, the disk structure equations are already coupled to the MRI ones, and the two sets must be solved simultaneously. Introducing a further interdependence with opacity adds a level of complexity that we set aside in this exploratory work. We do compare, a posteriori, our constant κ to the values implied by Zhu et al. (2012) for our equilibrium disk structure, to gauge the discrepancy between the two; in general, we find our value to be reasonable.

4. MRI

Our treatment of the MRI generally follows that of Mohanty et al. (2013), except we consider ionization by thermal collisions instead of by X-rays, and we do not include grains. Here we summarize the major points of our analysis. The physical conditions required for the MRI to operate are set out in Section 4.1; our treatment of thermal ionization and recombination is discussed in Section 4.2; and the calculation of the various resistivities (Ohmic, ambipolar, and Hall), which determine whether or not the MRI criteria are met, is described in Section 4.3.

4.1. Criteria for Active MRI

We discuss the necessary conditions for active MRI in Appendix A and only state the final results here. The Ohmic Elsässer number Λ is defined as

$$\Lambda \equiv \frac{v_{Az}^2}{\eta_O \Omega}, \tag{8}$$

where η_O is the Ohmic resistivity and v_{Az} the vertical component of the local Alfvén velocity ($\equiv B_z/\sqrt{4\pi\rho}$, where B_z is the vertical field strength and ρ the local gas density). Similarly, the ambipolar Elsässer number Am is defined as

$$Am \equiv \frac{v_A^2}{\eta_A \Omega}, \tag{9}$$

where η_A is the ambipolar resistivity and v_A the local total Alfvén velocity ($\equiv B/\sqrt{4\pi\rho}$, where B is the rms field strength).[3]

With these definitions, the conditions for sustaining active MRI are

$$\Lambda > 1 \tag{10}$$

---

[3] Our reasons for adopting v_{Az} in Equation (8) but v_A in Equation (9) are supplied in the discussion preceding Equation (33) and in footnote (8), in Appendix A.

and

$$\beta > \beta_{\min}(Am). \tag{11a}$$

Here $\beta \equiv P_{\rm gas}/P_{\rm B}$ is the plasma β parameter (with magnetic pressure $P_{\rm B} \equiv B^2/8\pi$), and the minimum allowed value of β—denoted by $\beta_{\min}$—is a function of the ambipolar Elsässer number Am (Bai & Stone 2011):

$$\beta_{\min}(Am) = \left[\left(\frac{50}{Am^{1.2}}\right)^2 + \left(\frac{8}{Am^{0.3}} + 1\right)^2\right]^{1/2}. \tag{11b}$$

Equation (10) encapsulates the reasonable condition that, when Ohmic resistivity dominates, the MRI is sustained when the growth rate of the fastest-growing MRI mode exceeds its dissipation rate. When ambipolar diffusion dominates, on the other hand, Bai & Stone (2011) find that, in the strong-coupling (single-fluid) limit applicable to protoplanetary disks (see discussion preceding Equation 36(a) in Appendix A), the MRI can operate at *any* value of Am, provided that the field is sufficiently weak. Equations 11(a), (b) then define what "sufficiently weak" means: it signifies that the plasma β parameter must exceed a minimum threshold $\beta_{\min}$. Specifically, it implies that the gas pressure must dominate over the magnetic pressure in the disk for the MRI to function (see discussion following Equation 37(b)). An "active zone" is where both conditions (10) and (11) are satisfied, allowing efficient MRI; a "dead zone" is where condition (10) is not met, so that Ohmic resistivity shuts off the MRI; and a "zombie zone" (following the nomenclature of Mohanty et al. 2013) is where condition 11(a) is not satisfied, so that ambipolar diffusion quenches the MRI.

Note that the effects of Hall diffusion are ignored in the above analysis. As discussed in Section 2 and Appendix A, in the presence of a net vertical background field, the Hall effect can amplify the MRI or suppress it, depending on whether the field is aligned or anti-aligned with the spin axis of the disk. Quantifying this effect is beyond the scope of this paper. However, we do investigate the Hall effect a posteriori, by calculating the Hall Elsässer number ($\chi \equiv v_A^2/(|\eta_H|\Omega)$; see Appendix A and Equation (35)) everywhere in our solutions. In any region where $\chi < 1$, which we call a "Hall zone," Hall diffusion has a strong effect on the MRI; we discuss the potentially critical implications of such regions for planet formation.

4.1.1. Choice of Magnetic Field Strength

Both the Ohmic and ambipolar conditions for active MRI, Equations (10) and (11), depend on the magnetic field strength: via v_{Az} in Λ_O and $P_{\rm B}$ in β. Indeed, for a given set of stellar parameters and a fixed accretion rate, we will see that there exist an infinite number of solutions, each corresponding to a different disk structure with a different field strength B.

The question then is how to determine an appropriate B. We do so by assuming that (a) *the magnetic field strength is constant with height across the active layer* and (b) *the MRI is maximally efficient, generating the strongest possible field that still allows the MRI to operate (i.e., still satisfies the constraint* $\beta > \beta_{min}$*).*

The same assumptions are made by Mohanty et al. (2013) and Bai (2011). A roughly constant B across the active layer is expected from MRI-driven turbulent mixing (Bai 2011, and references therein), justifying condition (a). Condition (b) encapsulates the notion that (in the absence of any other mechanism) the MRI turbulence will continue to amplify the field up to some maximum value $B_{\max}$ corresponding to $\beta_{\min}$, beyond which the MRI is quenched (i.e., the instability is self-regulated). Our implementation of this condition to derive equilibrium disk solutions is described in Section 5.

Finally, we note that numerical simulations of the MRI by Sano et al. (2004) indicate that the total rms field strength B and its vertical component B_z are related by $B^2 \sim 25B_z^2$, a condition we adopt. Thus, though our Ohmic MRI condition is defined in terms of $v_{Az} \propto B_z^2$ while the ambipolar condition is in terms of $P_{\rm B} \propto B^2$, one need specify only B or B_z, not both independently.

4.2. Thermal Ionization and Recombination

In the hot inner regions of the disk, ionization is dominated by thermal collisions, with the equilibrium level of thermal ionization of an atomic species a given by the Saha equation:

$$\frac{n_e\, n_{+,a}}{n_{0,a}} = \frac{1}{\lambda_e^3}\,\frac{g_e\, g_{+,a}}{g_{0,a}}\exp\left(\frac{-\mathcal{I}_a}{k_B T}\right). \tag{12}$$

Here n_e is the number density of free electrons, and $n_{0,a}$ and $n_{+,a}$ are the number densities of neutral atoms and singly ionized ions, respectively, of species a; $\lambda_e \equiv \sqrt{h^2/(2\pi m_e k_B T)}$ is the thermal de Broglie wavelength of electrons of mass m_e; g_e (=2), $g_{0,a}$, and $g_{+,a}$ are the degeneracy of states for free electrons, neutrals, and ions, respectively; and $\mathcal{I}_a$ is the ionization energy.

We note the following simplifications when only one, singly ionized species (e.g., an alkali metal; see below) participates in ionization/recombination. In this case, charge conservation requires $n_e = n_{+,a}$ and $n_{0,a} = n_a - n_e$ (where n_a is the total number density of species a). Since molecular hydrogen, with number density $n_{\rm H_2}$, forms the vast bulk of the gas, we adopt the standard expressions for fractional ionization, $x_e \equiv n_e/n_{\rm H_2}$, and the abundance of species a, $x_a \equiv n_a/n_{\rm H_2}$. Writing the entire right-hand side of the Saha equation above as $\mathcal{S}_a(T)$, a little algebra then yields $x_e = [-1 \pm \sqrt{1 + 4x_a(n_{\rm H_2}/\mathcal{S}_a)}\,]/[2(n_{\rm H_2}/\mathcal{S}_a)]$. This leads to two limiting physical solutions: when $n_{\rm H_2} \to 0$ (more precisely, when $4x_a(n_{\rm H_2}/\mathcal{S}_a) \ll 1$), we get $x_e \approx x_a$, and when $n_{\rm H_2} \to \infty$ (more precisely, when $4x_a(n_{\rm H_2}/\mathcal{S}_a) \gg 1$), we get $x_e \approx \sqrt{x_a \mathcal{S}_a/n_{\rm H_2}}$. Also note that, without any ionization of hydrogen itself, and with hydrogen being the most abundant species by far, we have $n_{\rm H_2} \approx n_n$ (number density of neutrals) $\approx n_{\rm tot}$ (total number density of particles). We use these results later.

In order of decreasing ionization potential $\mathcal{I}_a$, the important elements in the inner disk are He, H, Mg, Na, and K (Keith & Wardle 2014). The exponential in the Saha equation ensures the on/off behavior of thermal ionization, wherein most of the atoms of a species a become ionized over a narrow range of temperatures around the ionization temperature $T_a \equiv \mathcal{I}_a/k_B$. Thus, since we expect the disk temperature to generally decrease radially outward and we are concerned with the outer edge of the active zone, we only consider potassium (K) here, which has the smallest $\mathcal{I}_a$ and is thus ionized farthest out. Our adopted quantities for K are listed in Table 1; in this pilot study, we neglect its depletion into grains.

With a chemical network comprising collisional ionization/recombination of just one singly ionized element, the recombination rate is simply $dn_e/dt = k_{ei}\, n_e\, n_{+,a} = k_{ei}\, n_e^2$, where $k_{ei} = 3 \times 10^{-11}/\sqrt{T}$ cm^3 s^{-1} (Ilgner & Nelson 2014) is the rate coefficient for electron–ion collisions and the second equality follows from charge conservation. The recombination timescale is then (e.g., Bai 2011)

$$t_{\rm rcb} \sim \frac{n_e}{dn_e/dt} = \frac{1}{k_{ei}\, n_e}. \tag{13}$$

We will compare this timescale to the dynamical time $t_{\rm dyn}$ to verify whether our equilibrium solutions are in the strongly coupled limit described in Appendix A.

4.3. Resistivities

Armed with the equilibrium abundances of electrons, ions, and neutrals computed via the Saha equation, we derive the resistivities in the disk and thus examine where the disk is MRI active by the criteria of Section 4.1 (for a field strength B given by the considerations of Section 4.1.1). We follow Wardle (2007) in writing the Ohmic, Hall, and Pederson conductivities (σ_O, σ_H, and σ_P respectively) as

$$\sigma_O = \frac{ec}{B} \sum_j n_j\, |Z_j|\, \beta_j \tag{14}$$

$$\sigma_H = \frac{ec}{B} \sum_j \frac{n_j\, Z_j}{1 + \beta_j^2} \tag{15}$$

$$\sigma_P = \frac{ec}{B} \sum_j \frac{n_j\, |Z_j|\, \beta_j}{1 + \beta_j^2}, \tag{16}$$

where the summation is over all charged species j (in our case, $j = e$ for electrons and i for singly charged ions of K), with particle mass m_j, number density n_j, and charge $Z_j e$ (with $Z_j = \pm 1$ for us). The Hall parameter β_j (not to be confused with the plasma β parameter) is the ratio of the gyrofrequency of a charged particle of species j to its collision frequency with neutrals (of mean particle mass $m_n = \mu m_{\rm H}$ and density ρ_n):

$$\beta_j = \frac{|Z_j| eB}{m_j\, c} \frac{1}{\gamma_j\, \rho_n}. \tag{17}$$

Here $\gamma_j = \langle\sigma v\rangle_j/(m_j + m_n)$ is the drag coefficient and $\langle\sigma v\rangle_j$ the rate coefficient for collisional momentum transfer between charged species j and neutrals, making $\gamma_j \rho_n$ the collision frequency with neutrals. Note that $\beta_i \ll \beta_e$ (since $m_i \gg m_e$).

The resistivities may then be written as

$$\eta_O = \frac{c^2}{4\pi\sigma_O} \tag{18}$$

$$\eta_H = \frac{c^2}{4\pi\sigma_\perp} \frac{\sigma_H}{\sigma_\perp} \tag{19}$$

$$\eta_A = \frac{c^2}{4\pi\sigma_\perp} \frac{\sigma_P}{\sigma_\perp} - \eta_O, \tag{20}$$

where $\sigma_\perp \equiv \sqrt{\sigma_H^2 + \sigma_P^2}$ is the total conductivity perpendicular to the magnetic field.

If electrons and ions are the only charged species (which is the case for us, without grains), then the above equations imply (1) $\eta_H = \beta_e\, \eta_O$ and $\eta_A = \beta_i \beta_e \eta_O$; (2) consequently, while η_O is independent of the magnetic field strength B, η_H and η_A scale linearly and quadratically, respectively, with B; and (3) the ambipolar Elsässer number in Equation (9), $Am \equiv v_A^2/\eta_A \Omega$, reduces to (using $\beta_i \ll \beta_e$) $Am \approx \gamma_i \rho_i/\Omega$. The three diffusion regimes then correspond to (e.g., Wardle 2007) $\beta_i \ll \beta_e \ll 1$ (Ohmic: neither electrons nor ions are tied to the field, being coupled instead to the neutrals through frequent collisions), $\beta_i \ll 1 \ll \beta_e$ (Hall: electrons are tied to the field while ions are not), and $1 \ll \beta_i \ll \beta_e$ (ambipolar: both electrons and ions are tied to the field, and drift together through the sea of neutrals).

Table 1

Adopted Parameters for Potassium

A[a] (amu)	$x_{\rm K}$[a,b]	$\mathcal{I}_{\rm K}$[a] (eV)	$g_{+,{\rm K}}/g_{0,{\rm K}}$[c]
39.10	1.97 10^{-7}	4.34	1/2

Notes.

[a] Atomic mass (A), abundance ($x_{\rm K} \equiv n_{\rm K}/n_{\rm H_2}$), and ionization potential ($\mathcal{I}_{\rm K}$) from Keith & Wardle (2014).

[b] Keith & Wardle (2014) cite the abundance of K relative to H atoms as 9.87×10^{-8}; our value is relative to H molecules and thus double their value.

[c] Rouse (1961) cites $g_+/g_0 = 1/2$ for the alkali metal sodium; we adopt the same value for the alkali potassium.

To compute the resitivities, we use the rate coefficients from Wardle & Ng (1999):

$$\langle\sigma v\rangle_e = 10^{-15} \left(\frac{128\, k_B T_e}{9\pi m_e}\right)^{1/2} {\rm cm^3\ s^{-1}} \tag{21}$$

$$\langle\sigma v\rangle_i = 1.6 \times 10^{-9}\ {\rm cm^3\ s^{-1}}, \tag{22}$$

where T_e is the electron temperature, assumed here to equal the disk gas temperature given by Equation (4).

5. Calculation of $\bar{\alpha}$

Finally, we must connect the MRI formulation of accretion to the α-disk model. In particular, we must specify how to calculate the effective viscosity parameter $\bar{\alpha}$, defined by Equations (1) and (2), that goes into the Shakura–Sunyaev disk model. The derivation is supplied in Appendix B; we only state the main results here. At any radius in the disk, we expect a vertically layered structure: in the hot innermost disk close to the star, we expect an MRI-active zone straddling the midplane, bounded by a zombie zone close to the disk upper and lower surfaces; farther out, where the disk is cooler, we expect a dead zone straddling the midplane, a zombie zone close to the disk upper and lower surfaces, and an MRI-active zone sandwiched between the two.[4] For a vertically isothermal disk (as assumed here), $\bar{\alpha}$ at any radius is then given in general by

$$\bar{\alpha} = \frac{\sum_i (N_i\, \bar{\alpha}_i)}{N_{\rm tot}}, \tag{23}$$

---

[4] Such a layered disk model was first put forward by Gammie (1996) and has since been recovered in various semianalytic studies invoking both Ohmic and ambipolar diffusion and based on local shearing-box MHD simulations (e.g., Bai 2011; Dzyurkevich et al. 2013; Mohanty et al. 2013), as well as by global stratified 3D simulations invoking only Ohmic dissipation (Dzyurkevich et al. 2010). (Although all these studies concern larger radii in the disk where the ionization is primarily due to stellar irradiation, e.g., X-rays, instead of being thermally driven as in this paper, the basic physics for active MRI remains the same as outlined in Section 4.1.) As noted in Section 2, such a model becomes invalid if, in the presence of both Ohmic and ambipolar diffusion and a net vertical field, the MRI is shut off, angular momentum transport is driven by winds, and the entire vertical extent of the disk becomes laminar instead; our models in this paper do not speak to the latter situation.

where the summation is over i = MRI (active zone), DZ (dead zone), and ZZ (zombie zone). Here N_i is the one-sided column density of the ith zone, $N_{\rm tot} = \sum_i N_i$ is the total one-sided column density of the disk at that radius (i.e., from the surface to the midplane), and $\bar{\alpha}_i$ is the effective viscosity parameter within the ith zone (see below). Thus, for a vertically isothermal disk, $\bar{\alpha}$ at any radius is the column-weighted mean of the active, dead, and zombie effective viscosity parameters.

The different $\bar{\alpha}_i$ ($\bar{\alpha}_{\rm MRI}$, $\bar{\alpha}_{\rm DZ}$, and $\bar{\alpha}_{\rm ZZ}$) are specified as follows. Within the MRI-active zone, we have (see Appendix B)

$$\bar{\alpha}_{\rm MRI} = \frac{2}{3}\left(\frac{1}{2\langle\beta\rangle}\right), \tag{24}$$

where $\langle\beta\rangle = \langle P_{\rm gas}\rangle/P_{\rm B}$ is the plasma-beta parameter averaged over the thickness of the active layer (note that we assume that B and hence P_B are vertically constant, so the averaging is only over $P_{\rm gas}$). In the dead and zombie zones, where the MRI is quenched, various hydrodynamical processes can still produce residual (*non*-MRI) stresses; numerical simulations of these suggest an associated effective α in the approximate range $\sim 10^{-5}$ to 10^{-3} (e.g., Dzyurkevich et al. 2010, 2013, and references therein; Malygin et al. 2017, and references therein). Additionally, without carrying out detailed hydrodynamic simulations, we have no concrete way of judging how the effective α in the dead and zombie zones might differ. For simplicity, therefore, we assume that the effective viscosity parameter in the dead and zombie zones is the same (i.e., $\bar{\alpha}_{\rm DZ} = \bar{\alpha}_{\rm ZZ}$) and find equilibrium solutions for the disk structure for three different fiducial values of $\bar{\alpha}_{\rm DZ}$ spanning the range implied by the numerical solutions:

$$\bar{\alpha}_{\rm DZ}\ (=\bar{\alpha}_{\rm ZZ}) = 10^{-5}\ \text{or}\ 10^{-4}\ \text{or}\ 10^{-3}. \tag{25}$$

Importantly, note that $\bar{\alpha}_{\rm DZ}$ also sets a minimum value ("floor") on $\bar{\alpha}_{\rm MRI}$: when our calculations imply that a region is formally "MRI-active" (i.e., satisfies Equations (10) and (11)), but nevertheless has $\bar{\alpha}_{\rm MRI}$ less than our adopted $\bar{\alpha}_{\rm DZ}$, we expect that the residual hydrodynamic stresses there will dominate over the MRI stress. We therefore declare such a region to be dead by fiat and assign it an effective viscosity parameter equal to $\bar{\alpha}_{\rm DZ}$.

6. Accretion Rates

Within a given disk zone (MRI active, dead, or zombie), the local accretion rate (positive inward) at any radius is $\dot{M}_i = -2(r\Omega)^{-1}\partial[2\pi r^2 \int_{2h_i} T_{r\phi,i}\,dz]/\partial r$, where $2h_i$ is the thickness of the ith zone (summed over both sides of the midplane) and $T_{r\phi,i}$ the particular shear stress operating in that zone. For a vertically isothermal disk, this reduces to (see Appendix B)

$$\dot{M}_i = \frac{12\pi\ \mu m_{\rm H}}{r\Omega}\frac{\partial}{\partial r}(r^2 c_s^2\ N_i \bar{\alpha}_i), \tag{26}$$

where i = MRI, DZ, or ZZ; N_i is the one-sided column density of the ith zone; and the values of the various $\bar{\alpha}_i$ are specified in the previous section.

Similarly, the total accretion rate at any radius, i.e., the local sum of the rates through the different vertical layers, is $\dot{M} \equiv \sum_i \dot{M}_i = -2(r\Omega)^{-1}\partial[2\pi r^2 \int_{-\infty}^{+\infty} T_{r\phi}\,dz]/\partial r$. Now, in a real disk, the chemistry and ionization, and hence the thickness (column) of any zone and the field strength, will generally vary with radius, and there is no physically compelling reason to expect the accretion rate through any given zone (Equation (26) above) to be radially or temporally constant. In steady state, however, the *total* accretion rate must by definition be a constant in both time and radius (to prevent temporal changes in the local surface density). Imposing this condition on our solutions, the total accretion rate becomes (see Appendix B)

$$\dot{M} = \frac{3\pi\ \bar{\alpha}\ c_s^2\ \Sigma}{f_r\ \Omega}, \tag{27}$$

the standard expression for a constant accretion rate in a vertically isothermal α-disk. $\bar{\alpha}$ here is given by Equation (23), Σ ($=2\mu m_{\rm H}\ N_{\rm tot}$) is the total surface density summed over both sides of the midplane, and $f_r \equiv (1 - \sqrt{R_{\rm in}/r})$. As an aside, note that it is the combination $f_r\dot{M}$ that appears in the disk structure equations (Section 3), which is independent of $R_{\rm in}$ by Equation (27).

7. Method for Determining the Equilibrium Solution

At a given disk radius around a fixed stellar mass, specified input values of the accretion rate and mean viscosity parameter ($\dot{M}_{\rm in}$ and $\bar{\alpha}_{\rm in}$) determine the pressure, temperature, and density via the disk structure Equations (3)–(7). The latter quantities, combined with the Saha Equation (12), set the fractional ionization. The disk structure and ionization, together with a specified field strength B, then determine the resistivities (via Equations (14)–(22)) and hence the extent of the active layer via the MRI conditions (10)–(11). This in turn yields the output mean viscosity parameter and accretion rate ($\bar{\alpha}_{\rm out}$ and $\dot{M}_{\rm out}$) implied by the MRI (Equations (23)–(25) and (27)). We find self-consistent equilibrium solutions ($\dot{M}_{\rm out} = \dot{M}_{\rm in}$ and $\bar{\alpha}_{\rm out} = \bar{\alpha}_{\rm in}$) through a grid-based technique, as follows.

For a specified stellar mass M_* and disk radius r and a desired disk accretion rate $\dot{M}_{\rm in}$, we determine the disk structure and ionization for a range of input $\bar{\alpha}$: $\bar{\alpha}_{\rm in} = [\bar{\alpha}_{\rm DZ}, 1]$, spanning the gamut of plausible values given the assumed $\bar{\alpha}_{\rm DZ}$ in the dead zone. For each of these disk structures, we then derive the height of the active layer, and thus the MRI-implied $\dot{M}_{\rm out}$ and $\bar{\alpha}_{\rm out}$, for a range of field strengths: $B = [10^{-5}, 10^{3}]$ G, which covers the plausible range in stellar accretion disks. A self-consistent disk structure solution is then one for which $\dot{M}_{\rm out} = \dot{M}_{\rm in}$ and $\bar{\alpha}_{\rm out} = \bar{\alpha}_{\rm in}$.

How exactly such a solution is determined is illustrated in Figure 1 for a fiducial case: $M_* = 1\ M_\odot$, $\dot{M}_{\rm in} = 10^{-9}\ M_\odot\ {\rm yr}^{-1}$, $\bar{\alpha}_{\rm DZ} = 10^{-4}$, at radius $r = 0.02$ au. The x- and y-axes show $\bar{\alpha}_{\rm in}$ and $\bar{\alpha}_{\rm out}$, respectively, while the overplotted grayscale contour map shows the magnetic field strength B (with the white curves marking contours of constant B). The overlaid solid black contours are the output accretion rate $\dot{M}_{\rm out}$.

We see that, along the locus of equilibrium solutions (solid blue line, along which $\bar{\alpha}_{\rm out} = \bar{\alpha}_{\rm in}$ *and* $\dot{M}_{\rm out} = \dot{M}_{\rm in}$), increasing $\bar{\alpha}$ corresponds to increasing field strength B (this is easily seen by noticing that contours of constant $\dot{M}$ are steeper than the contours of constant B, so B changes—increases—as one marches up the blue solution locus with $\dot{M}$ constant). In other words, for any given $\bar{\alpha}$, there exists a field strength B that yields an equilibrium solution with the desired $\dot{M}$, up to some upper limit in $\bar{\alpha}$ (corresponding to an upper limit in B). How do we

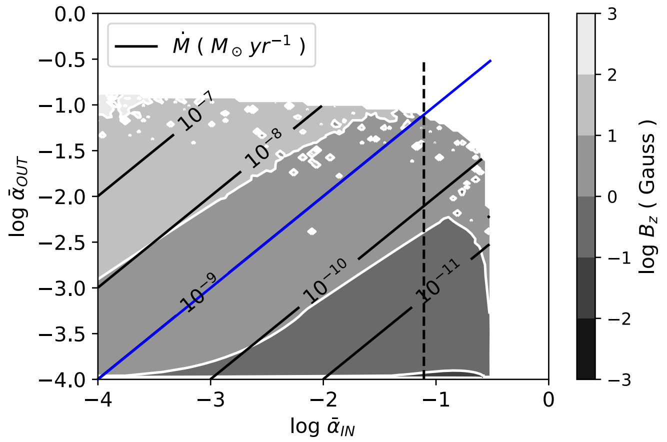


Figure 1. Output values $\dot{M}_{\rm out}$ (*solid black lines*) and $\bar{\alpha}_{\rm out}$ (*y*-axis) corresponding to input values for $\bar{\alpha}_{\rm in}$ (*x*-axis) and magnetic field strength *B* (gray scale with white contours), for fixed stellar and disk parameters $M_* = 1\ M_\odot$, $\dot{M}_{\rm in} = 10^{-9}\ M_\odot\ {\rm yr}^{-1}$, $\bar{\alpha}_{\rm DZ} = 10^{-4}$, and radius $r = 0.02$ au. The solid blue line indicates the locus of all solutions $\bar{\alpha}_{\rm in} = \bar{\alpha}_{\rm out}$; note that this line also overlies the required accretion rate for a self-consistent solution: $\dot{M}_{\rm out} = \dot{M}_{\rm in} = 10^{-9}\ M_\odot\ {\rm yr}^{-1}$. The intersection of the vertical dashed line with the blue line marks the position of the final adopted equilibrium solution, corresponding to the largest value of *B* that still allows the MRI to operate (there are no solutions with active MRI beyond this *B*, which is why this solution lies at the edge of the map).

choose a unique solution from among these infinite possibilities? We do so by invoking our assumption (see Section 4.1.1) that the MRI is maximally efficient, generating the strongest possible field that still allows the MRI to operate. Thus, we choose the maximum *B*, and thus the maximum $\bar{\alpha}$ (marked by a dashed vertical line), for which an equilibrium solution exists.

For a given M_* and $\dot{M}$, we repeat the above calculations for a range of radii *r*, to determine $\bar{\alpha}$ as a function of radius. Our calculations begin at a disk inner edge of $R_{\rm in} = R_*$. We continue working outward in radius until our derived equilibrium solution for $\bar{\alpha}$ falls to the assumed floor value $\bar{\alpha}_{\rm DZ}$. Beyond this radius, there is no active zone any more in our model, and we simply assume a constant $\bar{\alpha} = \bar{\alpha}_{\rm DZ}$.

8. Results

We first present a detailed discussion of our solution for the fiducial case ($M_* = 1\,M_\odot$, $\dot{M} = 10^{-9}\,M_\odot\,{\rm yr}^{-1}$, $\bar{\alpha}_{\rm DZ} = 10^{-4}$) in Section 8.2: the disk structure and location of the pressure maximum (Sections 8.1.1–8.1.4), behavior of the accretion flow in different layers (Section 8.1.5), the appearance of a viscous instability (Section 8.1.6), and the validity of various assumptions (Sections 8.1.7–8.1.8). We then briefly discuss the solutions arising from variations in our fiducial parameters ($\alpha_{\rm DZ}$, $\dot{M}$, and M_*), pointing out any salient differences along the way (Section 8.3–8.4). Piecewise polynomial fits to our $\bar{\alpha}$ and *B* results as a function of radius are provided in Appendix C for all cases.

8.1. Fiducial Model: $M_* = 1\,M_\odot$, $\dot{M} = 10^{-9}\,M_\odot\,yr^{-1}$, $\bar{\alpha}_{DZ} = 10^{-4}$

For this $M_* = 1\,M_\odot$ case, the stellar radius and effective temperature are $R_* = 2.33\,R_\odot$ and $T_{\rm eff} = 4350$ K, respectively (using the evolutionary models of Baraffe et al. [1998],[5] for a fiducial age of 1 Myr). In this and all following solutions, the disk inner radius is situated at the stellar surface (i.e., $R_{\rm in} = R_*$), and our MRI calculations stop at the radius where the effective viscosity parameter $\bar{\alpha}$ falls to the floor value $\bar{\alpha}_{\rm DZ}$ (i.e., where the pressure maximum forms). Beyond this radius, the disk structure is calculated assuming that the viscosity parameter remains constant at $\bar{\alpha} = \bar{\alpha}_{\rm DZ}$.

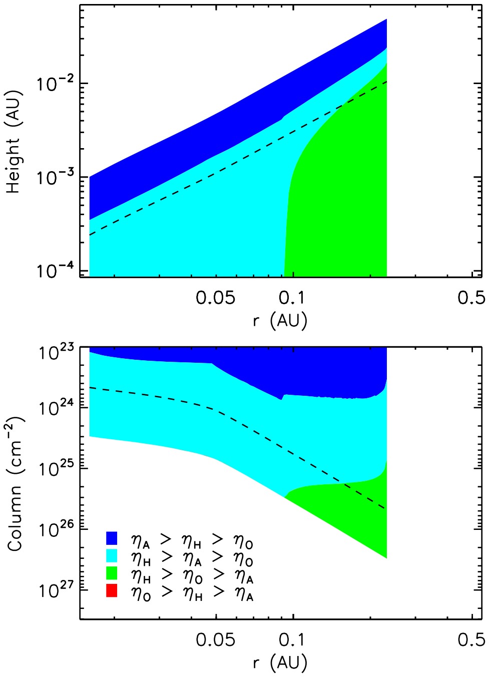


Figure 2. Relative importance of the Ohmic (η_O), Hall ($|\eta_H|$), and ambipolar (η_A) resistivities as a function of location in the inner disk, for our fiducial disk model. The top panel shows vertical location in units of actual height above the midplane; the bottom panel shows vertical location in units of column density measured from the disk surface. The dashed line in both panels indicates one pressure scale height. Note that there is *no* region where Ohmic resistivity dominates over both Hall and ambipolar resistivities. See Section 8.1.1.

8.1.1. Dominant Resistivities

Figure 2 shows the relative importance of the three resistivities—η_O, $|\eta_H|$, and η_A—as a function of location in the inner disk. Ambipolar diffusion dominates over Hall and Ohmic in the surface layers, while Hall resistivity dominates everywhere else at these radii. Ohmic resistivity is not dominant anywhere, though it is larger than ambipolar closer to the midplane at radii $\gtrsim$0.09 au. This distribution of resistivities is also depicted more quantitatively in Figure 3, where we plot η_O, $|\eta_H|$, and η_A as functions of scale height at various radii.

The physics underlying the above behavior can be extracted from Figure 4, where we plot the fractional ionization ($x_e \equiv n_e/n_{\rm H_2}$) and number density of neutral molecular hydrogen ($n_{\rm H_2}$) as functions of scale height at different radii. Recall that $n_{\rm H_2} \approx n_n$ (number density of neutrals) $\approx n_{\rm tot}$ (total number density), given the overwhelming relative abundance of hydrogen and the very low ionization fractions in general (since potassium, with total abundance $x_{\rm K} \equiv n_{\rm K}/n_{\rm H_2} \sim 2 \times 10^{-7}$, is the only ionized species here). Combining this with our results from Section 4.2 for one singly ionized species, we get $x_e \approx n_e/n_n \propto \sqrt{\mathcal{S}_{\rm K}(T)/n_n}$ when n_n is sufficiently high (with the subscript “K” on $\mathcal{S}$ denoting the specific case of

---

[5] Specifically, the iso.3 models with mixing length = 1.9× pressure scale height, as required to fit the Sun.

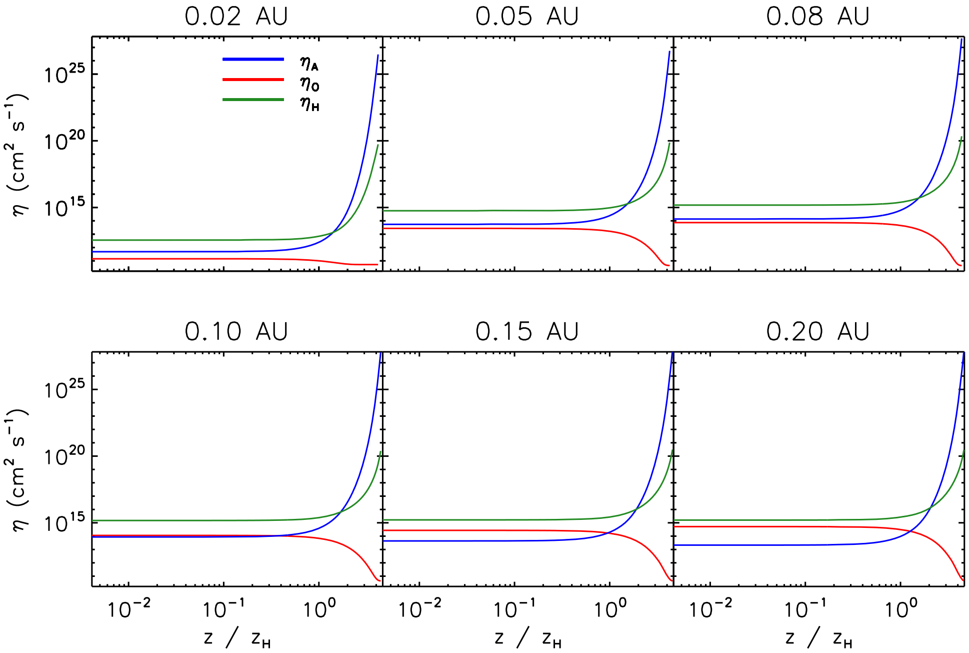


Figure 3. Ambipolar (η_A), Ohmic (η_O), and Hall ($|\eta_H|$) resistivities as a function of height above the midplane (in units of the local scale height z_H), at various radii for our fiducial disk model. See Section 8.1.1.

potassium). For the same conditions, and combining the latter relationship with results from Section 4.3, we also have $\eta_O \propto 1/x_e \propto \sqrt{n_n/\mathcal{S}_{\rm K}(T)}$, $|\eta_H| \propto B/(x_e n_n) \propto B/\sqrt{\mathcal{S}_{\rm K}(T) n_n}$, and $\eta_A \propto B^2/(x_e n_n^2) \propto B^2/\sqrt{\mathcal{S}_{\rm K}(T) n_n^3}$. Thus, at any fixed radius in Figure 4 (with $\mathcal{S}_{\rm K}(T)$ constant since vertically isothermal), the ionization fraction x_e increases rapidly above a scale height z_H as hydrostatic equilibrium causes $n_{\rm tot} \sim n_{\rm H_2} \sim n_n$ to drop, with all the potassium ionized ($x_e \rightarrow x_{\rm K}$ as $n_{\rm H_2} \rightarrow 0$; see Section 4.2) by a few × z_H. Consequently, at a given radius in Figure 3, η_O decreases with height above $\sim z_H$, while $|\eta_H|$ *increases* with height and η_A increases even faster (note that the field strength B is vertically constant at fixed radius in our calculations).

In summary, though a large fraction of the alkali atoms are ionized near the disk surface, the total density here is too low to collisionally couple either ions or electrons to the bulk fluid of neutrals, and hence ambipolar diffusion dominates; closer to the midplane, the density increases sufficiently to tie ions (but not electrons) to the neutrals, making Hall resistivity dominant, but the density is still too low for Ohmic resistivity to compete with either Hall or ambipolar diffusion. Beyond ~0.09 au, the rising density and falling temperature are sufficient (combined with a declining B; see Section 8.1.2 below) for Ohmic resistivity to exceed ambipolar diffusion near the midplane, but still not enough to allow Ohmic resistivity to exceed Hall diffusion here.

8.1.2. Active, Dead, and Zombie Zones

Figure 5 shows our derived locations of the MRI-active zone, the dead zone (where Ohmic resistivity shuts off the MRI: $\Lambda < 1$) and the zombie zone (where ambipolar diffusion cuts off the MRI: $\beta < \beta_{\rm min}$). We emphasize that the effects of Hall resistivity on the MRI are not accounted for here: we only consider the effects of Ohmic and ambipolar diffusion, even in regions where $|\eta_H|$ dominates over η_O and η_A. Nevertheless, we also overplot the Hall zone, where the Hall Elsässer number $\chi < 1$: this is where the Hall influence on the MRI is significant (see further below), and *should* be accounted for in future work. Figure 6 shows the associated field strength B as a function of radius, while Figure 7 shows the midplane radial behavior of the ionization fraction, plasma β parameter, and Ohmic Elsässer number.

We see from Figure 5 that, from the inner edge of the disk to ~0.05 au, the active zone extends from the midplane up to a roughly constant fraction of the scale height, bounded above by a zombie zone, while from 0.05 to 0.09 au the active zone narrows considerably, with the zombie zone pushing down increasingly toward the midplane. At 0.09 au, a dead zone rises up sharply from the midplane; from here on, the active zone is confined to a very thin and continuously narrowing layer sandwiched between the zombie zone above and dead zone below, until the MRI is completely choked off at ~0.25 au, at which point our calculations stop (beyond this radius, we assume a constant $\bar{\alpha} = \bar{\alpha}_{\rm DZ}$, leading to the formation of a pressure maximum at this radius; see following sections).

These trends in the active, dead, and zombie zones can be understood as follows. At a fixed height (in scale height units), η_O increases while the field strength B declines, going radially outward from the inner edge to ~0.05 au (Figures 3 and 6). The combined effect is to decrease the Ohmic Elsässer number Λ; however, it still remains high enough to allow the active zone to straddle the midplane (e.g., see midplane Λ in Figure 7, bottom panel). The weakening of the field over this radial span instead serves to keep the plasma β sufficiently large, $\beta > \beta_{\rm min}$, so that ambipolar diffusion does not cut off the MRI too close to the midplane and drive $\dot{M}$ below the desired steady-state value (e.g., see midplane β and $\beta_{\rm min}$ in Figure 7, middle panel). By 0.05 au, however, the midplane Λ has fallen to unity (Figure 7). Now the field B has two choices: either continue to weaken, making $\Lambda < 1$ at the midplane (i.e., creating a dead zone there) and thus driving the active zone upward, or strengthen instead, thereby keeping the MRI alive

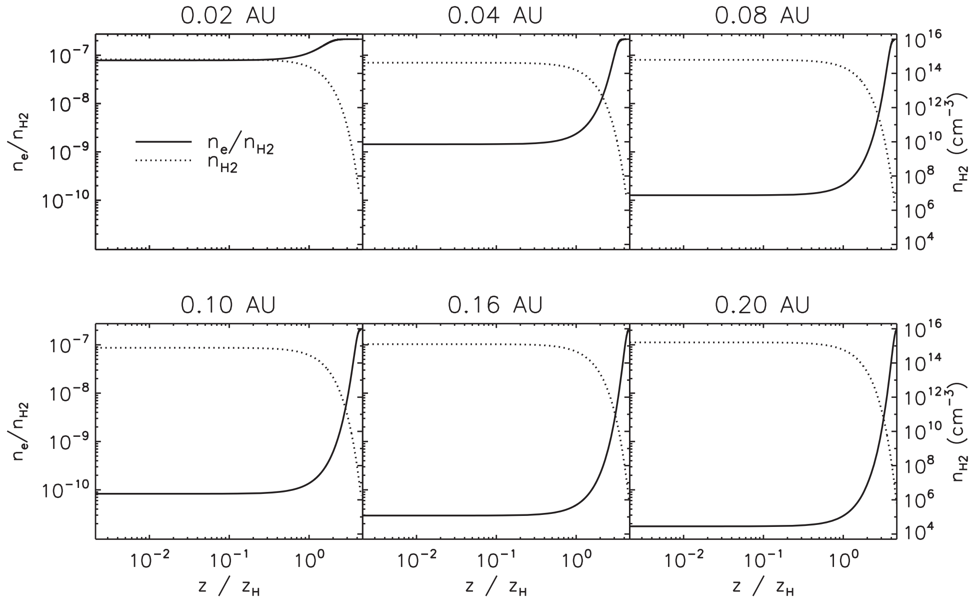


Figure 4. Fractional ionization (n_e/n_{H_2}) and molecular hydrogen density (n_{H_2}) as a function of height above the midplane (in units of the local scale height z_H), at various radii for our fiducial disk model. See Section 8.1.1.

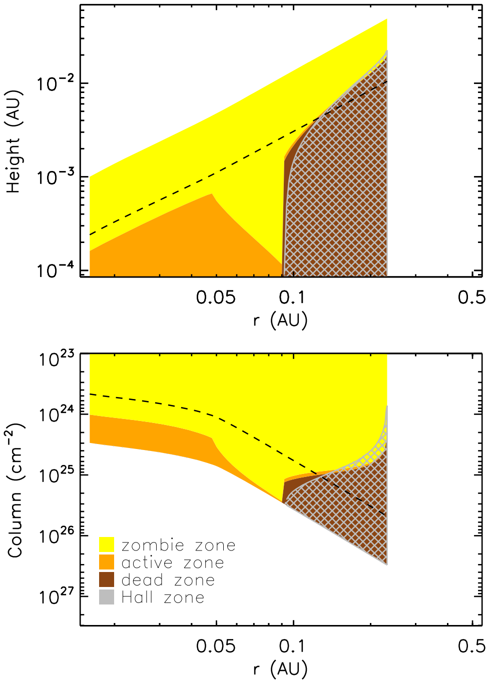


Figure 5. Various MRI zones in our fiducial disk model as a function of disk location. The top panel shows vertical location in units of height above midplane; the bottom panel shows it in units of column density. The dashed line indicates the disk scale height. Orange denotes the MRI-active zone (i.e., where $\Lambda_O > 1$ and $\beta > \beta_{\rm min}$), brown denotes the dead zone (where $\Lambda_O < 1$); yellow denotes the zombie zone (where $\beta < \beta_{\rm min}$), and the gray hashed region denotes the Hall zone (where $\chi < 1$). Note that, beyond 0.09 au, the active zone rises above the midplane and continues as a thin layer sandwiched between the dead and zombie zones, until it is finally quenched totally at the outer edge of our solution at ~0.25 au. See Section 8.1.2.

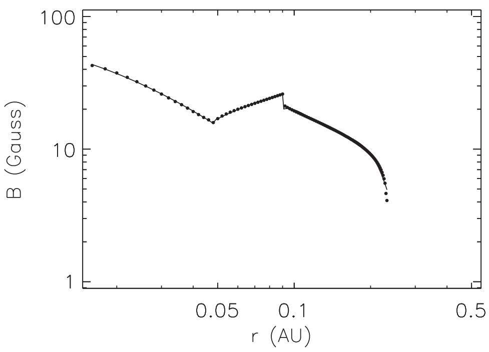


Figure 6. Magnetic field strength B as a function of radius, for our fiducial disk model. Filled circles are our model results, and the overplotted solid line is a combined piecewise polynomial fit to these results. Note that the jump at 0.09 au is not a physical discontinuity, but a result of our finite-grid radial sampling. See Section 8.1.2 and Table 1 in Appendix C for the polynomial fit parameters.

around the midplane, but suppressing β and thus allowing the zombie zone to descend toward the midplane. Since *we assume that the MRI is maximally efficient*, i.e., generates the strongest possible field that still allows the MRI to survive, it is the latter solution that is chosen (Figure 6), yielding the observed active and zombie zone shapes in Figure 5 over 0.05–0.09 au. The quantitative increase in B here (and thus change in $\bar{\alpha}$ and hence in surface density Σ; see following sections) is such that the $\dot{M}$ (by Equation (27)) remains at the required value.

By 0.09 au, however, the zombie zone has descended all the way to the midplane (i.e., $\beta = \beta_{\rm min}$ at the midplane; Figure 7). Now the field has no choice but to weaken again (Figure 6), in order to maintain any active zone at all. As B decreases, a dead zone develops at the midplane, the zombie zone lower boundary is impelled upward, and a thin active layer forms between the dead and zombie regions (Figure 5). This situation cannot continue indefinitely, though, since the dead zone upper boundary keeps rising with radius (as η_O continues to grow; Figure 3). Finally, at ~0.25 au, the MRI-active zone is squeezed shut completely, as the upper edge of the dead zone meets the lower edge of the zombie zone. No further changes in B can alter this, since the dead region would expand upward for smaller B, and the zombie region would expand downward for larger B. Thus, this is the radius where the effective viscosity parameter $\bar{\alpha}$ falls to its minimum value $\bar{\alpha}_{\rm DZ}$ (since the disk is

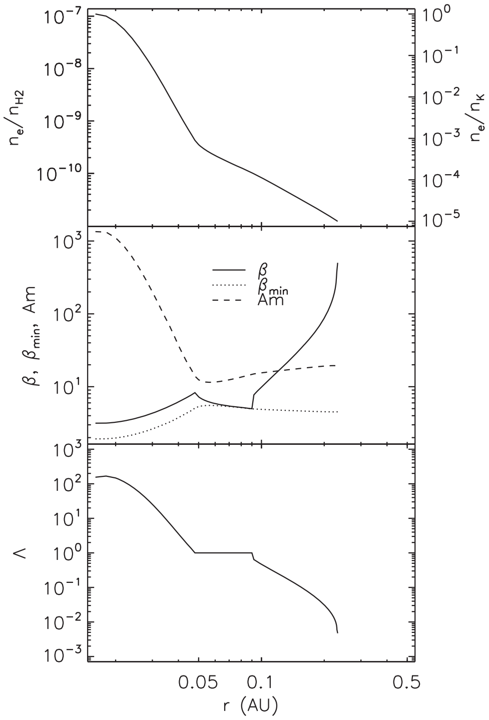


Figure 7. Various MRI related quantities in the midplane, plotted as a function of radius, for our fiducial disk model. Top: fractional abundance of electrons, expressed relative to both the number density of hydrogen molecules (n_e/n_{H_2}; left axis) and the number density of potassium nuclei (n_e/n_{K}; right axis). Middle: plasma β parameter, minimum value β_{min} required for active MRI, and ambipolar Elsässer number *Am*. Bottom: Ohmic Elsässer number Λ. See Section 8.1.2.

now fully MRI-dead vertically), and hence where the midplane pressure maximum forms.

The above result raises an important point missed in most earlier work: the midplane gas pressure does not reach its maximum *at* the inner edge of the dead zone (i.e., at ~0.09 au in this example), but rather somewhat radially *beyond* this edge (at ~0.25 au here). In other words, the midplane gas pressure achieves its maximum value *within* the dead zone. This is a straightforward consequence of two facts: (1) the midplane pressure in the Shakura–Sunyaev model is not a function of simply the local midplane value of α, but rather its *vertically averaged* value $\bar{\alpha}$ (see Equation (5) and discussion in Section 3); and (2) the active zone does not abruptly come to an end when a dead zone appears in the midplane, but instead climbs above the dead zone and continues outward for some distance, thereby pushing the location of minimum $\bar{\alpha}$ (and so maximum midplane pressure; see Figures 8 and 9 further below) beyond the DZIB. As such, pebbles drifting inward along the midplane will become trapped within the dead zone itself, where conditions are less turbulent than at the active/dead zone interface farther in, with potentially important implications for planet formation.

In the context of the location of the pressure maximum, we now discuss the potential importance of some effects ignored in our simplified treatment here.

Other Ionized Elements: We have only treated potassium here, with the justification that—as the element with the lowest ionization potential (4.34 eV) among the important species in the inner disk (Section 4.2)—it remains ionized farthest out and is thus most relevant to the location of the pressure maximum.

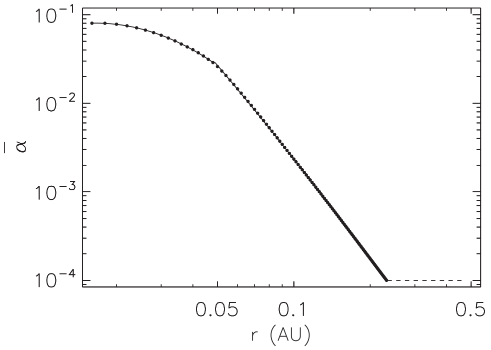


Figure 8. Vertically averaged viscosity parameter $\bar{\alpha}$, plotted as a function of radius for our fiducial disk model. Filled circles represent our model results; the overplotted solid line is a combined piecewise polynomial fit to these results. Our MRI calculations end at ~0.25 au, the radius at which $\bar{\alpha}$ falls to our adopted floor value $\bar{\alpha}_{\mathrm{DZ}} = 10^{-4}$. Beyond this radius we assume a constant $\bar{\alpha} = \bar{\alpha}_{\mathrm{DZ}}$, as indicated by the dashed horizontal line. See Table 1 in Appendix C for the polynomial fit parameters.

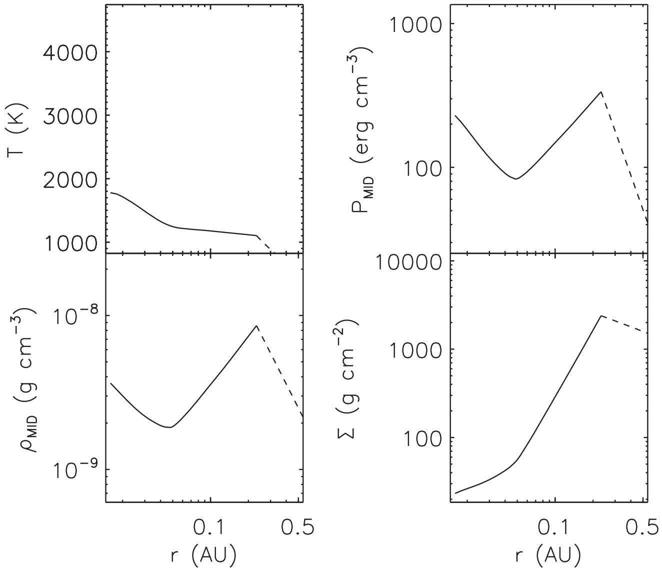


Figure 9. Various disk structure quantities plotted as a function of radius, for our fiducial disk model. Top left: (vertically constant) temperature. Top right and bottom left: midplane pressure and midplane density. Bottom right: surface density. Solid lines represent our model calculations, which end at the radius where $\bar{\alpha}$ falls to $\bar{\alpha}_{\mathrm{DZ}}$. Beyond this radius we assume a constant $\bar{\alpha} = \bar{\alpha}_{\mathrm{DZ}}$, obtaining the results shown here by the dashed lines.

Nevertheless, other elements with slightly higher ionization potentials may plausibly matter because their abundances are much higher. To check this, we carried out calculations for our fiducial model with sodium instead, which has an ionization potential (5.14 eV) only slightly larger than potassium's but is ~16 times more abundant. We find (not plotted) that, while the greater abundance of Na yields a significantly higher ionization fraction in regions where our original simulations showed K to already be highly ionized (in surface layers, and near the midplane close to the disk inner edge), the pressure maximum occurs slightly inward of its position with K; i.e., the latter is still set by the difference in ionization potentials. As such, while the precise shape of the active, dead, and zombie zones will vary somewhat when other atomic species are included with K, we do not expect the position of the pressure maximum to shift substantially. Implementing more complex chemical networks (with additional atomic and molecular species and grains) *will* be important for increasing the recombination rate and ensuring that we are in the strongly coupled limit (see Section 8.1.8 further below); we shall tackle this in an upcoming paper.

Importance of Dust: Dust grains affect both the opacity of the disk and the efficiency of the MRI. While our calculations are dust-free—in the sense that grain effects on the MRI are ignored—we have nonetheless assumed a constant opacity of $10\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, which is a reasonable value for the warm inner regions of dusty protoplanetary disks (see Hu et al. 2017). Concurrently, an a posteriori calculation of the opacities in our disk solution, using detailed opacity tables including grains, yields values very close to our assumed constant in all regions of interest except very close to the disk inner edge (see Section 8.1.7 further below). As such, grains are effectively included in our opacities, and treating them more precisely via opacity tables should not alter our results appreciably.

Inclusion of dust is very likely to be important for the MRI, however. Grains can drastically suppress the MRI, by soaking up electrons and thereby reducing the amount of negative charge tied to the magnetic field (since all but the very smallest grains [see below] are collisionally decoupled from the field themselves; e.g., Perez-Becker & Chiang 2011a; Bai 2013; Mohanty et al. 2013). Enhanced recombination on the charged grain surfaces also removes positive charge from the gas, further hampering the MRI. Lastly, MRI damping is exacerbated by the incorporation of the alkali atoms (which are the primary charge suppliers) into grains and their adsorption onto grain surfaces; we have currently ignored this effect, which can deplete metal abundances by an order of magnitude or more (e.g., Jenkins 2009; Keith & Wardle 2014). Concurrently, as Figure 9 shows, the disk temperatures in our solution are well below the dust sublimation temperature of $\sim$1500 K (at the extant densities) at radii $\gtrsim$0.05 au; as such, the pressure maximum and the DZIB in our current solution sit squarely within the radial range where dust is thermodynamically allowed. Moreover, while the pressure maximum traps relatively large grains ("pebbles")—the whole reason for invoking it for planet formation—smaller ones are increasingly well coupled to the gas and can thus flow through the trap; furthermore, it is these small grains that have the greatest impact on the MRI (because of their large collective surface area for electron adsorption). Therefore, we expect small grains to exist in our solution space, damping the MRI to some extent and moving both the DZIB and the pressure maximum radially inward of our currently predicted locations.

The *magnitude* of this effect depends, on the one hand, on the relative abundance of grains versus electrons. For dust grains with number density n_d and a fixed radius a, the grain abundance $x_d \equiv n_d/n_{\mathrm{H}_2}$ may be expressed as $x_d = (3R\mu m_{\mathrm{H}})/(4\pi\,\rho_{\mathrm{gr}}\,a^3)$, where R is the dust-to-gas ratio by mass and $\rho_{\mathrm{gr}} \approx 3\ \mathrm{g\,cm}^{-3}$ is the density of a single grain. For a standard ISM value of $R = 10^{-2}$, very small grains of size $a = 0.1\ \mu$m then imply $x_d \approx 3 \times 10^{-12}$: $\sim$10–30 times smaller than the ionization fraction $x_e \sim$ few $\times\ 10^{-11}$ to 10^{-10} that we infer over most of the active zone (both close to the midplane, at radii $\gtrsim$0.05 au, and higher up, at $\sim 1z_H$–$2z_H$, once a dead zone forms in the midplane; see Figure 4). Such grains will therefore put a significant dent in the number density of free electrons, and thus affect the MRI activity, if the adsorbed negative charge per grain is of order -10. Slightly smaller grains, $a = 0.03\ \mu$m, imply $x_d = 10^{-10} \gtrsim x_e$ and so will have a severe impact on the MRI even with $\lesssim$1 electron adsorbed per grain on average. Such grain sizes and charging are not unrealistic in disks (e.g., Perez-Becker & Chiang 2011a). We note that this calculation assumes that *all* the dust is sequestered in grains of a single size; a more realistic grain size distribution will reduce the effective dust-to-gas ratio in small grains and thus decrease x_d. This is plausibly a small correction, though, since the grain number density is likely to be dominated by the smallest particles (e.g., standard MRN distribution: $n_a \propto a^{-3.5}$; but see Birnstiel et al. 2011).

Furthermore, we have compared grain abundances here to the electron abundance derived assuming no depletion of potassium in the gas phase. If a sizable fraction of K is sequestered in dust instead (both by inclusion in molecules that make up dust grains and by the adsorption of neutral K atoms onto grains), then the x_e due to thermal ionization will be much smaller than we have inferred to start with, further reducing the MRI (though this effect will be tempered somewhat by ion and thermionic emissions, whereby neutral K collisions with grains *produce* free K^+ ions and/or electrons; see Desch & Turner 2015).

On the other hand, MRI damping by grains is mitigated to the extent that they are tied to the field (and thus act like ions), instead of being knocked off by collisions with neutrals. The Hall parameter β_j (Equation (17)) is a measure of the strength of the field coupling for any species j; noting that grains are much more massive than neutral gas particles, the relative coupling strength for grains versus ions is thus $\beta_{\mathrm{gr}}/\beta_i = (|Z_{\mathrm{gr}}|\langle\sigma v\rangle_i)/(|Z_i|\langle\sigma v\rangle_{\mathrm{gr}})$. The rate coefficient for ion–neutral collisions $\langle\sigma v\rangle_i$ is given in Equation (22), while that for grain-neutral collisions is (Wardle & Ng 1999) $\langle\sigma v\rangle_{\mathrm{gr}} = \pi\,a^2\sqrt{(128k_B T_n)/(9\pi m_n)}\ \mathrm{cm}^3\,\mathrm{s}^{-1}$, where T_n is the neutral temperature, which we assume equals the gas temperature T, and $|Z_i| = 1$ in our case. We thus get $\beta_{\mathrm{gr}}/\beta_i \approx 4 \times 10^{-4}\,|Z_{\mathrm{gr}}|(T/10^3\ \mathrm{K})^{-1/2}(a/0.1\ \mu\mathrm{m})^{-2}$. Hence, at the $T \sim$1000–2000 K in our disk solution (Figure 9), the 0.03–0.1 μm grains considered above will be far more decoupled from the field than the ions, even for grain charges $|Z_{\mathrm{gr}}| \sim 10$. We conclude that the net effect of abundant very small grains will be to significantly suppress the MRI and thus shift the pressure maximum inward of where we currently find it to be.

Relevance of X-rays: Here we have only considered thermal ionization and ignored photoionization by stellar X-rays. We estimate the effect of the latter as follows. Igea & Glassgold (1999, hereafter IG99) have calculated the ionization rate ζ_{X}, due to X-rays with photon energies of a few keV and ignoring grain effects, as a function of column density. They find that, while $\zeta_{\mathrm{X}} \propto L_{\mathrm{X}}/r^2$ (where L_{X} is the stellar X-ray luminosity and r the radial distance from the star), as expected, it is also "universal," in the sense that ζ_{X} plotted as a function of (vertical) column density is independent of the precise density structure of the disk. Moreover, in the absence of grains, the ionization fraction is given simply by $x_e = \sqrt{\zeta_{\mathrm{X}}/(n_{\mathrm{H}_2}k_{ie})}$, where k_{ie} is the recombination rate coefficient for ion–electron recombinations for the relevant dominant ions (e.g., see expressions for x_e in various limiting cases derived by Perez-Becker & Chiang 2011a). We use these facts to scale directly from IG99's results (correcting for the fact that they supply column densities in terms of hydrogen nuclei while we use hydrogen molecules instead).

The column density in our active region close to the midplane, at a mean radial distance $r \sim 0.05$ au, is $N_{H_2} \sim 3 \times 10^{24}\,\mathrm{cm}^{-2}$, while in the active region above the dead zone, at a mean $r \sim 0.1$ au, it is $N_{H_2} \sim 10^{25}\,\mathrm{cm}^{-2}$ (see Figure 5). At the same active region locations, we also have $x_e \sim 3 \times 10^{-10}$ and 10^{-10}, respectively, due to thermal ionization, and $n_{H_2} \sim 10^{14}\,\mathrm{cm}^{-3}$ (Figure 4). Concurrently, at 1 au, for $L_X = 10^{29}\,\mathrm{erg\,s}^{-1}$ and photon energies of 5 keV, IG99's Figure 5 implies $\zeta_X \sim 3 \times 10^{-17}\,\mathrm{s}^{-1}$ and $3 \times 10^{-18}\,\mathrm{s}^{-1}$ at $N_{H_2} \sim 3 \times 10^{24}\,\mathrm{cm}^{-2}$ and $10^{25}\,\mathrm{cm}^{-2}$, respectively (results for 3 and 8 keV photons are only marginally different). Assuming, as IG99 do, that molecular ions, specifically HCO$^+$, are dominant, and thus using a dissociative recombination rate coefficient of $k_{\mathrm{HCO^+},e} = 2.4 \times 10^{-7}/(T/300\,\mathrm{K})^{0.69}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$ (Woodall et al. 2007; Perez-Becker & Chiang 2011a) and scaling to our radii of interest, where $T \sim 10^3$ K, we then find that X-ray ionization implies $x_e \sim 3 \times 10^{-11}$ in our active region at 0.05 au and $x_e \sim 5 \times 10^{-12}$ in the active region at 0.1 au; these are roughly an order of magnitude smaller than x_e from thermal ionization cited above. We note that Bai & Goodman (2009) provide an analytic fit to IG99's ζ_X curves; we get the same results using their fitting formula.

However, while X-rays first produce molecular ions, charge transfer to metals is so rapid that it is metal ions that constitute the dominant ionic species, *if* the metal abundance is high (as it is in our non-depleted grainless case;[6] e.g., Fujii et al. 2011; Keith & Wardle 2014). In that case, in the absence of grains, it is the metal ion (M$^+$)–electron recombination rate coefficient, $k_{\mathrm{M^+},e} = 2.8 \times 10^{-12}/(T/300\,\mathrm{K})^{0.86}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$ (see Section 4.2), that must be used to calculate the X-ray-driven x_e. At the relevant temperatures $T \sim 10^3$ K, we see that $k_{\mathrm{M^+},e} \approx 10^{-5} \times k_{\mathrm{HCO^+},e}$ (i.e., metal ions recombine vastly slower than molecular ones); consequently, the x_e due to X-rays in our metal-abundant active regions will be more than 2 orders of magnitude higher than inferred above using HCO+, completely swamping the x_e from thermal ionization. Of course, metals may be severely depleted when grains are present; however, this will decrease the thermal ionization fraction too, so we expect X-rays to remain highly competitive with thermal ionization in activating the MRI in the inner disk.

Note, however, that once a dead zone forms in the midplane, the midplane column density quickly exceeds that in the overlying active zone by more than an order of magnitude (Figure 5). IG99's results then imply an X-ray-induced midplane x_e at least 3 orders of magnitude smaller than that deduced from X-rays in the active zone, and much smaller than the midplane x_e from thermal ionization. As such, X-ray ionization will not change our result that a dead zone eventually forms in the midplane and the active zone climbs up above it. However, by enhancing the ionization in the overlying active zone, and thus increasing the effective viscosity parameter $\bar{\alpha}$, X-rays will alter the location of the pressure maximum. These effects will be quantified in our upcoming work including X-rays (Jankovic et al. 2018, in preparation).

Finally, we point out that, in past work, X-ray ionization has widely been stated to be unimportant in the inner disk, with thermal ionization of alkali metals being the dominant process instead. Why then do we find X-rays to be at least as important as thermal collisions? The reason is that previous studies have drawn their conclusions based on the assumption of a surface density distribution that monotonically increases radially inward (e.g., the MMSN; Igea & Glassgold 1999). In that case, the surface density in the innermost regions is indeed too high for X-rays to penetrate to any significant depth in the disk. Here, however, we examine a posteriori the degree of X-ray ionization in our steady-state disk solution,[7] in which the surface density Σ is considerably *lower* inward of the pressure maximum (see Figure 9 further below). Such a turnover in the radial Σ profile is in fact a generic feature of steady-state models that invoke a radially changing α-viscosity to produce a pressure maximum in the disk (because the higher viscosity inward of the pressure maximum requires a lower Σ to drive a given $\dot{M}$, by Equation (27); e.g., see solutions by Kretke & Lin 2007, 2010). The severely depressed surface density in the inner disk then allows much greater X-ray penetration and ionization. Therefore, if protoplanetary disks *start* with a standard monotonic $\Sigma(r)$ profile, we conjecture that they will evolve as follows: initially, thermal ionization will dominate in the inner disk, driving it toward the steady-state solution we find, and thereby reducing the surface density in these regions; once the Σ here falls sufficiently (i.e., column densities drop to $\sim 10^{25}$–$10^{24}\,\mathrm{cm}^{-2}$), X-ray ionization will begin to complement, and perhaps overtake, the ionization due to thermal collisions, enhancing the MRI and thus the effective α. As argued above, we do not expect this to alter the qualitative features of our disk solution, but we do expect the precise locations of the DZIB and the pressure maximum to change from our current results.

Hall Effect: Here we have neglected the effects of Hall resistivity on the MRI. This does not prevent us, though, from calculating the Hall Elsässer number χ everywhere within our solution disk. The results are shown in Figure 5, where the cross-hatched region denotes the Hall zone, i.e., where $\chi < 1$, and hence where the Hall effect *is* important. We see that the Hall zone essentially overlaps with the Ohmic dead zone and also extends into the overlying active zone at radii $\gtrsim 0.15$ au. Thus, if the net vertical field is anti-aligned with the disk spin axis, we do not expect our solution to change very much: in this field configuration, the Hall effect damps magnetically driven radial angular momentum transport, so the active zone will end at (and the pressure maximum will thus be located at) ~0.15 au instead of ~0.25 au, while the dead zone (where Ohmic resistivity already quenches the MRI) will remain dead. If the net vertical field is aligned with the spin axis, on the other hand, the HSI can activate magnetically driven radial transport within the entire dead zone.

This suggests an explanation for the fact that close-in Earths/super-Earths are *not* seen around ~50% of stars. In general, one expects a net vertical background magnetic field threading the disk, due to either the stellar field or an external interstellar field. Moreover, one expects the alignment/anti-alignment of this field to be random relative to the disk angular momentum vector, with a roughly equal distribution of either geometry. Thus, in roughly half the systems, alignment between the field and disk spin axis should lead to the HSI activating the dead zone, which will remove the pressure barrier and thus suppress the formation of close-in small planets; in the other half of systems, anti-alignment will damp the HSI, allow the pressure barrier to form, and thus promote the formation of such planets.

---

[6] Note that it is the metal *abundance*, and not the ionization potential, that is the controlling factor here (because the keV X-ray energies greatly exceed the electron binding energies in the metals). As such, the relevant metal here is magnesium (with attendant ions Mg$^+$), and not potassium as in our thermal ionization calculations, since Mg is far more abundant than K: $x_{\mathrm{Mg}}/x_{\mathrm{K}} \approx 4 \times 10^2$ (e.g., Keith & Wardle 2014).

[7] Where the solution has been derived using the standard assumption of thermal ionization alone.

We shall address this mechanism quantitatively in future work; we only note here that our result—that $\chi < 1$ within the Ohmic dead zone—is in qualitative agreement with that of Bai (2017), who finds that the Hall effect is critical within the classical Ohmic dead zone (albeit at much larger radii than in our solutions).

8.1.3. $\bar{\alpha}(r)$

Figure 8 shows our solution for the vertically averaged viscosity parameter $\bar{\alpha}$ as a function of radius. In the innermost disk, $\bar{\alpha}$ saturates at ∼0.08 as the potassium becomes almost entirely ionized (see top panel of Figure 7). It then falls smoothly by nearly 3 orders of magnitude, reaching our adopted floor value of $\bar{\alpha}_{\rm DZ} = 10^{-4}$ at ∼0.25 au. Beyond this point, there is no MRI-active zone any more, and we assume a constant $\bar{\alpha} = \bar{\alpha}_{\rm DZ}$ (depicted by the dashed horizontal line in Figure 8).

8.1.4. Disk Structure and Pressure Maximum

Figure 9 shows the (vertically isothermal) temperature, midplane density, midplane pressure, and surface density as functions of radius for our fiducial disk model. Beyond ∼0.25 au, where $\bar{\alpha}$ falls to $\bar{\alpha}_{\rm DZ}$, we calculate these quantities assuming a constant $\bar{\alpha} = \bar{\alpha}_{\rm DZ}$ (as depicted by the dashed lines in Figure 9).

The salient results are as follows: (a) There is a clear maximum in the midplane gas pressure (and midplane gas density) at ∼0.25 au, where $\bar{\alpha}$ reaches its floor value of $\bar{\alpha}_{\rm DZ}$. Note that this location is radially well beyond the DZIB, which is located at ∼0.09 au; thus, the midplane pressure maximum is situated *within* the dead zone, for the reasons discussed earlier. (b) The surface density declines sharply inward of the pressure maximum, falling by 2 orders of magnitude toward the disk inner edge. This is a straightforward consequence of $\bar{\alpha}$ increasing inward in this region coupled with a constant $\dot{M}$, as discussed previously. (c) The temperature varies quite slowly in the inner disk in this fiducial model, by less than a factor of 2, and in particular remains lower than the dust sublimation temperature of ∼1500 K except near the disk inner edge. As such, small dust grains (which will be coupled to the gas rather than being trapped in the pressure maximum) are expected to have a significant effect on the MRI in this region, which we examine in a subsequent paper.

8.1.5. Accretion Rates in Active, Dead, and Zombie Zones

The total inward accretion rate (which by definition is radially constant in our steady-state solutions) is, at every radius, the sum of the accretion rates within the individual vertical layers of the disk (active, dead, and zombie). We calculate these individual $\dot{M}$ using Equation (26); the results are plotted in Figure 10. We see that the inward $\dot{M}$ through the active layer is practically the sole contributor to the total from the innermost radii out to ∼0.09 au, where the dead zone in the midplane first develops; the $\dot{M}$ through the overlying zombie zone (due to non-MRI torques) steadily increases over this radial span but is negligible compared to the active zone value. Once the Ohmic dead zone forms, the inward accretion through it (again, due to non-MRI torques) rapidly increases (as the thickness of this layer grows), while the $\dot{M}$ in the active and zombie zones correspondingly decrease. Indeed, beyond ∼0.15 au, the inward $\dot{M}$ in the dead zone *exceeds* the total value; this is compensated for by *decretion* (outward flow of mass) in the active and zombie zones, which ensures that the total inward accretion rate remains constant at the desired value ($10^{-9}\,M_\odot\,{\rm yr}^{-1}$ here).

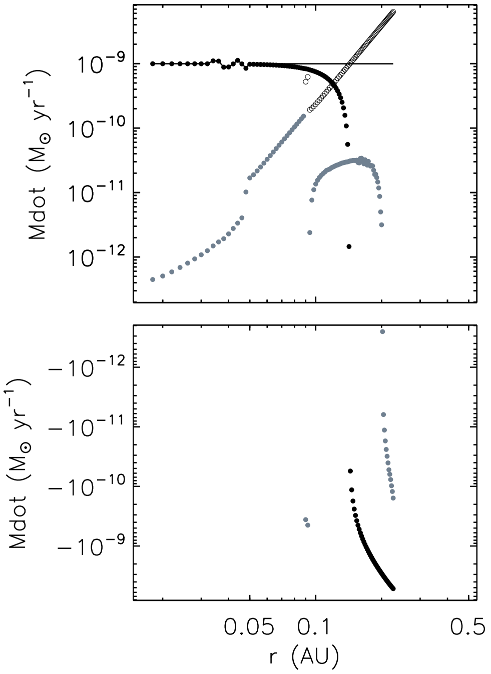


Figure 10. Accretion rates through the active zone (filled black circles), zombie zone (filled gray circles), and dead zone (open circles) for our fiducial model. The solid black line represents the sum of the three rates (i.e., the total accretion rate through the disk, held fixed at $\dot{M} = 10^{-9}\,M_\odot\,{\rm yr}^{-1}$ in this model). The top panel shows inward (positive) accretion rates, and the bottom panel shows outward (negative) rates. A few small anomalies—the minor jitter in the active zone (and thus total) rate around 0.04 au, and the anomalously large first two points in the dead zone accretion rate, at 0.09 au—result from our finite grid resolution at locations where the disk resistivities undergo sharp changes (we have left them in to show the limits of our precision). See Section 8.1.5.

A little reflection shows that in a nontrivial and nonpathological disk, i.e., one in which the disk properties vary radially in a physically plausible manner, such inconstancy of the accretion rates within the individual layers is unavoidable if the total $\dot{M}$ is to remain fixed: if we demand that the total value be invariant, then we do not have any separate justifiable knobs to turn to ensure that the individual contributing rates remain constant as well.

Does this phenomenon represent a growing instability? Certainly the buildup of mass at some locations, and excavation at others, that the radially varying accretion rates will generate in the individual layers will tend to drive the disk away from our equilibrium solution. However, these changes in the vertical density profile will occur over a local viscous timescale, given by $t_{\rm visc} \sim r^2/\bar{\nu}$ (where $\bar{\nu}$ is the vertically averaged local viscosity). By Equations (1) and (2) $\bar{\nu} = \bar{\alpha} c_s^2/\Omega$ in our vertically isothermal disk, so $t_{\rm visc} \sim \bar{\alpha}^{-1}(z_H/r)^{-2}\,\Omega^{-1} \sim \bar{\alpha}^{-1}(z_H/r)^{-2}\,t_{\rm dyn}$, where $t_{\rm dyn} \sim 1/\Omega$ is the dynamical timescale. Simultaneously, the disk will tend to relax back to a hydrostatic equilibrium vertical profile (which is assumed in our solution) on a timescale given by $t_H \sim z_H/c_s \sim 1/\Omega \sim t_{\rm dyn}$. Note that the instantaneous perturbations in the vertical density profile here do *not* represent a change in the total surface density Σ at any location: the latter remains constant (by Equation (27), since the total $\dot{M}$ is fixed at our steady-state value), i.e., the density perturbations sum to zero vertically. Thus, the disk will tend to relax to the *same* hydrostatic equilibrium vertical profile as in our solution. Now, in a normal thin disk, the disk aspect ratio $z_H/r \ll 1$, so for a standard $\bar{\alpha} < 1$, we have $t_{\rm visc} \gg t_{\rm dyn}$. Figure 11 demonstrates this explicitly for our disk: we see that $t_{\rm visc}$ is orders of magnitude larger than $t_{\rm dyn}$ over our radii of interest. Consequently, we expect the density perturbations introduced by the variable accretion rates to be vertically smoothed out, and hydrostatic equilibrium re-established, much more rapidly than these perturbations can grow; our steady-state solution will then remain valid in a (dynamical) time-averaged sense.

8.1.6. Viscous Instability

In our steady-state solutions, the surface density Σ and hence the accretion rate are temporally constant. Perturbations in Σ, however, may lead to a viscous instability as follows (see Pringle 1981). The general evolution equation for the disk surface density is

$$\frac{\partial \Sigma}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}\left[\left(\frac{dj}{dr}\right)^{-1}\frac{\partial}{\partial r}\left(r^3\,\bar{\nu}\,\Sigma\,\frac{d\Omega}{dr}\right)\right], \tag{28}$$

where j is the specific angular momentum at any disk location, and $\bar{\nu}$ is again the vertically averaged viscosity. For the specific case of a Keplerian disk, we have $\Omega = \sqrt{GM_*/r^3}$ and $j = r^2\Omega$, and the above reduces to

$$\frac{\partial \Sigma}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[3\,r^{1/2}\frac{\partial}{\partial r}(r^{1/2}\,\bar{\nu}\,\Sigma)\right]. \tag{29}$$

Changes in Σ will occur on a viscous timescale. We have already noted that vertical hydrostatic equilibrium is established over a timescale $t_{\rm dyn} \ll t_{\rm visc}$. Similarly, the disk will relax to thermal equilibrium over a time given by the ratio of the thermal energy content per unit area to the rate of viscous heating (=rate of cooling in equilibrium) per unit area: $t_{\rm th} \sim (P_{\rm gas}\,z_H)/(\bar{\nu}\Sigma\,\Omega^2) \sim c_s^2/(\bar{\nu}\,\Omega^2) \sim \bar{\alpha}^{-1}\,t_{\rm dyn}$. Thus, for $\bar{\alpha} < 1$, we have $t_{\rm dyn} < t_{\rm th} \ll t_{\rm visc}$ (as Figure 11 explicitly shows for our disk), and we expect the disk to be in both thermal and hydrostatic equilibrium over the timescales on which Σ varies. In this situation, the mean viscosity at a fixed radius will depend only on the local surface density, i.e., $\bar{\nu} = \bar{\nu}(\Sigma, r)$, and Equation (29) is a nonlinear diffusion equation for Σ. For steady-state solutions, the left-hand side of Equation (29) is zero; we wish to investigate the effect of a small perturbation about any such equilibrium solution Σ_0. Define $x \equiv \bar{\nu}\,\Sigma$. Then any small variation in the surface density, $\Sigma_0 \to \Sigma_0 + \delta\Sigma$, implies a variation $x_0 \to x_0 + \delta x$, with $\delta x = (\partial x/\partial\Sigma)\delta\Sigma$. Inserting the perturbed value of Σ into Equation (29) then gives the time evolution equation for the perturbation δx:

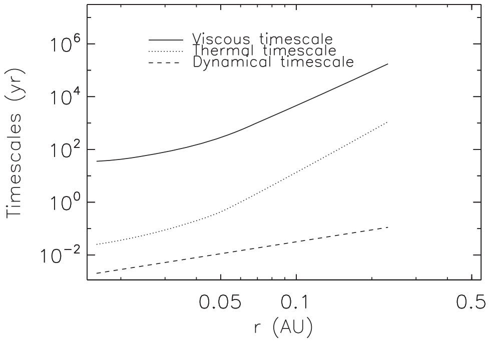


Figure 11. Viscous, thermal, and dynamical (orbital) timescales as a function of radius for our fiducial disk model. See Sections 8.1.5 and 8.1.6.

$$\frac{\partial(\delta x)}{\partial t} = \left(\frac{\partial x}{\partial \Sigma}\right)\frac{1}{r}\frac{\partial}{\partial r}\left[3\,r^{1/2}\frac{\partial}{\partial r}(r^{1/2}\,\delta x)\right]. \tag{30}$$

This linear diffusion equation for $\delta\, x$ is well behaved *if and only if* the diffusion constant $\partial x/\partial\Sigma$ is positive; instability results otherwise. Hence, using $\bar{\nu} = \bar{\alpha}c_s^2/\Omega$ in our disk to evaluate the diffusion constant, we arrive at the viscous instability condition:

$$\begin{aligned}\text{Instability} &\iff \frac{\partial x}{\partial \Sigma} < 0\\ &\iff \frac{\partial(\ln\bar{\alpha})}{\partial(\ln\Sigma)} + 2\,\frac{\partial(\ln c_s)}{\partial(\ln\Sigma)} < -1.\end{aligned} \tag{31}$$

A negative diffusion constant implies that surface density inhomogeneities will be amplified: overdense regions will grow denser, while underdense ones will become even more rarefied. In other words, an axisymmetric disk will tend to break up into rings.

To investigate whether our inner disk is viscously unstable, we proceed as follows. We *assume* that, given a local perturbation in surface density Σ, the local disk parameters ($\bar{\alpha}$, c_s, x, $\dot{M}$) tend toward their *steady-state* values corresponding to the *perturbed* value of Σ. This allows us to evaluate the instability criterion by comparing the different equilibrium solutions we have calculated. We also find it useful to change variables from x to $\dot{M}$, in order to connect to our steady-state solutions for different values of $\dot{M}$.

In general, $\partial\Sigma/\partial t = (2\pi r)^{-1}\partial\dot{M}/\partial r$. For steady state, $\dot{M}$ must be radially constant; in this case, combining the latter expression with Equation (29) yields the equilibrium solution $\dot{M}_0 = 3\pi\,\bar{\nu}\,\Sigma/f_r$ (equivalent to Equation (27) with our definition of $\bar{\nu}$). Thus, $\dot{M}_0 \propto x_0$ (with the constant of proportionality independent of Σ), and the instability condition $\partial x/\partial\Sigma < 0$ may be expressed as $\partial\dot{M}/\partial\Sigma < 0$, or equivalently as $\partial\Sigma/\partial\dot{M} < 0$. Evaluating the latter expression, we can write the instability criterion as

$$\begin{aligned}\text{Instability} &\iff \frac{\partial \Sigma}{\partial \dot{M}} < 0\\ &\iff \frac{\partial(\ln\bar{\alpha})}{\partial(\ln\dot{M})} + 2\,\frac{\partial(\ln c_s)}{\partial(\ln\dot{M})} > 1.\end{aligned} \tag{32}$$

In Figure 12, we plot the steady-state Σ solution as a function of radius, for various $\dot{M}$ spanning $\pm$ 0.3 dex around our fiducial value of $10^{-9}\,M_\odot\,{\rm yr}^{-1}$. We immediately see that, at

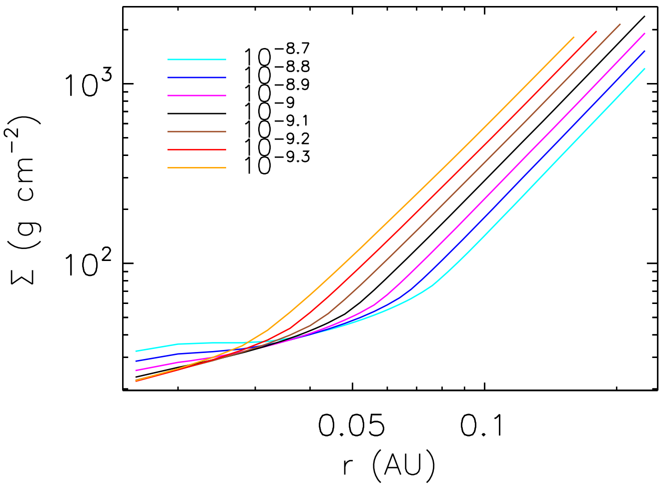


Figure 12. Steady-state solutions for the surface density Σ as a function of radius, for model parameters $M_* = 1\,M_\odot$, $\bar{\alpha}_{\rm DZ} = 10^{-4}$, and varying accretion rates $\dot{M} = 10^{-9.3}$–$10^{-8.7}\,M_\odot\,{\rm yr}^{-1}$ in steps of 0.1 dex. Over most of the disk (except the innermost regions), the surface density increases with decreasing accretion rate. See Section 8.1.6.

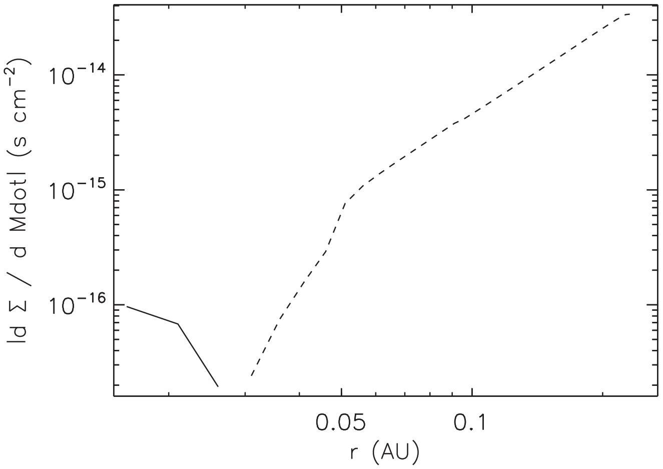


Figure 13. $|\partial\Sigma/\partial\dot{M}|$ as a function of radius for our fiducial disk model. The solid line denotes $\partial\Sigma/\partial\dot{M} > 0$ (viscously stable), while the dashed line denotes $\partial\Sigma/\partial\dot{M} < 0$ (viscously unstable). The disk is thus unstable at radii $r \gtrsim 0.03$ au. See Section 8.1.6.

any fixed radius beyond ~0.035 au, Σ increases as $\dot{M}$ decreases, i.e., $\partial\Sigma/\partial\dot{M} < 0$. Thus, *most of the disk is viscously unstable*. This is shown more explicitly in Figure 13, where we plot $\partial\Sigma/\partial\dot{M}$ (calculated by deriving the steady-state Σ for $\dot{M} = 10^{-9}\,M_\odot\,{\rm yr}^{-1} \pm 1\%$) against radius; the quantity is negative over all but the innermost disk regions. By Equation (32), the instability criterion may also be expressed as a condition on the summed change in $\bar{\alpha}$ and c_s^2 as a function of the change in $\dot{M}$. In Figure 14, we plot each of these two terms separately. It is apparent that the instability is caused primarily by the large change in $\bar{\alpha}$ with $\dot{M}$, with the change in sound speed making only a minor contribution. We shall see explicitly how $\bar{\alpha}$ changes with accretion rate in Section 8.3.

8.1.7. Opacity

In this work, we have *assumed* a constant opacity of $10\,{\rm cm}^2\,{\rm g}^{-1}$ throughout our calculation domain. Given the pressure and temperature structure derived thereby for our solution disk, we check the validity of this assumption a posteriori, by using the detailed tables of Zhu et al. (2012) to compute the opacities *predicted* as a function of pressure and temperature.

The results are plotted in Figure 15. We see that the predicted opacity over the bulk of our disk solution is 5–10 $\rm cm^2\,g^{-1}$ (primarily due to grains; see below), very close to our assumed value. The only exception is the innermost disk, at $\lesssim$0.03 au, where the expected opacities are 1–2 orders of magnitude lower (as grains disappear). However, this small inner region is not consequential to our results at larger radii, e.g., regarding the

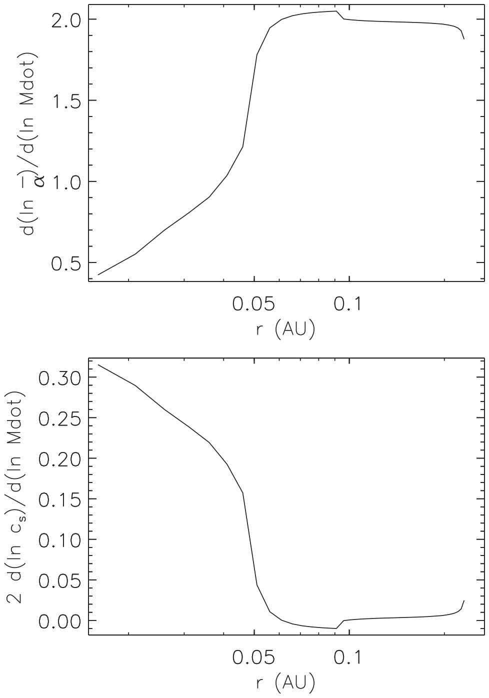


Figure 14. $\partial(\ln\bar{\alpha})/\partial(\ln\dot{M})$ (top) and $2\partial(\ln c_s)/\partial(\ln\dot{M})$ (bottom) as a function of radius, for our fiducial disk model. See Section 8.1.6.

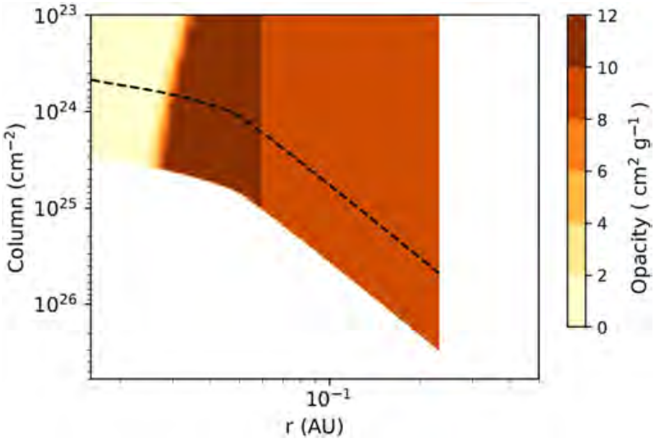


Figure 15. Rosseland mean opacity (in $\rm cm^2\,g^{-1}$) calculated a posteriori for our fiducial disk model, plotted as a function of disk location (with height in units of vertical column). The dashed curve denotes one pressure scale height. Over most of our region of interest in the disk, the derived opacity is within a factor of two of 10 $\rm cm^2\,g^{-1}$, consistent with our a priori adoption of this value everywhere. See Section 8.1.7.

DZIB and the pressure maximum. In summary, therefore, our disk solution is overall self-consistent vis-à-vis the adopted opacity.

Note that we have not explicitly included grains in our calculations. Nevertheless, our assumed opacity of $10\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ is the fiducial value adopted widely for dusty accretion disks and is validated over most of the disk by the opacity calculations above that do account for grains. In other words, grains are *implicitly* included in our opacities. On the other hand, dust will also markedly influence the chemistry and the MRI (see Section 8.1.2); these grain effects are ignored in this work (we treat them in a subsequent paper; Jankovic et al. 2018, in preparation).

8.1.8. Validity of the Strong-coupling Limit

The criteria we use for active MRI in the presence of ambipolar diffusion (Equations 11(a), (b)), derived from the MRI simulations of Bai & Stone (2011), require that we be in the strong-coupling limit, i.e., in the single-fluid regime. The conditions for the latter are (see Appendix A) (1) $\rho_n \gg \rho_i$ (which is always satisfied in our case wherein potassium is the only ionized species, since the abundance of K puts a hard upper limit of $\sim 10^{-7} \times m_K/m_{\mathrm{H}_2} \ll 1$ on ρ_i/ρ_n) and (2) $t_{\mathrm{rcb}} \ll t_{\mathrm{dyn}}$, where t_{rcb} is the recombination timescale. The latter condition expresses the requirement that ionization–recombination equilibrium be established on timescales shorter than the dynamical time on which other relevant disk physics (such as field amplification by Keplerian shear) occurs. Since ionization is generally very fast, it is the recombination time that forms the bottleneck in establishing ionization equilibrium, hence the criterion $t_{\mathrm{rcb}} \ll t_{\mathrm{dyn}}$. If this is not satisfied, then the MRI simulation results do not represent a steady state.

We use Equation (13) to calculate t_{rcb} everywhere in our solution disk, and we compare it to the local t_{dyn}; the results are shown in Figure 16. We find that in fact the required condition on t_{rcb} is met *only* in the innermost disk close to the midplane, and nowhere else. The reason is clear: with effectively only a single chemical species (K), there is only one, relatively slow, recombination channel; thus, t_{rcb} ($\propto \sqrt{T}/n_e$) only becomes small enough to fall below t_{dyn} at the smallest radii, where n_e is highest (see Figure 4; the weaker dependence on T, combined with the relatively small variation in T in our solution—see Figure 9—means that the temperature does not alter t_{rcb} very much). As such, our disk solution is to be interpreted only as an idealized case that holds *if* ionization equilibrium is established with a single alkali species. Whether such an equilibrium can indeed be reached, or maintained, when the disk and field are otherwise evolving on much shorter dynamical timescales is unclear.[8]

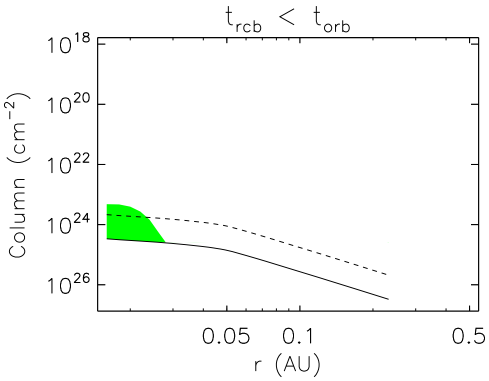


Figure 16. Recombination timescale (t_{rcb}) vs. dynamical timescale (t_{dyn}) as a function of location in our fiducial disk model. The solid black curve denotes the disk midplane, and the dashed curve denotes one pressure scale height. The green region is where $t_{\mathrm{rcb}} < t_{\mathrm{dyn}}$ (and thus where the single-fluid approximation is valid); in the rest of our fiducial disk, $t_{\mathrm{rcb}} > t_{\mathrm{dyn}}$. See Section 8.1.8.

Nevertheless, our disk solutions are useful for two reasons. First, actual disks should support far more complex chemical networks, including both molecular ions and grains in addition to metal ions. With the much larger number of recombination channels available in such physically realistic circumstances, we do expect the time to attain ionization equilibrium to usually be shorter than the dynamical one (e.g., Bai 2011). In that case, as long as $\bar{\alpha}$ follows the general form in our solutions (high value at very small radii, and tapering off with increasing distance), our results, regarding the behavior of the various zones and the trends in the MRI and accretion rates, should remain qualitatively applicable (though the quantitative locations of the pressure maximum and so forth will certainly change). Second, our analysis provides a general *method* for self-consistently solving the problem of an α-disk coupled to the MRI (and for checking the validity of the solution a posteriori, as done here). This methodology will remain applicable, whatever the specifics of the chemical network.

8.2. Variations in $\bar{\alpha}_{\mathrm{DZ}}$

Figures 17–19 show our disk solutions for the same M_* and $\dot{M}$ as the fiducial case, but with $\bar{\alpha}_{\mathrm{DZ}} = 10^{-3}$ and 10^{-5} (instead of 10^{-4}). These results closely resemble the fiducial solution, but with a couple of important quantitative differences.

First, because $\bar{\alpha}$ declines with increasing radius, the pressure maximum (located where $\bar{\alpha}$ hits the floor value $\bar{\alpha}_{\mathrm{DZ}}$) occurs at a smaller radius ($\sim$0.12 au) for $\bar{\alpha}_{\mathrm{DZ}} = 10^{-3}$ compared to the fiducial case ($\sim$0.25 au); conversely, it is at a larger radius ($\sim$0.4 au) for $\bar{\alpha}_{\mathrm{DZ}} = 10^{-5}$. in fact, as Figure 20 shows, the radial location of the pressure maximum as a function of the floor value $\bar{\alpha}_{\mathrm{DZ}}$ is approximately a power law: $r_{P_{\max}} \propto \alpha_{\mathrm{DZ}}^{-1/4}$. This follows from the fact that our solutions $\bar{\alpha}$ as a function of radius approximately decrease as power laws beyond 0.1 au, and the location of the pressure maximum corresponds to the radius at which this power law falls to the floor value $\bar{\alpha}_{\mathrm{DZ}}$ (right panels of Figure 18).

Second, while all three $\bar{\alpha}_{\mathrm{DZ}}$ solutions are quite similar at radii inward of the pressure maxima, they are not exactly the same: specifically, the field strengths in the three cases diverge beyond $\sim$0.09 au (left panel of Figure 18), which is where the dead zone first arises (compare Figures 5 and 17). This stems from the fact that the total $\dot{M}$ at any radius is the sum of the individual accretion rates through the active, dead, and

---

[8] Answering this question rigorously requires a general two-fluid simulation (of which the one-fluid regime is a special case), including source and sink terms for the ions in order to account for an evolving ionization fraction (we thank X. Bai for useful discussions on this issue). Note that the idealized two-fluid simulations of Hawley & Stone (1998, hereafter HS98; see Appendix A) assume a *fixed* ion fraction and so do not address this issue directly. Nonetheless, if $t_{\mathrm{rcb}} \gg t_{\mathrm{dyn}}$, then the ionization fraction may be assumed to be approximately constant over t_{dyn}, with all the relevant species being completely ionized (since the ionization timescale alone is very short). In this sense, the HS98 results may be applied to a disk like ours, with only a single alkali species, with the specification that all the alkali atoms be ionized. We cannot, however, apply the HS98 results to the ionization fractions *we* have derived assuming Saha equilibrium, because $t_{\mathrm{rcb}} \gg t_{\mathrm{dyn}}$ means that Saha equilibrium is simply not established over the dynamical timescales relevant to the HS98 simulations. At any rate, as discussed in the main text above, we do not expect a chemical network comprising only one alkali to be generally representative of real disks, so we do not pursue this line of inquiry further here.

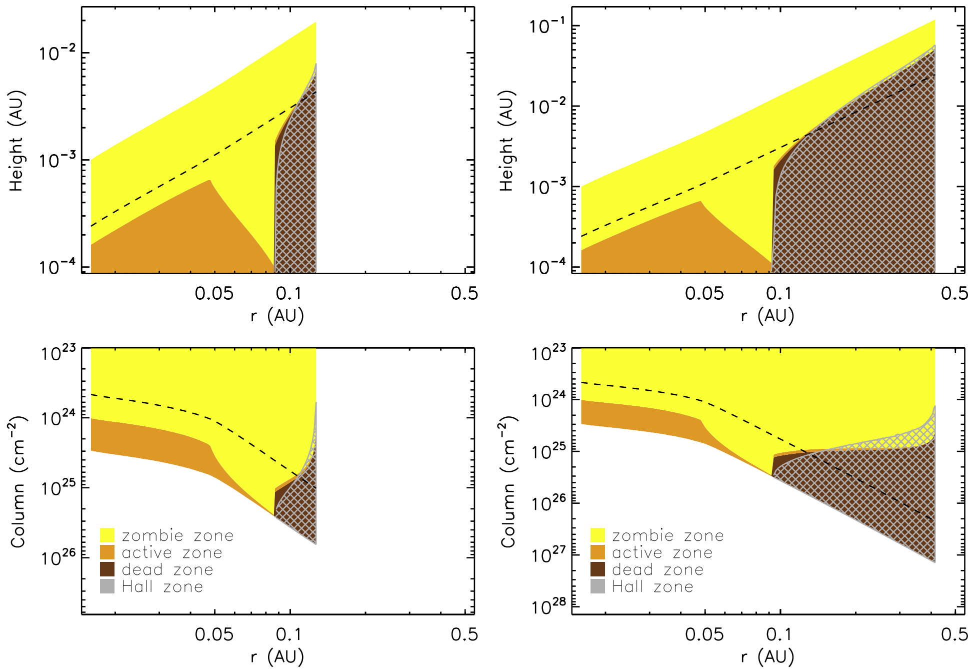


Figure 17. Same as Figure 5, but now for $\bar{\alpha}_{\rm DZ} = 10^{-3}$ (left) and 10^{-5} (right).

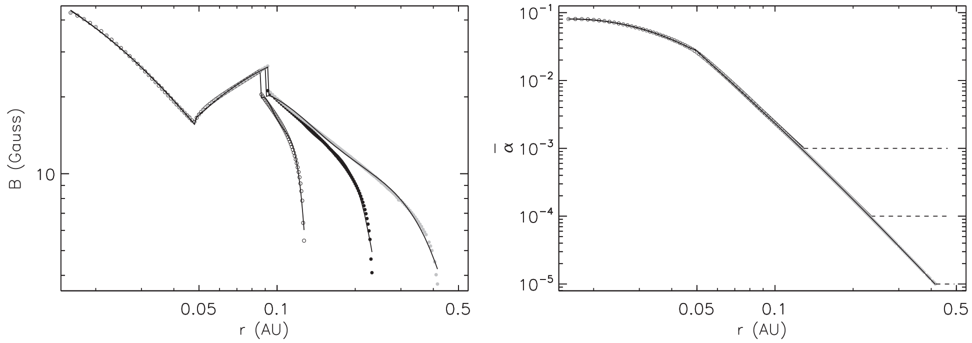


Figure 18. Left: field strength B as a function of radius. Right: $\bar{\alpha}$ as a function of radius. In both plots, results for $\bar{\alpha}_{\rm DZ} = 10^{-3}$ (open circles) and $\bar{\alpha}_{\rm DZ} = 10^{-5}$ (filled gray circles) are overplotted on the results for our fiducial model with $\bar{\alpha}_{\rm DZ} = 10^{-4}$ (filled black circles; the fiducial results are the same ones shown in Figures 6 and 8, respectively).

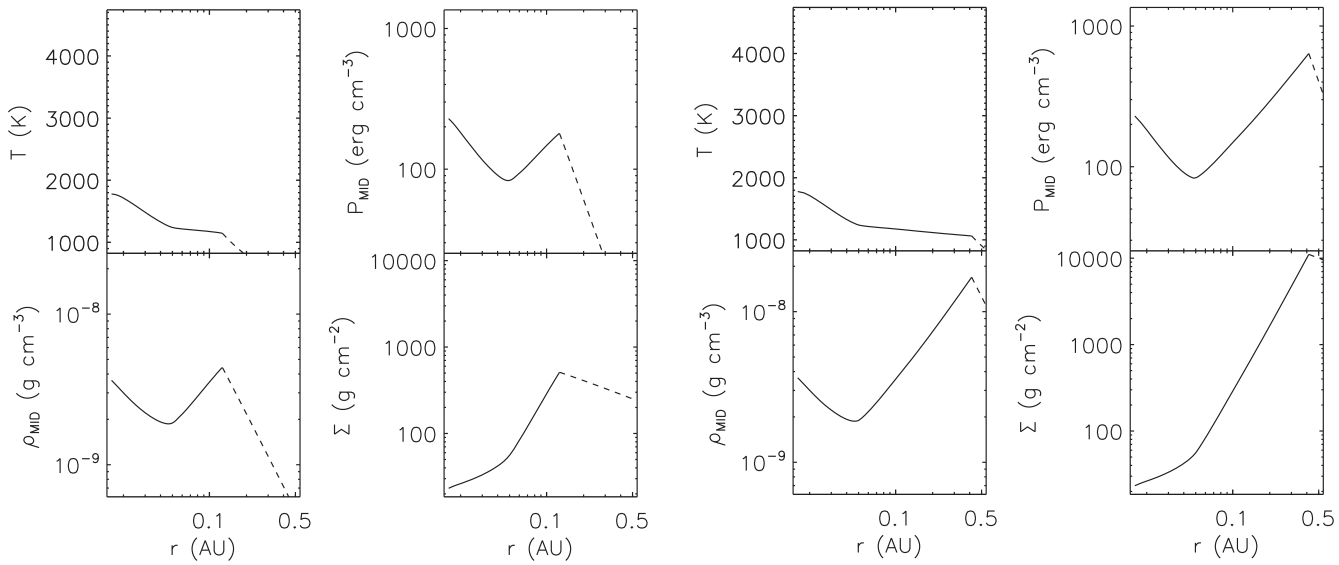


Figure 19. Various disk parameters as a function of radius: same as Figure 9, but now for $\bar{\alpha}_{\rm DZ} = 10^{-3}$ (left) and 10^{5} (right).

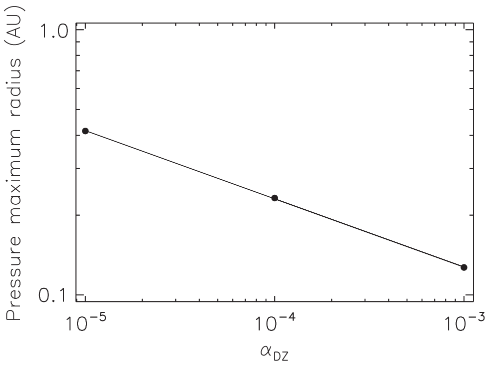


Figure 20. Radial location of the pressure maximum as a function of $\bar{\alpha}_{\rm DZ}$, showing the approximate power-law dependence $r_{P_{\rm max}} \propto \bar{\alpha}_{\rm DZ}^{-1/4}$.

zombie layers at that location. As discussed in Section 8.1.5 for the fiducial case, the accretion rate in the low-density zombie zone makes a negligible contribution to the total; hence, at radii where the active zone dominates in the midplane, the $\dot{M}$ through it (controlled by the *B*-field) is essentially constant at the fixed total rate. Thus, the field strength at these radii remains the same for the three $\bar{\alpha}_{\rm DZ}$ cases considered (which all have the same total $\dot{M}$). Once a dead zone forms in the high-density midplane, however, the accretion rate through it makes a significant and radially increasing contribution to the total rate; the rate through the active zone then compensates (in order to keep the total $\dot{M}$ fixed) by declining rapidly with radius, facilitated by a steep decrease in the field strength (see Figure 10). Since the accretion rate through the dead zone increases with $\bar{\alpha}_{\rm DZ}$, a higher (lower) $\bar{\alpha}_{\rm DZ}$ leads to a steeper (shallower) fall-off in field strength (and thus in the active zone accretion rate) with radius, as depicted in Figure 18.

8.3. Variations in $\dot{M}$

Figures 21–24 show our disk solutions for the same M_* and $\bar{\alpha}_{\rm DZ}$ as the fiducial case, but with $\dot{M} = 10^{-8}$ and 10^{-10} $M_\odot$ yr^{-1} (instead of 10^{-9} $M_\odot$ yr^{-1}). The salient deviations here from the fiducial case are all rooted in the fact that a higher $\dot{M}$ elevates the viscous heating rate, leading to a larger ionization fraction at a given location.

First (Figures 21 and 24), the pressure maximum is pushed out to $\sim$0.7 au when $\dot{M} = 10^{-8}$ $M_\odot$ yr^{-1} and in to $\sim$0.07 au when $\dot{M} = 10^{-10}$ $M_\odot$ yr^{-1}, compared to $\sim$0.25 au for the fiducial accretion rate. An increase (decrease) in ionization fraction yields a higher (lower) $\bar{\alpha}$ at a fixed radius, so the pressure maximum (achieved where $\bar{\alpha}$ falls to $\bar{\alpha}_{\rm DZ}$) occurs at a larger (smaller) radius for a given $\bar{\alpha}_{\rm DZ}$. As Figure 25 shows, the radial location of the pressure maximum as a function of the accretion rate is approximately a power law: $r_{P_{\rm max}} \propto \dot{M}^{1/2}$.

Second (Figure 21), for the higher $\dot{M} = 10^{-8}$ $M_\odot$ yr^{-1}, the inner edge of the dead zone recedes to a larger radius ($\sim$0.4 au, versus 0.09 au for the fiducial case). The inner edge of the Hall zone is pushed out as well, but not as much, resulting in this zone now intruding on the MRI-active zone. For the lower $\dot{M} = 10^{-10}$ $M_\odot$ yr^{-1}, on the other hand, the dead zone extends all the way to the disk inner edge; the active zone only occurs sandwiched between the dead and zombie zones and never extends to the midplane. Interestingly, the very low ionization fractions in this solution also allow the appearance of a region where Ohmic resistivity η_O dominates over both η_H and η_A (red sliver at the disk outer edge in Figure 22; the only time such a region appears in our solutions).

Third (Figure 23, right panel), $\bar{\alpha}$ saturates at $\sim$0.1 at small radii as the accretion rate increases to $\gtrsim 10^{-9}$ $M_\odot$ yr^{-1} (left panel of Figure 22). The saturation occurs because, at these $\dot{M}$, potassium is almost completely ionized at small radii over the entire vertical extent of the disk (e.g., see top left panel of Figure 4, which shows that, near the disk inner edge in the fiducial case, $x_e \approx (1–2) \times 10^{-7}$ from the midplane to the disk surface: very close to the maximum possible value of x_e in our disks, equal to the abundance of K, of $\sim 2 \times 10^{-7}$; for 10^{-8} $M_\odot$ yr^{-1}, x_e [not shown] is even closer to this upper limit at small radii). This explains why we found, in Section 8.1.6, that the innermost regions of our fiducial disk are viscously *stable*: this instability is mainly controlled by the change in $\bar{\alpha}$ with $\dot{M}$ (see Equation (32) and Figure 14), and this change is by definition very small when $\bar{\alpha}$ is close to saturation. Note as well that $\bar{\alpha}$ is saturated out to a much larger radius for 10^{-8} $M_\odot$ yr^{-1} compared to the fiducial case (because the ionization fraction grows with $\dot{M}$), implying that the inner disk becomes viscously stable over an increasing radial extent as the accretion rate climbs.

8.4. Variations in M_*

Figures 26–28 show our disk solutions for the same $\dot{M}$ and $\bar{\alpha}_{\rm DZ}$ as the fiducial case, but with $M_* = 0.1\,M_\odot$ (instead of $1\,M_\odot$). Note that the inner edge of the disk, assumed to lie at the stellar surface in our calculations, is also smaller in this case ($R_{\rm in} = R_* \approx 1\,R_\odot$, compared to $\sim 2.3\,R_\odot$ for the fiducial mass).

We see that the solutions for the two different stellar masses are nearly identical, except that the solutions for the lower mass are compressed radially inward by a roughly constant multiplicative factor (i.e., shifted inward by a constant additive factor, on the logarithmic radial scale in the plots). This is explained by the functional form of the fundamental parameters ρ, $P_{\rm gas}$, and T in a steady-state α-disk (Equations (4)–(6)). Specifically, the dependence of each of these parameters on the stellar mass M_* and orbital radius r can be expressed as a dependence on the combined parameter M_*/r^3 (the additional dependence on r via f_r is negligible for $r \gg R_{\rm in}$). Dependencies on $\bar{\alpha}$ and the opacity κ do not change this fact, since the latter quantities are themselves functions of ρ, $P_{\rm gas}$, and T. As such, for a fixed accretion rate, the solution at any radius r_a, for a stellar mass M_{*a}, is identical to that at radius $r_b \equiv r_a(M_{*b}/M_{*a})^{1/3}$ for a stellar mass M_{*b}.

9. Discussion and Conclusions

The IOPF mechanism depends on the presence of a midplane pressure maximum, arising initially from the change in viscosity between the MRI-active innermost disk and the adjacent dead zone. We have investigated the formation and location of this first pressure maximum by solving the coupled equations for MRI-driven viscosity with thermal ionization and an α-disk structure in steady state. We examine a range of disk accretion rates (10^{-10} to 10^{-8} $M_\odot$ yr^{-1}) and stellar masses (0.1–1 $M_\odot$). Within the dead and zombie zones, where the viscosity comes from non-MRI hydrodynamic and/or gravitational stresses, we assume a constant viscosity parameter $\bar{\alpha}_{\rm DZ}$ (which also sets a "floor" on the MRI-driven α), set to a fiducial value in the range 10^{-3} to 10^{-5}. We find the following:

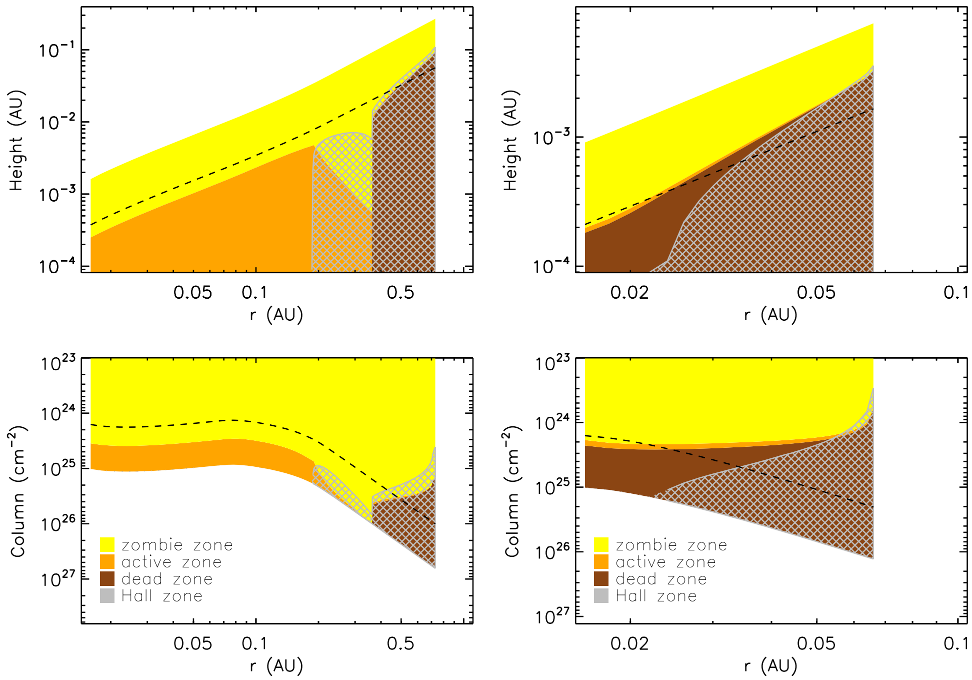


Figure 21. Same as Figure 5, but now for an accretion rate $\dot{M} = 10^{-8}\,M_{\odot}\,\mathrm{yr}^{-1}$ (left) and $10^{-10}\,M_{\odot}\,\mathrm{yr}^{-1}$ (right).

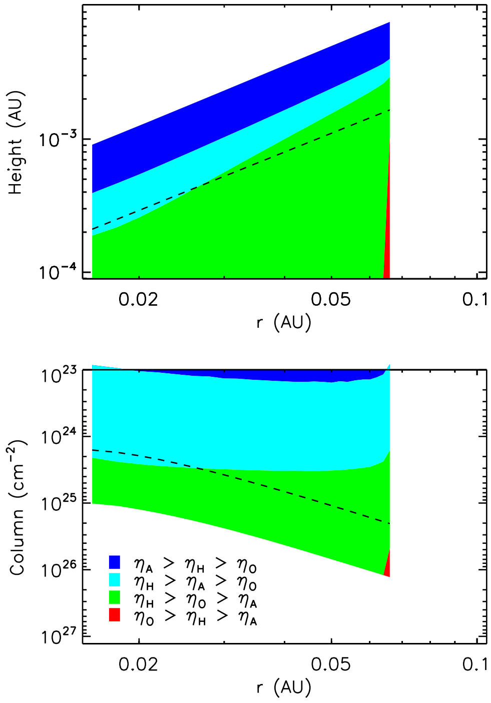


Figure 22. Same as Figure 2, but for an accretion rate $\dot{M} = 10^{-10}\,M_{\odot}\,\mathrm{yr}^{-1}$. This is the only one of our various disk models in which an Ohmic-dominated region arises (red sliver in bottom right corner of both panels).

(1) A midplane pressure maximum does form, but it is located *within* the dead zone, rather than at the DZIB as usually assumed. This is a general consequence of two factors: first, the midplane pressure does not depend on the local value of α, but rather on its vertically averaged effective value $\bar{\alpha}$; second, the MRI-active zone does not end abruptly at the DZIB, but instead continues outward above the dead zone, so that $\bar{\alpha}$ falls to its minimum value $\bar{\alpha}_{\mathrm{DZ}}$ (causing a pressure maximum) *beyond* the DZIB.

(2) The radial location of the pressure maximum has approximately power-law dependencies on α_{DZ}, stellar mass, and accretion rate: $r_{P_{\max}} \propto \alpha_{\mathrm{DZ}}^{-1/4}$, $M_*^{1/3}$, and $\dot{M}^{1/2}$.

(3) Inward of the pressure maximum, the surface density Σ in our steady-state solutions *decreases* radially inward, instead of increasing monotonically as usually assumed (e.g., in the MMSN). This is a general feature of all solutions with a constant accretion rate and an α that increases inward (since a lower Σ is required to produce the same $\dot{M}$ with a larger α). The very low Σ that results in these inner disk regions has two consequences (points 4 and 5 below).

(4) At these low Σ, Hall diffusion rather than Ohmic resistivity dominates near the midplane. Specifically, for the range of M_*, $\dot{M}$, and α_{DZ} considered here, the Hall Elsässer number $\chi < 1$ within the Ohmic dead zone in our solutions. As such, in the presence of a net vertical background field aligned with the disk spin axis, the Hall effect can reactivate the dead zone, thereby removing the pressure maximum and suppressing the IOPF mechanism. This might explain why close-in small planets are *not* found in roughly half of all systems: any background stellar or interstellar field threading the disk will be randomly aligned/anti-aligned with the disk spin axis, yielding alignment in statistically half the cases.

(5) At these low Σ, X-ray ionization can become competitive with thermal ionization, contrary to the standard assumption that the X-rays may be ignored here. In our analysis, where X-rays are not included (their effects are only investigated a posteriori), the MRI-active zone eventually ends where thermal ionization peters out, and $\bar{\alpha}$ falls to the floor value α_{DZ} (forming a pressure maximum there). In real disks, we expect that the MRI-active layer above the dead zone will eventually become X-ray supported, and thus continue outward to join up with the active layer in the midplane beyond the outer boundary of the dead zone

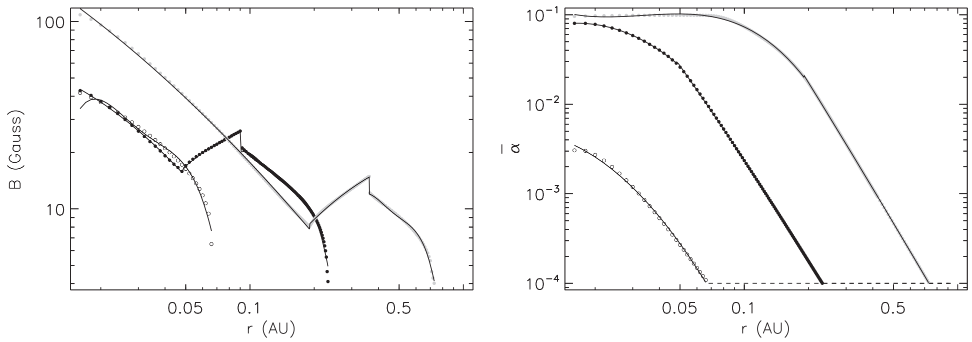


Figure 23. Left: field strength B as a function of radius. Right: $\bar{\alpha}$ as a function of radius. In both plots, results for $\dot{M} = 10^{-10}\,M_\odot\,\mathrm{yr}^{-1}$ (open circles) and $\dot{M} = 10^{-8}\,M_\odot\,\mathrm{yr}^{-1}$ (filled gray circles) are overplotted on the results for our fiducial model with $\dot{M} = 10^{-9}\,M_\odot\,\mathrm{yr}^{-1}$ (filled black circles; the fiducial results are the same ones shown in Figures 6 and 8, respectively).

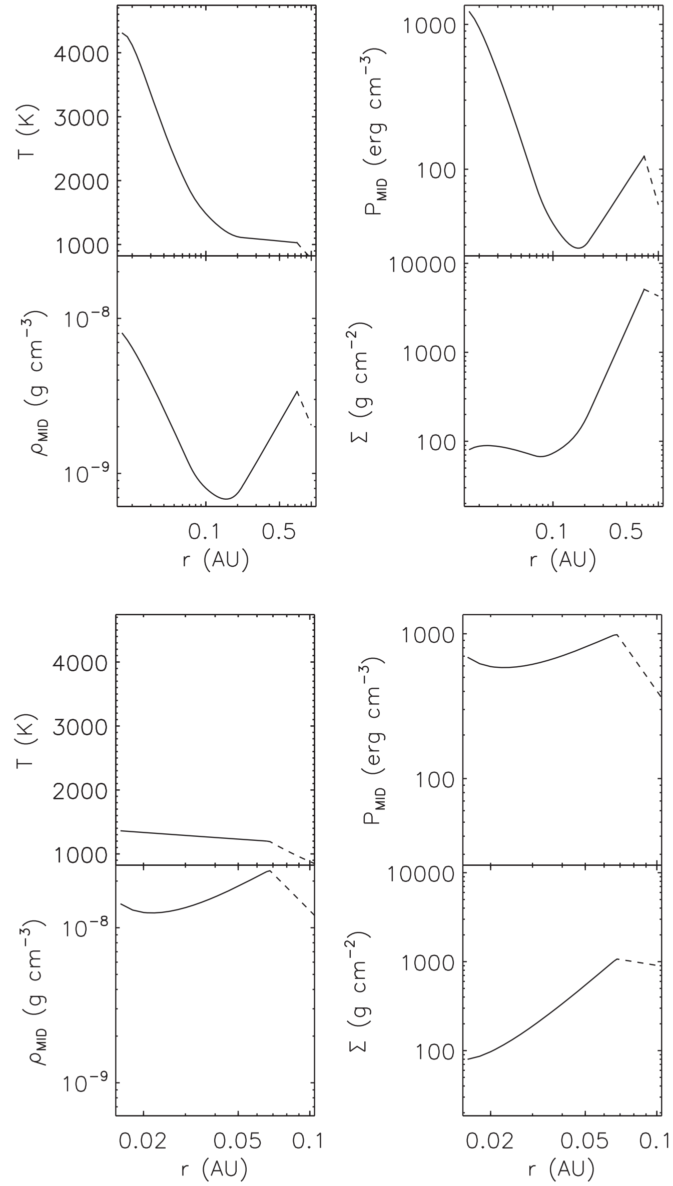


Figure 24. Various disk parameters as a function of radius: same as Figure 9, but now for an accretion rate $\dot{M} = 10^{-8}\,M_\odot\,\mathrm{yr}^{-1}$ (top) and $10^{-10}\,M_\odot\,\mathrm{yr}^{-1}$ (bottom).

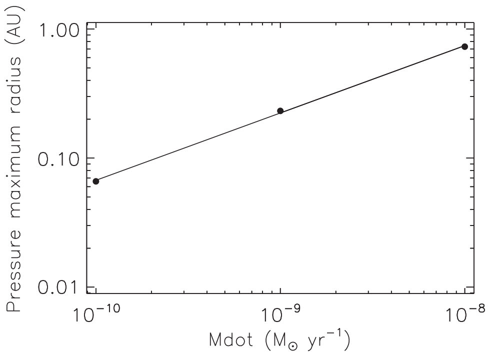


Figure 25. Radial location of the pressure maximum as a function of the accretion rate $\dot{M}$, showing the approximate power-law dependence $r_{P_{\max}} \propto \dot{M}^{1/2}$.

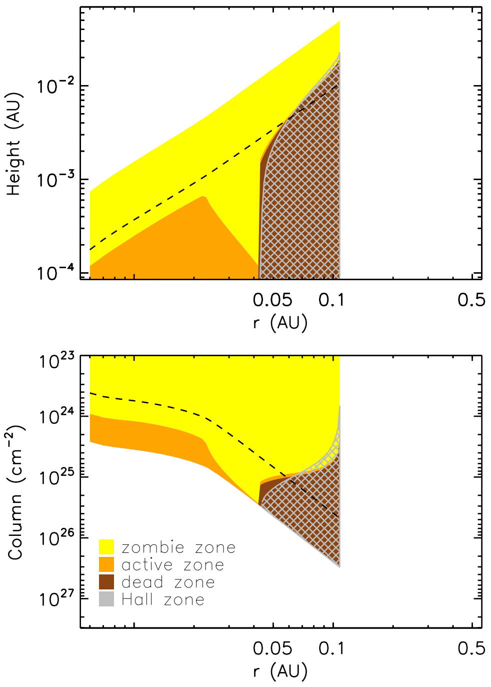


Figure 26. Same as Figure 5, but now for a stellar mass $M_* = 0.1\,M_\odot$.

(e.g., see disk solutions with X-ray-driven MRI by Mohanty et al. 2013). In this case, the minimum value of $\bar{\alpha}$ will be somewhat higher than α_{DZ} (since the disk never becomes completely dead vertically); what this precise value is, and where it is achieved (and thus a pressure barrier is formed), will be X-ray dependent.

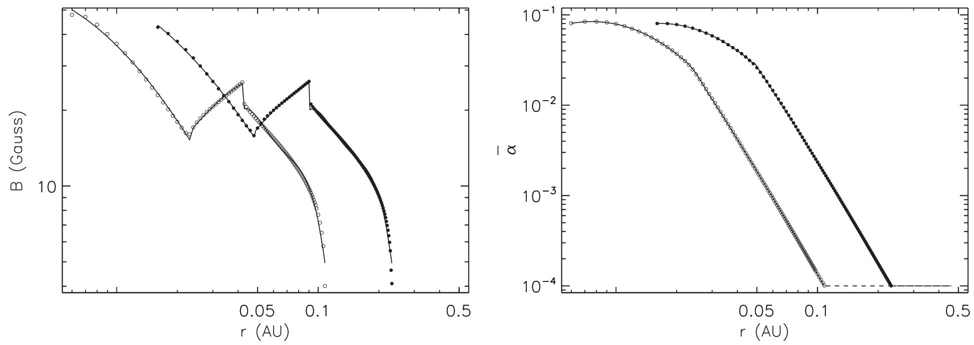


Figure 27. Left: field strength B as a function of radius. Right: $\bar{\alpha}$ as a function of radius. In both plots, results for $M_* = 0.1\, M_\odot$ (open circles) are overplotted on the results for our fiducial model with $M_* = 1\, M_\odot$ (filled black circles; the fiducial results are the same ones shown in Figures 6 and 8, respectively).

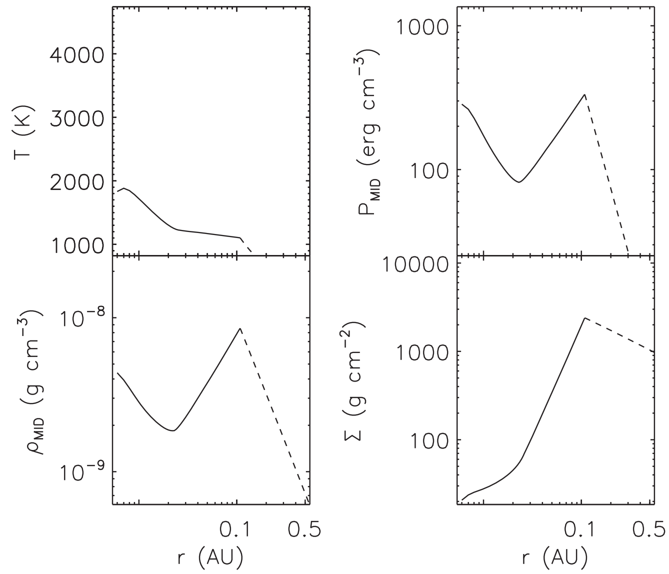


Figure 28. Various disk parameters as a function of radius: same as Figure 9, but now for a stellar mass $M_* = 0.1\, M_\odot$.

(6) A linear stability analysis of our equilibrium disk solutions indicates that most of the inner disk is viscously unstable ($\partial\Sigma/\partial\dot{M} < 0$), with the inner edge of this unstable region moving outward with increasing $\dot{M}$. The instability is driven primarily by the change in $\bar{\alpha}$ (due to variations in the ambipolar and Ohmic diffusivities) as a function of $\dot{M}$. To zeroth order, this instability will cause the inner disk to break up into rings. A more detailed nonlinear analysis, together with the inclusion of more realistic disk physics (i.e., inclusion of grain effects on the MRI, and a more rigorous treatment of the disk thermal structure, ionization including X-rays, and opacities), is required to verify the presence of the viscous instability.

We thank Sourav Chatterjee, Xiao Hu, Zhaohuan Zhu, Mordecai-Mark Mac Low, and Patrick Hennebelle for very helpful discussions. S.M. and J.T. acknowledge support from a Royal Society International Exchange grant IE131607 and from NASA ATP grant NNX15AK20G (JCT). M.R.J. acknowledges support from the Imperial College PhD Scholarship and the Dositeja stipend from the Fund for Young Talents of the Serbian Ministry for Youth and Sport.

Appendix A
Discussion of Conditions for Active MRI

Our treatment of the conditions for active MRI generally follows that of Mohanty et al. (2013); we summarize the salient points here. Magnetic torques are important for mass and angular momentum transport in the disk only if the gas is sufficiently coupled to the field, i.e., if gas motions can generate magnetic stresses faster than they can diffuse away owing to a finite resistivity η. In a Keplerian disk, these stresses arise owing to orbital shear, so the relevant timescale for field generation is the orbital period, i.e., the dynamical timescale $t_{\rm dyn} \sim 1/\Omega$.

For local tangled fields generated by MRI-induced turbulence, the height of the thin disk sets an upper limit on the wavelength of MRI modes, and hence on the dissipation timescale, so it is the vertical direction that is relevant. For a vertical mode with wavenumber k, the Ohmic dissipation rate is $\sim k^2\eta_O$, while the growth rate is kv_{Az}, where v_{Az} is the vertical component of the local Alfvén velocity ($v_{Az} \equiv B_z/\sqrt{4\pi\rho}$, for a local vertical field strength B_z and gas density ρ; MRI simulations by Sano et al. (2004) indicate that $B_z^2 \sim B^2/25$, where B is the rms field strength). Since the maximum growth rate is Ω, the wavenumber of the fastest-growing mode is $k = \Omega/v_{Az}$. Stipulating that the growth rate of this mode exceed its dissipation rate then yields the Ohmic Elsässer number criterion for active MRI:

$$\Lambda \equiv \frac{v_{Az}^2}{\eta_O \Omega} > 1, \tag{33}$$

whether the net background field is vertical, toroidal, or zero (Sano & Stone 2002). Moreover, we will see below that efficient MRI additionally requires the gas pressure in the disk ($P_{\rm gas}$) to substantially exceed the magnetic pressure ($P_{\rm B}$). For gas with sound speed c_s, $P_{\rm gas}$ ($\propto c_s^2$) $\gg P_{\rm B}$ ($\propto v_A^2$) implies $c_s/v_A \gg 1$, guaranteeing that the wavelength $\sim v_{Az}/\Omega$ ($< v_A/\Omega$) of the fastest-growing mode is indeed much smaller than the disk scale height $z_H \sim c_s/\Omega$.

The rationale for the Ohmic Elsässer criterion may be understood more clearly by considering the general induction equation for magnetic fields (e.g., Wardle 2007):

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \nabla \times [\eta_O(\nabla \times \boldsymbol{B}) + \eta_H(\nabla \times \boldsymbol{B}) \times \hat{\boldsymbol{B}} + \eta_A(\nabla \times \boldsymbol{B})_\perp], \tag{34}$$

where $\boldsymbol{v}$ and $\boldsymbol{B}$ are the neutral velocity and magnetic field vectors, respectively, "$\hat{}$" denotes a unit vector, and "$\perp$" indicates the component of a vector perpendicular to $\boldsymbol{B}$. The first expression on the right is the inductive term (I), while the second, third, and fourth terms represent Ohmic (O), Hall (H), and ambipolar (A) diffusion, respectively. Clearly, in magnitude, $I \sim \Omega B$ while $O \sim \eta_O B/L^2$; moreover, as noted above, $L \sim v_{Az}/\Omega$ is the characteristic length scale of the fastest-growing MRI mode (and thus the relevant scale for field diffusion too). Thus $\Lambda \sim I/O$, and Equation (33) simply expresses the intuitive notion that, for robust MRI-driven field amplification when Ohmic resistivity is the prime diffusive channel, the ratio of the inductive to the Ohmic term must exceed unity. A region where Ohmic diffusion kills the MRI, and thus Equation (33) is *not* satisfied (i.e., $\Lambda < 1$), is called a dead zone.

Analogously, when the Hall term dominates the diffusivities on the right-hand side of the induction equation, we expect it to strongly affect the MRI when the Hall Elsässer number χ satisfies

$$\chi \equiv \frac{v_A^2}{|\eta_H|\Omega} < 1. \tag{35}$$

We call a region satisfying Equation (35) the Hall zone. The *nature* of the Hall term's effect on the MRI, however, is very different from that of Ohmic resistivity: in the presence of a net background vertical field threading the disk, the nondissipative character of Hall diffusion implies that it may amplify or suppress the MRI depending on whether the field is aligned or anti-aligned with the rotation axis of the disk (i.e., whether $\boldsymbol{B} \cdot \boldsymbol{\Omega} > 0$ or <0; this behavior can be understood by noticing that flipping the direction of $\boldsymbol{B}$ changes the sign of every term in the induction equation, Equation (34), *except* the Hall one (e.g., Wardle 1999; Balbus & Terquem 2001). These issues have been explored in a linear analysis by Wardle & Salmeron (2012), and in various more recent nonlinear simulations (see discussion of simulation results in Section 2); Xu & Bai (2016) have also explored similar Hall diffusion effects when, in the presence of a net field, grains cause a flip in the sign of η_H that mimics a reversal in field polarity.

Including the quantitative effect of Hall diffusion on the MRI is thus nontrivial and beyond the scope of this exploratory paper; as such, *we ignore it* here. We do calculate the magnitude of all three resistivities (η_O, η_H, and η_A) and show their relative strengths over our region of interest; however, in Hall-dominated areas, we use either the Ohmic Elsässer criterion (Equation (33)) or the ambipolar condition (discussed below; Equations (36) and 37(a)) to evaluate the MRI efficiency, depending on whether η_O or η_A is the next-strongest resistivity.

For ambipolar diffusion, the Elsässer number Am is again defined analogously to the Elsässer number Λ for Ohmic diffusion, but with η_A replacing η_O[9]:

$$Am \equiv \frac{v_A^2}{\eta_A \Omega}. \tag{36}$$

Note that Am is independent of the field strength B, since $v_A^2 \propto B^2$ and so is η_A (see Section 4.3).[10] Wardle (1999) argued that the appropriate criterion for efficient MRI, when ambipolar diffusion dominates instead of Ohmic, should in fact mirror Equation (33), i.e., $Am > 1$. When electrons and ions are the only charged species, the latter condition reduces to $\gamma_i \rho_i/\Omega > 1$ (see Section 4.3), where γ_i is the neutral–ion collisional drag coefficient and ρ_i is the ion density. This implies that the MRI can flourish in the presence of ambipolar diffusion (i.e., the field, to which the ions and electrons are frozen, will be sufficiently coupled to the mainly neutral fluid) only if a neutral particle collides at least once per orbit with an ion. This condition has often been used to investigate ambipolar-dominated disk regions (e.g., Turner et al. 2010).

On the other hand, Hawley & Stone (1998) suggested that the above criterion is too lenient. Their 3D local shearing-box simulations, using an idealized two-fluid approximation (ions + neutrals; ionization and recombination are not considered, so ion and neutral numbers are individually conserved), indicated that efficient MRI with ambipolar diffusion requires neutral–ion collisions to be at least 100 times more frequent, i.e., $\gamma_i \rho_i/\Omega \gtrsim 100$.

However, more recently, Bai & Stone (2011) have argued that it is not the two-fluid approximation but the strong-coupling limit that is most applicable to protoplanetary disks. This limit holds when two criteria are met: (a) the neutral density vastly exceeds the ion density, $\rho_n \gg \rho_i$, a condition invariably satisfied in these disks; and (b) the recombination timescale is much shorter than the orbital period (dynamical timescale), $t_{\rm rcb} \ll t_{\rm dyn} (\sim 1/\Omega)$, which Bai (2011) demonstrates is true over most of the disk as well (see discussion below). In this case, the ion inertia may be neglected, the ion density is controlled entirely by ionization–recombination equilibrium with the neutrals, and the problem reduces to a single-fluid (of neutrals) approximation. In this strongly coupled limit, with the ratio of the inductive to ambipolar term further given by the general expression for Am in Equation (35) (instead of just the reduced value $\gamma_i\rho_i/\Omega$), Bai & Stone (2011) find that the MRI can operate at *any* value of Am, provided that the field is sufficiently weak. Specifically, the MRI can be sustained as long as the plasma β parameter, $\beta \equiv P_{\rm gas}/P_B$, satisfies

$$\beta > \beta_{\rm min}, \tag{37a}$$

where the minimum value of β is a function of Am,

$$\beta_{\rm min}(Am) = \left[\left(\frac{50}{Am^{1.2}}\right)^2 + \left(\frac{8}{Am^{0.3}} + 1\right)^2\right]^{1/2}, \tag{37b}$$

and $P_{\rm B} = B^2/8\pi$. Note from Equation 37(b) that $\beta_{\rm min}$ approaches $(50/Am^{1.2})$ for $Am \lesssim 1$ and asymptotes to 1 from above as $Am \to \infty$. Hence, condition 37(a) for active MRI demands that the gas pressure dominate over the magnetic pressure in the disk, as stated earlier. Following Mohanty et al. (2013), we denote locations where Equations 37(a), (b) are *not* satisfied (i.e., $\beta < \beta_{\rm min}$), and thus the MRI is quenched by ambipolar diffusion, as zombie zones.

Bai (2011) shows that $t_{\rm rcb} \ll t_{\rm dyn}$, i.e., the strong-coupling limit applies, when complex chemical networks are invoked and grains are abundant; he cautions that this limit may not hold in simpler formulations, as recombination pathways become limited. This warning is potentially germane to us, since, in our simplified treatment of thermal ionization here, the chemical network comprises only one channel ($M \rightleftharpoons M^+ + e^-$, where M is a single species of alkali metal), and grains are moreover omitted. We proceed by first *assuming* that the strong-coupling limit holds for us as well, and thus we use conditions 37(a), (b) to determine whether the MRI can operate in ambipolar-dominated regions; we then *check* whether $t_{\rm rcb} \ll t_{\rm dyn}$ holds in these regions in the disk solutions derived, to verify consistency (see discussion and Equation (15) in Section 4.2, and detailed discussion in Section 8.1.8).

---

[9] We use v_A here instead of the v_{Az} employed in the Ohmic Elsässer number definition in Equation (33), since we will adopt (in Equations 37(a), (b)) the results of the numerical simulations by Bai & Stone (2011), who use the total Alfvén velocity to define Am.

[10] A typo in the text of Mohanty et al. (2013) suggests that Am depends on B through both η_A and $v_A{}^2$; while this is formally true, the two dependencies in fact cancel out. This does not vitiate any results in Mohanty et al. (2013), since their actual calculations of Am are correctly implemented.

Appendix B
Connecting the MRI and α-disk Formulations

For a general shear stress $T_{r\phi}$ in the disk, the viscosity parameter α_T is defined by the relation

$$T_{r\phi}\left(\equiv \nu\rho r\frac{d\Omega}{dr}\right) = -\alpha_T P_{\rm gas} = -\alpha_T\rho c_s^2, \tag{38}$$

where the first equivalence in parentheses is the definition of $T_{r\phi}$ and ν is the viscosity. The negative sign on the right-hand-side terms enforces the convention that the α-parameter be positive (since $T_{r\phi} \propto d\Omega/dr < 0$ in a Keplerian disk). We have labeled the α-parameter with the subscript "T" to explicitly denote that it is defined here in terms of $T_{r\phi}$, instead of in terms of the viscosity ν as done in the Shakura–Sunyaev model (Equation (1) in the main text; we connect the two definitions further below). For any particular driver of shear stress (e.g., the MRI), we will find it mathematically convenient to define an *effective* viscosity parameter $\bar{\alpha}_T$, given by the pressure-weighted vertical average of α_T:

$$\bar{\alpha}_T \equiv \frac{\int_{2h}\alpha_T P_{\rm gas}\,dz}{\int_{2h}P_{\rm gas}\,dz} = \frac{\int_{2h}\alpha_T\rho\,dz}{\int_{2h}\rho\,dz}, \tag{39}$$

where the integrals are over the total thickness of the layer (summed over both sides of the midplane) where the specified shear stress operates. The second equality above, which defines $\bar{\alpha}_T$ as a density-weighted vertical average, holds when the disk is vertically isothermal (so that c_s is constant with height). Using Equation (38), this yields the useful form

$$\int_{2h}T_{r\phi}\,dz = -\bar{\alpha}_T\int_{2h}P_{\rm gas}\,dz = -\bar{\alpha}_T c_s^2\int_{2h}\rho\,dz, \tag{40}$$

where the first equality is general and the second again true for a vertically isothermal disk. For the specific case of MRI-driven turbulence, the vertical integral of the turbulent stress is given by (e.g., Wardle 2007)

$$\int_{2h}T_{r\phi,\,\rm MRI}\,dz = -\frac{h}{2\pi}\langle -B_r B_\phi\rangle \approx -\frac{hB^2}{8\pi}, \tag{41}$$

where B_r and B_ϕ are the radial and azimuthal components of the field, respectively, $\langle -B_r B_\phi\rangle \equiv -(2h)^{-1}\int_{2h}B_r B_\phi dz$, and B is the rms field strength. The second equality flows from the result of MRI simulations by Sano et al. (2004) that $B^2 \sim 4\langle -B_r B_\phi\rangle$. Replacing the vertical integral on the left by using the first equality in Equation (40), noting that the vertical average of $P_{\rm gas}$ over the MRI layer is $\langle P_{\rm gas}\rangle \equiv \left(\int_{2h}P_{\rm gas}\,dz\right)/2h$, recognizing that $P_{\rm B} = B^2/8\pi$ is the magnetic pressure, and using the definition of the plasma-beta parameter $\beta \equiv P_{\rm gas}/P_{\rm B}$, we finally arrive at

$$\bar{\alpha}_{T_{\rm MRI}} \approx \frac{1}{2\langle\beta\rangle} \tag{42}$$

for MRI-driven turbulent stresses (as denoted by the subscript on $\bar{\alpha}_T$ on the left-hand side). Here $\langle\beta\rangle \equiv \langle P_{\rm gas}/P_{\rm B}\rangle = \langle P_{\rm gas}\rangle/P_{\rm B}$ is the vertical average of the plasma-beta parameter over the active layer thickness $2h$ (the second equality comes from our assumption that the field strength is constant over this thickness). Bai & Stone (2011) also arrive at Equation (42); it is essentially a restatement of the assertion above that $B^2 \sim 4\langle -B_r B_\phi\rangle$, as they discuss.

Now, at any radial location in the disk, we expect the vertical structure to be multilayered, with the most general structure being a dead zone (where Ohmic resistivity suppresses the MRI) straddling the midplane, a zombie zone (where ambipolar diffusion shuts off the MRI) near the disk top and bottom surfaces, and an MRI-active layer sandwiched in between. The shear stress within the MRI-active and inactive layers is driven by different physical mechanisms, and hence $\bar{\alpha}_T$ within these layers will be (very) different. It is therefore convenient, in analogy with Equation (39) for the individual disk layers, to define an average viscosity parameter $\bar{\alpha}_{T_{\rm avg}}$ over the *entire* thickness of the disk:

$$\bar{\alpha}_{T_{\rm avg}} \equiv \frac{\int_{-\infty}^{+\infty}\alpha_T P_{\rm gas}\,dz}{\int_{-\infty}^{+\infty}P_{\rm gas}\,dz} = \frac{\int_{-\infty}^{+\infty}\alpha_T\rho\,dz}{\int_{-\infty}^{+\infty}\rho\,dz}. \tag{43}$$

For a vertically isothermal disk, which is assumed in this paper (and where the second equality above applies), $\bar{\alpha}_{T_{\rm avg}}$ can be put in a very simple form by noting that

$$\int_{-\infty}^{+\infty}T_{r\phi}\,dz = -\bar{\alpha}_{T_{\rm avg}}c_s^2\int_{-\infty}^{+\infty}\rho\,dz = \sum_i\left(\int_{2h_i}T_{r\phi,i}\,dz\right), \tag{44}$$

where the first equality comes from combining Equations (38) and (43). The second equality simply breaks up the vertical integral over the total disk thickness into a sum of integrals over zones with different shear-stress mechanisms; $2h_i$ and $T_{r\phi,i}$ denote, respectively, the thickness of the ith zone (summed over both sides of the midplane) and the form of the shear-stress tensor there. Using the second equality in Equation (40) to replace the individual integrals under the summation above, and dividing throughout by the mean molecular mass μ, we get

$$\bar{\alpha}_{T_{\rm avg}} = \frac{\sum_i(N_i\,\bar{\alpha}_{T_i})}{N_{\rm tot}}, \tag{45}$$

where $N_i \equiv \int_{h_i}\rho/\mu\,dz$ is the (one-sided) column density of each ith zone, and $N_{\rm tot} \equiv \int_0^{+\infty}\rho/\mu\,dz = \sum_i N_i$ is the (one-sided) total column density from the surface to the midplane (we assume that the disk is symmetric about the midplane). Thus, for a vertically isothermal disk, $\bar{\alpha}_{T_{\rm avg}}$ is the *column-weighted vertical average* of the effective viscosity parameters $\bar{\alpha}_{T_i}$ within each zone (MRI active, dead, and zombie; we will denote these zones by i = MRI, DZ, and ZZ, respectively).

Now, the parameter α that is used to derive the Shakura–Sunyaev disk equations is defined in terms of the viscosity ν (Equation (1) in the main text), while α_T is defined in terms of the shear stress $T_{r\phi}$ (Equation (38)). Combining these two equations with the definition $T_{r\phi} \equiv \nu\, \rho\, r\, d\Omega/dr$, we see that

$$\alpha = \frac{2}{3}\, \alpha_T, \tag{46}$$

where the factor of 2/3 comes from $d\Omega/dr$ in a Keplerian disk.

Moreover, the Shakura–Sunyaev equations (Equations (3)–(6) in the main text) are derived on the basis of vertically integrated quantities (Σ and $\dot{M}$; see H16). As such, it is not α that enters directly into these equations, but more precisely the *effective* parameter $\bar{\alpha}$, which is a vertical average over the entire disk thickness defined analogously to Equation (43):

$$\bar{\alpha} \equiv \frac{\int_{-\infty}^{+\infty} \alpha P_{\rm gas}\, dz}{\int_{-\infty}^{+\infty} P_{\rm gas}\, dz} = \frac{\int_{-\infty}^{+\infty} \alpha \rho\, dz}{\int_{-\infty}^{+\infty} \rho\, dz}. \tag{47}$$

Combining this with Equations (43), (46), and (45) yields

$$\bar{\alpha} = \frac{2}{3}\, \bar{\alpha}_{T_{\rm avg}} = \frac{\frac{2}{3}\sum_i (N_i\, \bar{\alpha}_{T_i})}{N_{\rm tot}}, \tag{48}$$

where the last equality holds for the vertically isothermally case. Note that $\bar{\alpha}_{T_{\rm MRI}} \approx 1/(2\langle\beta\rangle)$ by Equation (42). In the dead and zombie zones, the effective parameters $\bar{\alpha}_{T_{\rm DZ}}$ and $\bar{\alpha}_{T_{\rm ZZ}}$ are set by hydrodynamic and/or gravitational instabilities, and we set their values guided by the results of numerical simulations (see below). Furthermore, without detailed simulations of how the viscosity in the dead and zombie zones might differ, we assume that the effective viscosity parameters are the same in both zones: $\bar{\alpha}_{T_{\rm DZ}} = \bar{\alpha}_{T_{\rm ZZ}}$. Then, for the vertically isothermal conditions that we adopt, we may write

$$\bar{\alpha} = \frac{N_{\rm MRI}\, \bar{\alpha}_{\rm MRI} + (N_{\rm DZ} + N_{\rm ZZ})\bar{\alpha}_{\rm DZ}}{N_{\rm tot}}, \tag{49}$$

where $\bar{\alpha}_{\rm MRI} \equiv 2\bar{\alpha}_{T_{\rm MRI}}/3 \approx 1/(3\langle\beta\rangle)$ and $\bar{\alpha}_{\rm DZ}$ $(=\bar{\alpha}_{\rm ZZ}) \equiv 2\bar{\alpha}_{T_{\rm DZ}}/3$. Based on simulations (e.g., Dzyurkevich et al. 2010, 2013, and references therein; Malygin et al. 2017, and references therein), we adopt a fiducial value of $\bar{\alpha}_{\rm DZ} = 10^{-3}$, 10^{-4}, or 10^{-5}.

Finally, the accretion rate (positive inward) due to the shear stress within any ith zone (MRI active, dead, or zombie) is given by

$$\dot{M}_i = -\frac{2}{r\Omega}\frac{\partial}{\partial r}\left[2\pi r^2 \int_{2h_i} T_{r\phi,i}\, dz\right]. \tag{50}$$

For the vertically isothermal case, we can replace the stress integral using the last equality in Equation (40), which yields

$$\dot{M}_i = \frac{12\pi\, \mu m_{\rm H}}{r\Omega}\frac{\partial}{\partial r}(r^2 c_s^2\, N_i \bar{\alpha}_i), \tag{51}$$

where i = MRI, DZ, or ZZ, and $\bar{\alpha}_{\rm MRI}$ and $\bar{\alpha}_{\rm DZ}$ $(=\bar{\alpha}_{\rm ZZ}$ by assumption here) are defined above. We use this formula to calculate the accretion rates within the individual disk zones. Note that, for the specific case of accretion within the MRI zone, we can combine Equations (41) and (50) to alternatively write

$$\dot{M}_{\rm MRI} \approx \frac{1}{2r\Omega}\frac{\partial}{\partial r}(r^2 h B^2). \tag{52}$$

This shows explicitly how, given a disk structure and chemistry (ionization), the field strength B controls the accretion rate through the MRI-active layers: directly via its appearance in the above formula, and indirectly by influencing the magnitudes of the Ohmic Elsässer number Λ and the plasma β parameter, which in tandem set the active layer thickness h.

For a general shear stress $T_{r\phi}$, the total accretion rate at any radius is (analogous to Equation (50) but now integrated over the entire disk thickness)

$$\dot{M} = -\frac{2}{v_K}\frac{\partial}{\partial r}\left[2\pi r^2 \int_{-\infty}^{+\infty} T_{r\phi}\, dz\right]. \tag{53}$$

If the total accretion rate is *radially constant* (as we shall demand for our equilibrium solutions), then, multiplying throughout by v_K and integrating both sides over radius, from the disk inner edge $R_{\rm in}$ out to any desired radius r, we get

$$\dot{M} = -\frac{2\pi}{f_r\, \Omega}\int_{-\infty}^{+\infty} T_{r\phi}\, dz. \tag{54}$$

The factor $f_r \equiv (1 - \sqrt{R_{\rm in}/r})$ is the same one that appears in the Shakura–Sunyaev equations in Section 3; it arises from the radial integral of v_K. Note that there is no equivalent contribution from the disk inner edge when radially integrating the $\partial/\partial\, r$ term on the right-hand side of Equation (53), since $T_{r\phi} \propto d\Omega/dr = 0$ at the inner edge: in the α-disk model, $R_{\rm in}$ is by definition the location where the angular velocity Ω plateaus and turns over.

Finally, combining the first equality in Equations (44) and (48) to replace the vertical integral of the shear stress in Equation (54) above, and using the definition of surface density $\Sigma \equiv \int_{-\infty}^{+\infty} \rho\, dz$, we arrive at

$$\dot{M} = \frac{3\pi\, \bar{\alpha}\, c_s^2\, \Sigma}{f_r\, \Omega}, \tag{55}$$

the standard expression for a constant accretion rate in a vertically isothermal α-disk model.

Appendix C
Polynomial Fits to Solutions for $\bar{\alpha}(r)$ and $B(r)$

We fit our $\log\bar{\alpha}(\log r)$ and $\log B(\log r)$ solutions with piecewise polynomials over one, two, or three distinct intervals in radius r. The fits are of the form $y = c_0 + c_1 x + c_2 x^2 + \ldots$, where $y = \log\bar{\alpha}$ or $\log B$, $x = \log r$, and c_n is the nth polynomial coefficient. We list the radius intervals and polynomial coefficients for our various disk models in Tables 2–7. The starting radius for the innermost interval for all models is $R_{\rm in} = R_*$ ($=1\, R_\odot$ for $M_* = 0.1\, M_\odot$ and $2.3\, R_\odot$ for $M_* = 1\, M_\odot$).

Table 2
Disk Model: M_*=1 $M_\odot$, $\dot{M}$ = 10^{-9} $M_\odot$ yr^{-1}, $\alpha_{\rm DZ} = 10^{-4}$

Fitted Function	Interval End Radius (au)	c_0	c_1	c_2	c_3	c_4
$\log\bar{\alpha}(\log r)$	0.048	−8.5526396	−9.0739395	−3.1434803	−0.22451420	
	0.232	−6.6333076	−4.3837696	−0.37942259		
$\log B(\log r)$	0.048	−1.3996038	−2.7297839	−0.57806547		
	0.090	0.90110720	−1.4531042	−0.92303502		
	0.232	−65.562920	−307.15327	−531.13562	−407.62824	−116.78615

Table 3
Disk Model: $M_* = 0.1$ $M_\odot$, $\dot{M}$ = 10^{-9} $M_\odot$ yr^{-1}, $\alpha_{\rm DZ} = 10^{-4}$

Fitted Function	Interval End Radius (au)	c_0	c_1	c_2	c_3	c_4
$\log\bar{\alpha}(\log r)$	0.023	−12.472096	−12.102624	−3.8119263	−0.3015077	
	0.108	−8.103168	−4.5772352	−0.34497049		
$\log B(\log r)$	0.023	−2.6963071	−3.4619552	−0.66750054		
	0.042	0.40642769	−1.9598652	−0.89520128		
	0.108	−233.05517	−779.1073	−972.0254	−538.41475	−111.58188

Table 4
Disk Model: $M_* = 1$ $M_\odot$, $\dot{M}$ = 10^{-10} $M_\odot$ yr^{-1}, $\alpha_{\rm DZ} = 10^{-4}$

Fitted Function	Interval End Radius (au)	c_0	c_1	c_2	c_3	c_4
$\log\bar{\alpha}(\log r)$	0.066	−11.604954	−9.0560947	−2.2069353		
$\log B(\log r)$	0.066	−169.31517	−454.52007	−454.24907	−201.59358	−33.457131

Table 5
Disk Model: $M_* = 1$ $M_\odot$, $\dot{M}$ = 10^{-8} $M_\odot$ yr^{-1}, $\alpha_{\rm DZ} = 10^{-4}$

Fitted Function	Interval End Radius (au)	c_0	c_1	c_2	c_3	c_4
$\log\bar{\alpha}(\log r)$	0.190	−6.7811175	−12.02901	−8.2440562	−1.8594328	
	0.730	−4.5536009	−4.1222938	−0.17861849		
$\log B(\log r)$	0.190	−0.18901295	−1.6613468	−0.22499278		
	0.362	1.2930281	−0.098087489	−0.85474294		
	0.730	−1.2770216	−25.759443	−116.39124	−238.42566	−180.2177

Table 6
Disk Model: $M_* = 1$ $M_\odot$, $\dot{M}$ = 10^{-9} $M_\odot$ yr^{-1}, $\alpha_{\rm DZ} = 10^{-5}$

Fitted Function	Interval End Radius (au)	c_0	c_1	c_2	c_3	c_4
$\log\bar{\alpha}(\log r)$	0.048	−8.9655034	−9.8830202	−3.6673762	−0.33667805	
	0.415	−6.5991431	−4.2982932	−0.33100659		
$\log B(\log r)$	0.048	−1.4088427	−2.7388476	−0.58018993		
	0.092	0.99828119	−1.2919053	−0.85648004		
	0.415	−3.0815134	−20.240474	−39.070548	−34.357312	−11.153956

Table 7
Disk Model: $M_* = 1$ $M_\odot$, $\dot{M}$ = 10^{-9} $M_\odot$ yr^{-1}, $\alpha_{\rm DZ} = 10^{-3}$

Fitted Function	Interval End Radius (au)	c_0	c_1	c_2	c_3	c_4
$\log\bar{\alpha}(\log r)$	0.048	−8.9910621	−9.9615069	−3.7341083	−0.35427178	
	0.127	−6.7619972	−4.6539156	−0.51395667		
$\log B(\log r)$	0.048	−1.3612921	−2.6728148	−0.55758751		
	0.086	0.80147769	−1.6020543	−0.97716225		
	0.127	−3619.2649	−14562.646	−21960.398	−14712.958	−3694.7273

ORCID iDs

Subhanjoy Mohanty https://orcid.org/0000-0002-9896-8313
Jonathan C. Tan https://orcid.org/0000-0002-3389-9142
James E. Owen https://orcid.org/0000-0002-4856-7837

References

Bai, X.-N. 2011, ApJ, 739, 50
Bai, X.-N. 2013, ApJ, 772, 96
Bai, X.-N. 2014, ApJ, 791, 137
Bai, X.-N. 2015, ApJ, 798, 84
Bai, X.-N. 2017, ApJ, 845, 75
Bai, X.-N., & Goodman, J. 2009, ApJ, 701, 737
Bai, X.-N., & Stone, J. M. 2011, ApJ, 736, 144
Bai, X.-N., & Stone, J. M. 2013, ApJ, 769, 76
Balbus, S., & Terquem, C. 2001, ApJ, 552, 235
Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214
Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, A&A, 337, 403
Baruteau, C., Crida, A., Paardekooper, S. J., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 667
Baruteau, C., & Masset, F. 2008a, ApJ, 672, 1054
Baruteau, C., & Masset, F. 2008b, ApJ, 678, 483
Birnstiel, T., Ormel, C. W., & Dullemond, C. P. 2011, A&A, 525, A11
Chatterjee, S., & Tan, J. C. 2014, ApJ, 780, 53
Chatterjee, S., & Tan, J. C. 2015, ApJL, 798, L32
Chiang, E., & Laughlin, G. 2013, MNRAS, 431, 3444
Chiang, E., & Murray-Clay, R. 2007, NatPh, 3, 604
Cossou, C., Raymond, S. N., & Pierens, A. 2013, A&A, 553L, 2
Cossou, C., Raymond, S. N., & Pierens, A. 2014, in IAU Symp. 299, Exploring the Formation and Evolution of Planetary Systems (Cambridge: Cambridge Univ. Press), 360
Coughlin, J. L., Mullally, F., Thompson, S. E., et al. 2016, ApJS, 224, 12
Desch, S. J., & Turner, N. J. 2015, ApJ, 811, 156
Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
Dzyurkevich, N., Flock, M., Turner, N. J., Klahr, H., & Henning, T. 2010, A&A, 515, A70
Dzyurkevich, N., Turner, N. J., Henning, T., & Kley, W. 2013, ApJ, 765, 114
Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Frank, J., King, A., & Raine, D. J. 2002, Accretion Power in Astrophysics (3rd ed.; Cambridge: Cambridge Univ. Press)
Fang, J., & Margot, J. L. 2012, ApJ, 761, 92
Fang, J., & Margot, J. L. 2013, ApJ, 767, 115
Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
Fujii, Y. I., Okuzumi, S., Inutsuka, S., et al. 2011, ApJ, 743, 53
Gammie, C. E. 1996, ApJ, 457, 355
Goldreich, P., & Schlichting, H. E. 2014, AJ, 147, 32
Gressel, O., Turner, N. J., Nelson, R. P., & McNally, C. P. 2015, ApJ, 801, 84
Hansen, B., & Murray, N. 2012, ApJ, 751, 158
Hansen, B., & Murray, N. 2013, ApJ, 775, 53
Hawley, J., & Stone, J. 1998, ApJ, 501, 758
Hu, X., Zhu, Z., Tan, J. C., et al. 2017, ApJ, 857, 20
Hu, X., Zhu, Z., Tan, J. C., & Chatterjee, S. 2016, ApJ, 816, 19
Hubeny, I. 1990, ApJ, 351, 632
Igea, J., & Glassgold, A. E. 1999, ApJ, 518, 848 (IG99)
Jenkins, E. B. 2009, ApJ, 700, 1299
Ilgner, M., & Nelson, R. P. 2014, A&A, 445, 205
Keith, S. L., & Wardle, M. 2014, MNRAS, 440, 89
Kley, W., & Nelson, R. P. 2012, ARA&A, 50, 211
Kretke, K. A., & Lin, D. N. C. 2007, ApJL, 664, L55
Kretke, K. A., & Lin, D. N. C. 2010, ApJ, 721, 1585
Kretke, K. A., & Lin, D. N. C. 2012, ApJ, 755, 74
Kretke, K. A., Lin, D. N. C., Garaud, P., & Turner, N. J. 2009, ApJ, 690, 407
Kunz, M. W. 2008, MNRAS, 285, 1494
Lambrechts, M., Johansen, A., & Morbidelli, A. 2014, A&A, 572A, 35
Latter, H. N., & Balbus, S. 2012, MNRAS, 424, 1977
Lesur, G., Kunz, M. W., & Fromang, S. 2014, A&A, 566, A56
Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, ApJS, 197, 8L
Malygin, M. G., Klahr, H., Semenov, D., Henning, Th., & Dullemond, C. P. 2017, arXiv:1704.06786
Masset, F. S., Morbidelli, A., Crida, A., & Ferreira, J. 2006, ApJ, 642, 478
Matsumura, S., Pudritz, R. E., & Thommes, E. W. 2009, ApJ, 691, 1764
Mohanty, S., Ercolano, B., & Turner, N. J. 2013, ApJ, 764, 65
Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2015, ApJS, 217, 31
Nelson, R. P., & Papaloizou, J. C. B. 2003, MNRAS, 339, 993
Ogihara, M., Morbidelli, A., & Guillot, T. 2015, A&A, 578A, 36
Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29
Perez-Becker, D., & Chiang, E. 2011a, ApJ, 727, 2
Perez-Becker, D., & Chiang, E. 2011b, ApJ, 735, 8
Pringle, J. E. 1981, ARA&A, 19, 137
Raymond, S. N., & Cossou, C. 2014, MNRAS, 440, L11
Rouse, C. A. 1961, ApJ, 134, 435
Sano, T., Inutsuka, S., Turner, N. J., & Stone, J. 2004, ApJ, 605, 321
Sano, T., & Stone, J. 2002, ApJ, 577, 534
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Simon, J. B., Lesur, G., Kunz, M. W., & Armitage, P. J. 2015, MNRAS, 454, 1117
Toomre, A. 1964, ApJ, 139, 1217
Turner, N. J., Carballido, A., & Sano, T. 2010, ApJ, 708, 188
Turner, N. J., & Sano, T. 2008, ApJL, 679, L131
Umebayashi, T., & Nakano, T. 1988, PThPS, 96, 151
Wardle, M. 1999, MNRAS, 307, 849
Wardle, M. 2007, Ap&SS, 311, 35
Wardle, M., & Ng, R. 1999, MNRAS, 303, 239
Wardle, M., & Salmeron, R. 2012, MNRAS, 422, 2737
Wood, K., Wolff, M. J., Bjorkman, J. E., & Whitney, B. 2002, ApJ, 564, 887
Xu, R., & Bai, X. N. 2016, ApJ, 819, 68
Youdin, A. N., & Goodman, J. 2005, ApJ, 629, 459
Zhang, X., Liu, B., Lin, D. N. C., & Li, H. 2014, ApJ, 797, 20
Zhu, Z., Hartmann, L., Nelson, R. P., & Gammie, C. F. 2012, ApJ, 746, 110
Zhu, Z., Stone, J. M., & Rafikov, R. R. 2013, ApJ, 768, 143

Angular Momentum and Topological Dependence of Kepler's Third Law in the Broucke-Hadjidemetriou-Hénon Family of Periodic Three-Body Orbits

Marija R. Janković
Faculty of Physics, Belgrade University, Studentski Trg 12, 11000 Belgrade, Serbia

V. Dmitrašinović
Institute of Physics, Belgrade University, Pregrevica 118, Zemun, P.O. Box 57, 11080 Belgrade, Serbia

(Received 3 July 2015; published 10 February 2016)

We use 57 recently found topological satellites of Broucke-Hadjidemetriou-Hénon's periodic orbits with values of the topological exponent k ranging from $k = 3$ to $k = 58$ to plot the angular momentum L as a function of the period T, with both L and T rescaled to energy $E = -0.5$. Upon plotting $L(T/k)$ we find that all our solutions fall on a curve that is virtually indiscernible by the naked eye from the $L(T)$ curve for nonsatellite solutions. The standard deviation of the satellite data from the sixth-order polynomial fit to the progenitor data is $\sigma = 0.13$. This regularity supports Hénon's 1976 conjecture that the linearly stable Broucke-Hadjidemetriou-Hénon orbits are also perpetually, or Kol'mogorov-Arnol'd-Moser, stable.

DOI: 10.1103/PhysRevLett.116.064301

Introduction.—Numerical studies of periodic three-body orbits have increased their output over the past few years—more than 40 new orbits—and their "satellites" have been discovered, Refs. [1–4]. Unlike periodic two-body orbits, which are all ellipses, and thus are all topologically equivalent, the noncolliding three-body periodic orbits have one of infinitely many different topologies. Montgomery, Ref. [5], had devised an algebraic method to associate a free-group element ("word") w with a three-body orbit's topology, and thus to label and classify such periodic orbits; for an elementary introduction to this method, see Ref. [6]. That classification method has recently acquired practical importance in the identification of new three-body orbits, Refs. [1,3,4].

A number of newly discovered orbits, Refs. [1–4], were of the so-called topological satellite type. Such satellite orbits are also known as "bifurcation" in the older literature, Refs. [2,7], where they were only loosely defined in terms of their presumed origin. It was only in Ref. [3] that a precise definition of a topological satellite was given. When this definition was applied to the figure-eight satellites [8], reported in Ref. [3], it led to the discovery of a remarkable "topological Kepler's third law"-like regularity for arbitrary orbits with vanishing angular momenta, Ref. [9]. The immediate question is whether this regularity persists when the angular momentum does not vanish.

The present Letter is an attempt to answer that question, albeit in a single, specific family of three-body orbits, viz., in the Broucke-Hadjidemetriou-Hénon (BHH) family [10–16], that has the simplest nontrivial topology (free group element $w = a$). The main reason for selecting only this family of orbits is that it is the most thoroughly studied family thus far: it is the only family of orbits with a previously determined dependence of the period T on the angular momentum L of (nonsatellite, or progenitor) periodic orbits, Refs. [10–16]. No such, or comparable, study of any of the remaining known families exists to our knowledge at this moment. Moreover, the BHH family is one of only two families [17] of periodic three-body orbits that have been observed in astronomy: all known "hierarchical" triple star systems belong to BHH orbits. Moreover, the Sun-Earth-Moon system may be viewed as a BHH solution, albeit with highly asymmetrical mass ratios.

The first step towards this goal, the one of finding as many different BHH satellite orbits as possible, has already been accomplished in Ref. [18]. Previously, Davoust and Broucke, Ref. [7], had found one (the first $k = 3$) satellite of one retrograde BHH orbit. Reference [18] extended the search for retrograde BHH satellite orbits systematically up to values $k \leq 19$ of the topological exponent k, and more haphazardly up to $k = 58$. Thus, several different types of BHH satellites with identical values of k were discovered [19], as were a few prograde BHH satellites; see the Supplemental Material [20] and the Web site [21]. Prograde BHH satellites have not been studied systematically, as yet, mostly due to their paucity at the values of the angular momentum covered in the searches in Ref. [18]. Presently, it is not known how many satellites ought to exist, and under which conditions. It is interesting, however, that the observed satellites correspond only to linearly stable BHH progenitor orbits. This is in line with Hénon's 1976 conjecture [15,20] about Kol'mogorov-Arnol'd-Moser (KAM) stability of linearly stable BHH orbits.

Then, motivated by the findings reported in Ref. [9], we checked for similar regularities of satellite BHH orbits with nonzero angular momentum. First, we formulated the topological dependence of Kepler's third law for three-body orbits with nonzero angular momenta, and second, we

© 2016 American Physical Society

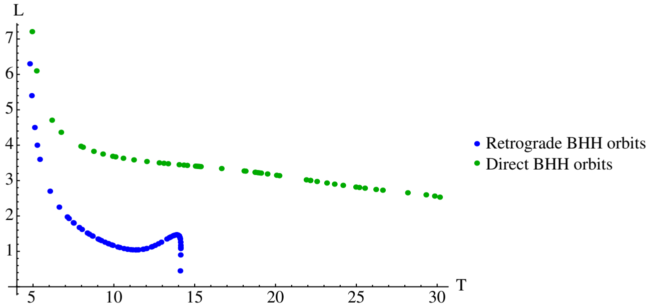


FIG. 1. $L(T)$ curves for direct or prograde (green, upper set of points) and retrograde (blue, lower set of points) BHH orbits, all at fixed energy $E = -0.5$.

tested it on the presently known satellites of the retrograde BHH family. We found a striking result: all of our retrograde BHH satellites fall on a single (continuous) curve $L(T/k)$, Fig. 3, that is practically indiscernible by the naked eye from the $L(T)$ curve, Fig. 1, for nonsatellite (progenitor) retrograde BHH solutions, whereas the "topologically uncorrected" curve $L(T)$ looks very different; see Fig. 2. A quantitative measure of this (dis)agreement is shown in terms of corresponding standard deviations.

Preliminaries.—Broucke [7,10,11], Hadjidemetriou [12–14], and Hénon [15,16] (BHH) explored a set of periodic planar three-body orbits with equal mass bodies. These orbits form two continuous curves in the L-T plane whose lower (retrograde) terminus ("end") is the collinear collision (Schubart) orbit, and both the retrograde and the direct $L(T)$ curves approach the same high-L limit at their upper termini, Fig. 1.

Although BHH write of two families of orbits—direct, or prograde, and retrograde—all of these orbits belong to a single topological family: during one period the orbit completes a single "loop" around one of the poles on the shape sphere. This loop can be described by the conjugacy class of the fundamental group or free group element a, according to the topological classification used in Refs. [1,6]. It turns out, however, that there are numerous relative periodic orbits with topology a^k, with $k = 2, 3, \ldots$. Such orbits are sometimes called satellites [2,3], whereas other authors call them "bifurcation orbits" [7].

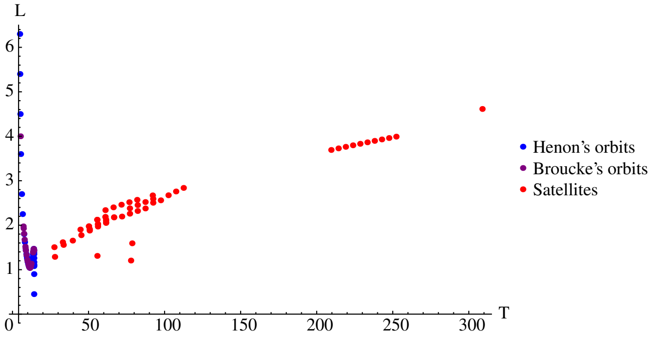


FIG. 2. $L(T)$ dependence of retrograde BHH orbits (blue dots of different hues) and their satellites (red), with various values of k, all at fixed energy $E = -0.5$. The data are from Table I.

Scaling laws for three bodies.—It is well known that Kepler's third law (for two bodies) follows from the spatiotemporal scaling laws, which, in turn, follow from the homogeneity of the Newtonian gravity's static potential, Ref. [22]. These scaling laws read $\mathbf{r} \to \lambda\mathbf{r}$, $t \to \lambda^{3/2}t$, and, consequently, $\mathbf{v} \to \mathbf{v}/\sqrt{\lambda}$. The (total) energy scales as $E \to \lambda^{-1}E$, the period T as $T \to \lambda^{3/2}T$, and angular momentum as $L \to \lambda^{1/2}L$, i.e., differently than either the period T, or "size" R, which is the reason why only the vanishing angular momentum $L = 0$ is a "fixed point" under scaling. For this reason, we use scale-invariant angular momentum $L_r = L|E|^{1/2}$, scale-invariant period $T_r = T|E|^{3/2}$ and, for simplicity's sake, equal masses. Thus, we may replace the "mean size" $\bar{R}$ of the three-body system in Kepler's third law $T \propto \bar{R}^{3/2}$ with the inverse absolute value of energy $|E|^{-1}$, i.e., $T \propto |E|^{-3/2}$, or equivalently $T|E|^{3/2} = T_r = \text{const}$.

The "constant" on the right-hand side of this equation is not a universal one in the three-body case, as it is in the two-body case (where it depends only on the masses and the Newtonian coupling G). It may depend both on the family w of the three-body orbit, described by the free-group word w, and on the scale-invariant angular momentum $L_r = L|E|^{1/2}$ of the orbit, see Refs. [15,16], as follows:

$$T^{(w)}|E|^{3/2} = T_r^{(w)} = f(L^{(w)}|E|^{1/2}) = f(L_r^{(w)}),$$

or as an inverse function,

$$L_r^{(w)} = L^{(w)}|E|^{1/2} = f^{-1}(T^{(w)}|E|^{3/2}) = f^{-1}(T_r^{(w)}).$$

Thus, the curve $L_r^{(w)}(T_r^{(w)}) = L^{(w)}|E|^{1/2}(T^{(w)}|E|^{3/2})$ as a function of $T_r^{(w)} = T^{(w)}|E|^{3/2}$ is a fundamental property of any family w of periodic orbits. For the BHH family the $L(T)$ curve, for fixed energy $E = -0.5$ orbits, based on the data from Refs. [10–16], is shown in Fig. 1.

We wish to see if the zero-angular-momentum relation $T_r(w^k) = kT_r(w)$, Ref. [9], or some similar statement holds also at nonzero angular momentum. The analogon of this relation for orbits with nonzero angular momenta would be a simple relation between $L(T)$ curves for the progenitor orbit $L_r(T_r)$ and its kth satellite $L_r^{(w^k)}(T_r^{(w^k)})$:

$$L_r^{(w)}(T_r^{(w)}) = L_r^{(w^k)}(T_r^{(w^k)}/k). \tag{1}$$

We shall test this relation in the BHH family of solutions, and in order to do so, we use the BHH satellite orbits from Ref. [18].

$L(T)$ curves for BHH satellites.—The L-T plots of different-k satellite orbits are scattered over a large region and do not intersect the BHH progenitor family of orbits'

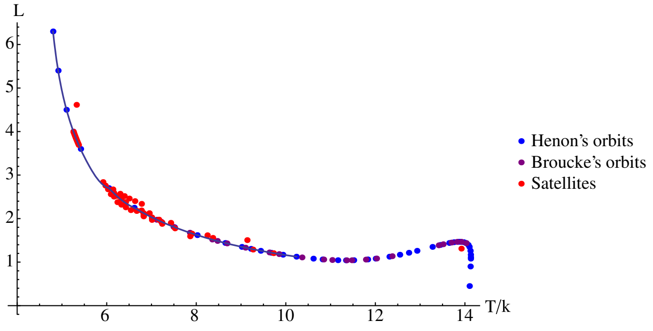


FIG. 3. $L(T' = T/k)$ dependence at fixed energy $E = -0.5$ for the aggregate set of retrograde BHH orbits (blue dots of different hues) and their satellites (red dots) with various values of k, together with the fitted interpolating function (blue solid). The data are from Table I.

$L(T)$ curve when plotted as a function of the (undivided) period T, see Fig. 2. Note the large span of periods T in the data, Table I, and in Fig. 2, as well as two large "gaps" in the data. These gaps are due to the exigencies of the search reported in Ref. [18], which was not conducted with the intention of testing the hypothetical topological Kepler's third law. The values in Table I have been rounded off to five significant decimal places. So, the numerical error is less than one part in 10 000. Such an error would be invisible in Figs. 2, 3, and 4, meaning that the "size of the points" in these figures is larger than the expected error. After dividing the period T (at fixed energy) by the topological exponent k, $T' = T/k$, we can see in Fig. 3 that the satellite orbits' $L(T/k)$ curve (the angular momentum L as a function of topologically rescaled period T/k) approximately coincides with the $L(T)$ curve of BHH retrograde orbits. It seems that such an appearance of order out of apparent disorder cannot be an accident.

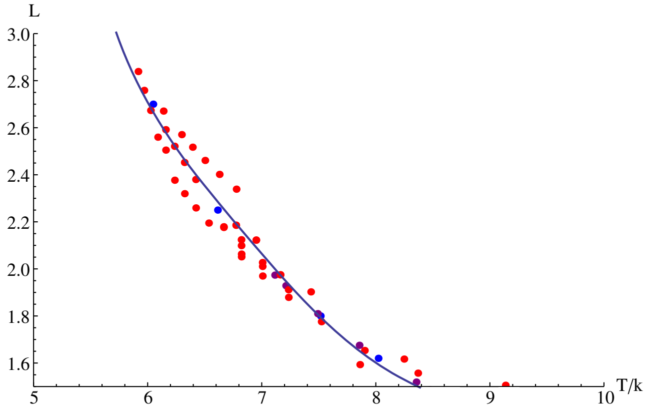


FIG. 4. Enlargement of the $L \in [1.5, 3]$ region of the retrograde BHH orbits (blue dots) and their satellites (red dots) with various values of k $L(T' = T/k)$ dependence at fixed energy $E = -0.5$. Note that the size of the dots on the diagram exceeds the corresponding numerical uncertainties ("error bars").

Next, in Fig. 3 we look more closely at the section of the $L(T)$ curve of progenitor BHH retrograde orbits in which we have found all but one of our satellites. We have interpolated Hénon's [15] 18 stable retrograde data points with a piecewise polynomial fit in this part of the $L(T)$ curve. The standard deviations from this interpolated curve were calculated for (1) Broucke's 10 progenitor retrograde orbits [10,11] and (2) the 56 out of 57 new satellite orbits from Table I (excluding one orbit that lies near the "shoulder" at $T = 14$ in Fig. 3), with the following results. (1) $\sigma = 0.0034$ for Broucke's orbits, and (2) $\sigma = 0.1269$ for satellite orbits. This difference of 2 orders of magnitude between these two numbers clearly indicates that the rescaled satellites' periods do *not* coincide *exactly* with the progenitor ones, but only approximately.

Moreover, when one assembles Hénon's and Broucke's [10,11] retrograde orbits in one set and fits the aggregate data by a polynomial of the sixth degree, Fig. 3, the standard deviation of the fit is $\sigma = 0.0313$, whereas the

TABLE I. Properties of satellite orbits in the retrograde branch of the BHH family. Here, k is the topological power of the orbit, T is its period, and L is its angular momentum. All orbits have the same energy $E = -(1/2)$. For the raw data and a discussion of numerical errors, see the Supplemental Material [20].

T	L	k	T	L	k
27.800 80	1.288 15	3	71.538 38	2.460 95	11
27.411 57	1.505 52	3	77.074 74	2.259 18	12
32.992 45	1.616 82	4	77.060 60	2.379 81	12
33.479 35	1.557 01	4	76.731 11	2.517 18	12
55.678 84	1.310 00	4	82.213 27	2.319 68	13
39.511 02	1.653 31	5	82.199 18	2.452 31	13
45.138 27	1.775 68	6	81.882 58	2.570 68	13
44.586 32	1.902 40	6	87.317 60	2.376 87	14
50.646 60	1.879 00	7	87.303 60	2.520 98	14
50.638 90	1.911 39	7	92.384 79	2.504 86	15
50.141 13	1.974 52	7	92.377 38	2.591 66	15
50.141 28	1.975 37	7	92.082 10	2.670 70	15
56.060 83	1.969 71	8	97.432 10	2.559 79	16
55.604 11	2.121 89	8	102.450 58	2.673 31	17
77.813 66	1.205 44	8	107.449 64	2.758 61	18
56.052 69	2.010 54	8	112.429 18	2.838 83	19
56.049 53	2.027 09	8	209.487 95	3.692 20	39
55.604 30	2.122 89	8	214.258 15	3.727 85	40
61.399 03	2.051 28	9	219.023 02	3.762 83	41
60.968 89	2.185 81	9	223.782 78	3.797 19	42
61.390 86	2.098 90	9	228.537 63	3.830 94	43
61.386 76	2.123 97	9	233.287 75	3.864 12	44
60.968 79	2.185 32	9	238.033 32	3.896 75	45
60.999 96	2.338 82	9	242.774 50	3.928 85	46
61.396 97	2.063 00	9	247.511 46	3.960 44	47
66.666 44	2.179 17	10	252.244 33	3.991 55	48
66.666 89	2.176 08	10	308.853 30	4.614 04	58
66.297 61	2.401 65	10			
78.610 58	1.593 25	10			
71.897 15	2.194 81	11			

standard deviation of all satellite orbits from this polynomial curve is $\sigma = 0.1315$, roughly four times bigger. It is (statistically) clear that the satellites do not follow exactly the same $L(T)$ curve as the progenitors, but the deviation is not large. This constitutes the evidence for the analogon of the topological dependence of Kepler's third law for the $L \neq 0$ case, Ref. [9].

Finally, we note that all of our newly found satellite orbits fall into a region of the progenitor $L(T)$ curve that corresponds to stable progenitor BHH orbits, with one possible exception (the red point near the shoulder at $T = 14$ in Fig. 3, that "sits" on the border point between stable and unstable regions). We have not found any other satellites in this, the second stable region of BHH retrograde orbits. In Fig. 4 we show the fine structure in the satellites' $L(T/k)$ curve, that remains to be studied in finer detail and be better understood.

We have not studied the direct or prograde (sub)family of BHH orbits, as Ref. [18] did not search for their satellites, but found four almost inadvertently. Certainly, that task ought to be completed in the future.

Summary, conclusions and outlook.—We have used 57 new satellite orbits from Ref. [18], in the family of Broucke-Hadjidemetriou-Hénon, Refs. [10–16], relative periodic solutions to the planar three body problem. Thence followed a striking relation between their kinematic and topological properties.

BHH orbits constitute a family with a simple topology, described by the free group element a according to the classification on the shape sphere, and their satellites are orbits of the topology a^k. The BHH orbits' angular momenta L and periods T form a continuous curve $L(T)$, at fixed energy. Our satellite orbits form a scattered set of points on the same $L(T)$ plot, but all of them exhibit the property that after their period T is divided by their topological order k, they approximately fall on the $L(T)$ curve of the original ($k = 1$) BHH orbits.

This study was motivated by the discovery, Ref. [9], of a relation between the topology and periods among the satellites of the figure-eight orbit, Ref. [3], and one other type ("moth I"—"yarn" in Ref. [1]), of three-body orbits at vanishing angular momentum. This Letter shows that Kepler's third law's topological dependence also holds for orbits with $L \neq 0$, albeit only approximately. It remains to be seen just precisely what this discrepancy depends on.

These results are even more striking if one remembers that among our results there are several distinct types of satellite orbits of the same topological power k, some with quite different values of L and T, which all display this property. A closer look at the $L(T/k)$ curve revealed a fine structure, which should be investigated in higher detail in the future. An extension of the search conducted in Ref. [18] into hitherto unexplored regions of the L-T plane ought to provide (new) data that will further test our result.

Our results indirectly confirm Hénon's 1976 conjecture, see page 282 in Ref. [15], reproduced in the Supplemental Material [20], that the linearly stable BHH orbits are also nonlinearly, or perpetually, or KAM stable. Such KAM stability implies the existence of quasiperiodic orbits with periods that conform to the quasiperiodicity condition (i.e., with periods that are "almost commensurate" with the BHH progenitor's period), as predicted by the KAM theorem, Refs. [23–25].

Our study opens several new questions. (1) The most commonly observed hierarchical triple star systems belong to the BHH family. Are there BHH topological satellites among astronomically observed three-body systems? It is important to extend the present study to the realistic case of three different masses: some early work has already been done in this direction by Broucke and Boggs, Ref. [10], and by Hadjidemetriou and Christides, Ref. [13]. (2) In recent years there have been formal "proofs of existence" given for at least some BHH orbits, Refs. [26,27]. This begs the question: can one "prove existence" of their satellite orbits, and, if yes, of how many satellites, and under which conditions?

M. R. J. was a recipient of the "Prof. Dr. Djordje Živanović" scholarship, awarded jointly by the Faculty of Physics and the Institute of Physics, Belgrade University, and was also supported by a City of Belgrade studentship (Gradska stipendija grada Beograda). The work of V. D. was supported by the Serbian Ministry of Science and Technological Development under Grants No. OI 171037 and No. III 41011. The computing cluster Zefram (zefram.ipb.ac.rs) at the Institute of Physics Belgrade has been used for numerical calculations.

[1] M. Šuvakov and V. Dmitrašinović, Three Classes of Newtonian Three-Body Planar Periodic Orbits, Phys. Rev. Lett. **110**, 114301 (2013).

[2] C. Simó, Dynamical properties of the figure eight solution of the three-body problem, in *Celestial Mechanics*, edited by A. Chenciner, R. Cushman, C. Robinson, and Z. J. Xia, Contemporary Mathematics Vol. 292 (AMS, Providence, RI, 2002), p. 209.

[3] M. Šuvakov, Numerical search for periodic solutions in the vicinity of the figure-eight orbit: Slaloming around singularities on the shape sphere, Celest. Mech. Dyn. Astron. **119**, 369 (2014).

[4] M. Šuvakov and M. Shibayama, Three topologically nontrivial choreographic motions of three bodies, Celest. Mech. Dyn. Astron., doi:10.1007/s10569-015-9657-9 (2015).

[5] R. Montgomery, The N-body problem, the braid group, and action-minimizing periodic solutions, Nonlinearity **11**, 363 (1998).

[6] M. Šuvakov and V. Dmitrašinović, A guide to hunting periodic three-body orbits, Am. J. Phys. **82**, 609 (2014).

[7] E. Davoust and R. Broucke, A manifold of periodic orbits in the planar general three-problem with equal masses, Astron. Astrophys. **112**, 305 (1982).

[8] Satellite orbits of the figure eight were first observed in Ref. [2] and further investigated in Refs. [3,4].

[9] V. Dmitrašinović and M. Šuvakov, Topological dependence of Kepler's third law for planar periodic three-body orbits with vanishing angular momentum, Phys. Lett. A **379**, 1939 (2015).

[10] R. Broucke and D. Boggs, Periodic orbits in the planar general three-body problem, Celest. Mech. **11**, 13 (1975).

[11] R. Broucke, On relative periodic solutions of the planar general three-body problem, Celest. Mech. **12**, 439 (1975).

[12] J. D. Hadjidemetriou, The continuation of periodic orbits from the restricted to the general three-body problem, Celest. Mech. **12**, 155 (1975).

[13] J. D. Hadjidemetriou and Th. Christides, Families of periodic orbits in the planar three-body problem, Celest. Mech. **12**, 175 (1975).

[14] J. D. Hadjidemetriou, The stability of periodic orbits in the three-body problem, Celest. Mech. **12**, 255 (1975).

[15] M. Hénon, A family of periodic solutions of the planar three-body problem, and their stability, Celest. Mech. **13**, 267 (1976).

[16] M. Hénon, Stability of interplay motions, Celest. Mech. **15**, 243 (1977).

[17] The other one being the Lagrange family of orbits.

[18] M. R. Janković and M. Šuvakov, Extension of the Broucke-Hadjidemetriou-Henon family of periodic orbits in the Newtonian planar three-body problem (to be published).

[19] The presence of multiple satellites with the same topology is not the first known instance of its kind. There are (many) different satellites of the figure-eight orbit with identical values of k, see Refs. [3,4], albeit with zero angular momentum.

[20] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.116.064301 for The search method is described, tables with the data regarding the new BHH satellites, together with the numerical uncertainties, are shown. We briefly discuss Hénon’s conjecture about KAM stability of BHH orbits.

[21] http://three-body.ipb.ac.rs/bh.php; http://three-body.ipb.ac.rs/bhh.satellites.php.

[22] L. D. Landau and E. M. Lifshitz, *Mechanics*, 3rd ed. (Butterworth-Heinemann, Oxford, 1976).

[23] A. N. Kolmogorov, Preservation of conditionally periodic movements with small change in the Hamilton function, Dokl. Akad. Nauk SSSR **98**, 527 (1954).

[24] V. I. Arnold, Proof of A. N. Kolmogorov's theorem on the preservation of quasiperiodic motions under small perturbations of the Hamiltonian, Russ. Math. Surv. **18**, 9 (1963).

[25] J. Moser, On invariant curves of area-preserving mappings on an annulus, Nachr. Akad. Wiss. Goettingen 2A **1**, 1 (1962).

[26] K.-C. Chen, Existence and minimizing properties of retrograde orbits to the three-body problem with various choices of masses, Ann. Math. **167**, 325 (2008).

[27] K.-C. Chen and Y.-C. Lin, On action-minimizing retrograde and prograde orbits of the three-body problem, Commun. Math. Phys. **291**, 403 (2009).

Physics of Plasmas

Numerical solution of the quantum Lenard-Balescu equation for a non-degenerate one-component plasma

Cite as: Phys. Plasmas **23**, 092119 (2016); https://doi.org/10.1063/1.4963254
Submitted: 28 April 2016 . Accepted: 08 September 2016 . Published Online: 29 September 2016

Christian R. Scullard, Andrew P. Belt, Susan C. Fennell, Marija R. Janković, Nathan Ng, Susana Serna, and Frank R. Graziani

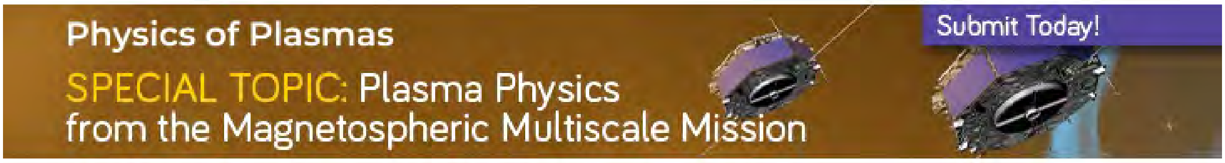
View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

On the quantum Landau collision operator and electron collisions in dense plasmas
Physics of Plasmas **23**, 032706 (2016); https://doi.org/10.1063/1.4944392

Properties of the Lenard–Balescu collision operator: A numerical study
Physics of Plasmas **9**, 430 (2002); https://doi.org/10.1063/1.1418720

An electron conductivity model for dense plasmas
The Physics of Fluids **27**, 1273 (1984); https://doi.org/10.1063/1.864744


Physics of Plasmas
SPECIAL TOPIC: Plasma Physics from the Magnetospheric Multiscale Mission
Submit Today!

© 2016 Author(s).

PHYSICS OF PLASMAS **23**, 092119 (2016)

Numerical solution of the quantum Lenard-Balescu equation for a non-degenerate one-component plasma

Christian R. Scullard,[1,a)] Andrew P. Belt,[2,b)] Susan C. Fennell,[2,c)] Marija R. Janković,[2,d)] Nathan Ng,[2,e)] Susana Serna,[3] and Frank R. Graziani[1]
[1]*Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
[2]*Institute for Pure and Applied Mathematics, UCLA, Los Angeles, California 90095, USA*
[3]*Departament de Matematiques, Universitat Autònoma de Barcelona, 08193 Bellaterra-Barcelona, Spain*

(Received 28 April 2016; accepted 8 September 2016; published online 29 September 2016)

We present a numerical solution of the quantum Lenard-Balescu equation using a spectral method, namely an expansion in Laguerre polynomials. This method exactly conserves both particles and kinetic energy and facilitates the integration over the dielectric function. To demonstrate the method, we solve the equilibration problem for a spatially homogeneous one-component plasma with various initial conditions. Unlike the more usual Landau/Fokker-Planck system, this method requires no input Coulomb logarithm; the logarithmic terms in the collision integral arise naturally from the equation along with the non-logarithmic order-unity terms. The spectral method can also be used to solve the Landau equation and a quantum version of the Landau equation in which the integration over the wavenumber requires only a lower cutoff. We solve these problems as well and compare them with the full Lenard-Balescu solution in the weak-coupling limit. Finally, we discuss the possible generalization of this method to include spatial inhomogeneity and velocity anisotropy. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4963254]

I. INTRODUCTION

The Landau equation, or its equivalent formulation in terms of the Fokker-Planck equation,[1] is a valuable tool in the study of out-of-equilibrium weakly coupled plasmas.[2,3] The assumption of small-angle binary scattering between the particles is well-suited to Coulomb interactions at high temperature and low density. However, this approximation results in a divergence at small impact parameters, and the neglect of screening leads to a divergence for large particle separations due to the long-range nature of the Coulomb interaction. As is well-known, these divergences require cutoffs, which, in practice, mean choosing a Coulomb logarithm and thereby adding a level of ambiguity to the calculation. Although a more realistic calculation does contain such a logarithmic term, there are other terms potentially the same order as $\log \Lambda$ that we are discarding by using Landau/Fokker-Planck. To include these terms requires a more sophisticated collision operator. An option is a quantum version of the Boltzmann equation,[4] which naturally handles strong collisions, avoiding the small-angle approximation. Another candidate is the quantum Lenard-Balescu (QLB) equation, which accounts for both quantum diffraction, solving the large-k (wavenumber) divergence, and dynamical screening, giving convergence as $k \to 0$, in a natural way and thus requires no input Coulomb logarithm. This equation has been used extensively to calculate various plasma properties at weak coupling, such as transport coefficients[5–8] and temperature equilibration rates.[9,10] These computations do not require a time-dependent solution of the QLB equation, and indeed the latter has rarely been attempted; the quantum Lenard-Balescu equation is far more complicated than Landau/Fokker-Planck, which itself is not trivial to solve.[11,12] We present here a numerical solution of the quantum Lenard-Balescu equation for a velocity-isotropic, spatially homogeneous, one-component plasma.

The paper is organized as follows. In Section II, we describe the equation in detail, and in Section III, we introduce our solution method, which, for reasons discussed there, is very different from those traditionally used to solve the Fokker-Planck equation. In Sections IV and V, we describe how we solve the most difficult problem, the integration over the dielectric function. Our solution method can easily be applied to several simpler kinetic equations, such as the Landau equation, and we enumerate these in Section VI and give the minor modifications needed for each. In Sections VII–IX, we describe our initial conditions, the numerical solution of the ordinary differential equations that arise from our method, and we show the relaxation to equilibrium of various initial distributions. In the remainder of the paper, we discuss possible generalizations of the method to handle anistropy in velocity and inhomogeneity in space.

II. QUANTUM LENARD-BALESCU EQUATION

The equation we will solve is the non-degenerate quantum[13] Lenard-Balescu[14,15] equation for a one-component plasma

$$\frac{\partial f}{\partial t} = C_{QLB}(f), \tag{1}$$

where

[a)] scullard1@llnl.gov

[b)] Present address: University of Tennessee, Knoxville, Tennessee 37996, USA.

[c)] Present address: University of Limerick, Limerick, Ireland.

[d)] Present address: University of Belgrade, Studentski Trg 12, 11000 Belgrade, Serbia.

[e)] Present address: University of Maryland, College Park, Maryland 20742, USA.

$$\begin{aligned} C_{QLB}(f) = -\frac{1}{4\pi^2\hbar^2}\int d^3\mathbf{v}'\int d^3\mathbf{k}\,&\frac{|\phi(k)|^2}{\left|\epsilon\left(k,\mathbf{k}\cdot\mathbf{v}+\frac{\hbar k^2}{2m}\right)\right|^2} \\ &\times\delta\left[\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')+\hbar k^2/m\right] \\ &\times\left[f(\mathbf{v})f(\mathbf{v}')-f(\mathbf{v}+\hbar\mathbf{k}/m)f\left(\mathbf{v}'-\hbar\mathbf{k}/m\right)\right], \end{aligned} \tag{2}$$

where m is the particle's mass, $\hbar$ is Planck's constant, and we use the Coulomb potential

$$\phi(k) = \frac{4\pi e^2}{k^2}. \tag{3}$$

The dielectric function is given by

$$\epsilon(k,\omega) = 1 - \frac{4\pi e^2}{k^2}\chi(k,\omega) \tag{4}$$

and $\chi(k,\omega)$ is the free-particle response function

$$\begin{aligned} \chi(k,\omega) &= \chi^{(0)}(k,\omega) \\ &\equiv \lim_{\eta\to 0^+}\int d^3\mathbf{v}\,\frac{f(\mathbf{v})-f(\mathbf{v}+\hbar\mathbf{k}/m)}{\hbar\omega-\hbar\mathbf{v}\cdot\mathbf{k}-\frac{\hbar^2k^2}{2m}+i\eta}. \end{aligned} \tag{5}$$

This equation is valid when the system is non-degenerate, i.e., when

$$\theta \equiv \frac{2mk_BT}{\hbar^2(3\pi^2n)^{2/3}} \gg 1, \tag{6}$$

where n is the number density, and weakly coupled

$$\Gamma \equiv \frac{e^2(4/3\pi n)^{1/3}}{k_BT} \ll 1. \tag{7}$$

When the former condition is violated, additional factors of $1-f$ appear in the integrand in (2), and the latter is required for the validity of the random phase approximation. Generally speaking, this equation describes high-temperature, low-density plasmas.

The presence of the distribution in the response function is a serious complication. Even worse, integrals over the dielectric function often contain very narrow peaks and their numerical integration can be tricky even at equilibrium[16,17] let alone for arbitrary distributions. These difficulties, coupled with the fact that the Landau equation, despite its deficiencies, yields distributions that are likely qualitatively correct at weak coupling, have kept the Lenard-Balescu (LB) equation from being studied numerically in any serious way in plasma physics. There are, however, several examples of its solution in the context of carrier scattering in semiconductors (e.g., Binder *et al.*,[18] and see the book of Bonitz[19] for further discussion and references). We know of only one previous attempt in the plasma context: Dolinsky's pioneering 1965 solution of the classical LB equation[20] using a discretization method in velocity. This work predates the advent of conservative velocity discretization schemes even for the Fokker-Planck equation, but it is not completely clear that such methods are generically well-suited to the Lenard-Balescu equation anyway because of the need to integrate accurately over the features of the dielectric function. This issue could certainly use a more thorough investigation. In any case, the classical equation considered by Dolinsky is divergent at large k and, unlike the quantum version, an artificial cutoff is needed. Besides Dolinksy, there is also the somewhat related work of Ricci and Lapenta,[21] in which they consider a one-dimensional version of the Lenard-Balescu equation. While certainly interesting, their system is primarily of theoretical value (it cannot equilibrate, for example). Although many sophisticated techniques are now available for the Landau and Boltzmann equations,[22–27] enabling solution in multiple spatial and velocity dimensions with several different particle species, we are only capable, for the moment, of a solution of the QLB equation for a spatially homogeneous, one-component plasma with an isotropic velocity distribution. In Section X, we will discuss how the method can be generalized.

As we explain in detail in Section V, after the initial condition has been chosen, only one dimensionless combination of the various physical parameters is really important in the subsequent evolution. We therefore do not lose much by specializing to electrons, so that m in the above equations is equal to the electron mass, m_e, and fixing the number density which we shall henceforth call n_e.

III. METHOD

Because of the difficulties associated with the dielectric function, we choose to steer clear of discretization in velocity. Instead, we use an expansion in Laguerre polynomials

$$f(v,t) = f^{\mathrm{eq}}(v)\sum_{n=0}^{\infty}A_n(t)L_n^{\left(\frac{1}{2}\right)}\left(\frac{u\beta m_e v^2}{2}\right), \tag{8}$$

where

$$f^{\mathrm{eq}}(v) \equiv n_e\left(\frac{m_e\beta}{2\pi}\right)^{3/2}\exp\left(-\frac{m_e\beta v^2}{2}\right) \tag{9}$$

is the Maxwell distribution, $\beta\equiv 1/k_BT$, k_B is Boltzmann's constant, T the temperature of the final equilibrium state, and n_e is the particle number density. The parameter $u\in[1,2]$ will be discussed in detail below. Multiplying by the Maxwell distribution is convenient because it is the stationary solution of this form of the QLB equation, and thus in equilibrium we will simply have

$$A_n = \delta_{n0}. \tag{10}$$

In other words, the action of the collision operator is to attempt to drive down all coefficients with $n>0$. Because we are multiplying by the Maxwell distribution, the Laguerre orthogonality property proves useful

$$\int_0^\infty x^\alpha e^{-x}L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)dx = \frac{\Gamma(n+\alpha+1)}{n!}\delta_{n,m}. \tag{11}$$

For example, if we choose $u=1$, conservation of particles and energy correspond to the simple identities

$$A_0 = 1 \ \text{[conservation of particles]}, \tag{12}$$

$$A_1 = 0 \ \text{[conservation of energy]}, \tag{13}$$

provided we make the choice $\alpha = 1/2$, as we have in (8). Because the QLB equation conserves particles and energy, the time derivatives of these two coefficients are identically zero, so if these identities hold for the initial distribution, then they hold for all times. The temperature that appears in the expansion is that of the final equilibrated state, which can easily be related to the total (kinetic) energy. The Lenard-Balescu equation also conserves momentum, but this is identically zero when we have isotropy in velocity.

Clearly, the expansion (8) with $u = 1$ has many advantages. However, we do pay some price for them. The Laguerre polynomials are orthogonal with respect to the weight $w(x) = x^{1/2}e^{-x}$, and for a function $f(x)$ to be representable by a series of these polynomials, it must be square integrable with respect to this weight, i.e.,

$$\int_0^\infty x^{1/2}e^{-x}|f(x)|^2 dx < \infty. \tag{14}$$

But because we actually have an expansion of the form

$$f(x) = e^{-x}\sum_n A_n L_n^{(\frac{1}{2})}(x), \tag{15}$$

we have the more stringent requirement that $e^x f(x)$ be square integrable, or

$$\int_0^\infty x^{1/2}e^{x}|f(x)|^2 dx < \infty. \tag{16}$$

Say for example, $f(x) = e^{-x/\gamma}$, then the integral (16) is

$$\int_0^\infty x^{1/2}e^{x(1-2/\gamma)}dx, \tag{17}$$

which converges only when $0 < \gamma < 2$. For the purposes of this work, the requirement that distributions fall off faster than $e^{-x/2}$ is not particularly problematic. We consider only equilibration problems, in which the end state is the Maxwell distribution, $A_n = \delta_{n0}$, and thus, if the initial distribution can be represented, then the subsequent evolution can as well. To be more precise, if (16) is satisfied for the initial time, then it is satisfied for all times. We will not prove this, but it seems very unlikely that the integral in (16) would be initially finite but then diverge as the distribution becomes more Maxwellian (it is, of course, finite for the Maxwell distribution itself). We will have more to say about this in Section X, where we show that choosing $u = 2$ in (8) restores completeness at the expense of complicating the collision integrals and conservation conditions.

To solve the equation, we truncate the expansion (8) at some $n_{\max}$, which will be as large as 40 in the present work. The ordinary differential equations that result are of the form

$$\frac{dA_n}{dt} = \sum_{l=0}^{n_{\max}}\sum_{k=0}^{n_{\max}} C_{lk}^n(\{A\})A_l A_k. \tag{18}$$

The coefficients C_{lk}^n are integrals over the dielectric function and depend on all the A_n, which we denote $\{A\}$, and therefore must be computed on the fly. We describe in Section V how we evaluate these coefficients, but first we turn to the dielectric function.

IV. DIELECTRIC FUNCTION

It is convenient to define the dimensionless variables

$$X^2 \equiv \frac{\hbar^2\beta k^2}{4m_e}, \tag{19}$$

$$Y^2 \equiv \frac{m_e\beta\omega^2}{k^2}, \tag{20}$$

in terms of which we will write all of our results. The non-equilibrium dielectric function is derived in Appendix A. This is given in terms of its expansion coefficients, like the distribution itself, where the first term is the non-degenerate equilibrium dielectric function and the terms containing A_n for $n > 1$ give the non-equilibrium part. In terms of X and Y, this expansion is

$$\epsilon(X, Y) = 1 + \frac{\eta^2}{X^3}w^Q(X, Y), \tag{21}$$

where $w^Q(X, Y) = w_r^Q(X, Y) + iw_i^Q(X, Y)$ is a complex function whose real and imaginary parts are given by

$$w_r^Q(X, Y) = \frac{1}{\sqrt{2}}\sum_{k=0}^{\infty}A_k\left[-Y_- M\left(k+1, \frac{3}{2}; -Y_-^2\right) + Y_+ M\left(k+1, \frac{3}{2}; -Y_+^2\right)\right], \tag{22}$$

$$w_i^Q(X, Y) = \sqrt{\frac{\pi}{2}}\frac{1}{2}\sum_k A_k \times\left[e^{-Y_-^2}L_k^{(-\frac{1}{2})}(Y_-^2) - e^{-Y_+^2}L_k^{(-\frac{1}{2})}\left(Y_+^2\right)\right], \tag{23}$$

where $M(a, b; z)$ is the confluent hypergeometric function, $Y_\pm \equiv (Y \pm X)/\sqrt{2}$, and

$$\eta \equiv \lambda_Q/\lambda_D \tag{24}$$

with

$$\lambda_Q^2 \equiv \frac{\hbar^2\beta}{4m_e}, \tag{25}$$

$$\lambda_D^2 \equiv \frac{1}{4\pi e^2 n_e\beta}. \tag{26}$$

Thus, η is the ratio of the equilibrium thermal de Broglie and Debye wavelengths (note that these are calculated in terms the final equilibrium temperature). The inverse of this ratio is usually denoted $\Lambda = 1/\eta$. At weak coupling, which is where the QLB equation is accurate, $\eta \ll 1$. We will exploit the smallness of η when we compute the coefficients.

Because it greatly simplifies the analysis without detracting from the important physics, we take the limit

$\hbar \to 0$ in the dielectric function. This is equivalent to expanding (21) in X,

$$w^Q(X,Y) \approx w^Q(0,Y) + Xw(Y), \tag{27}$$

where

$$w(Y) \equiv \left. \frac{\partial w^Q(X,Y)}{\partial X} \right|_{X=0}. \tag{28}$$

Clearly, $w^Q(0,Y)=0$ and we can compute $w(Y)$ from (28) by making use of the hypergeometric contiguous relation

$$z\frac{\partial M(a,b;z)}{\partial z} = (b-1)[M(a,b-1;z) - M(a,b;z)]. \tag{29}$$

The dielectric function is then

$$\epsilon_{\text{cl}}(X,Y) = 1 + \frac{\eta^2}{X^2} w(Y) \tag{30}$$

with the real and imaginary parts of $w(Y)$ given by

$$w_r(Y) = \sum_{k=0}^{\infty} A_k M\left(k+1, \frac{1}{2}; -\frac{Y^2}{2}\right), \tag{31}$$

$$w_i(Y) = \sqrt{\frac{\pi}{2}} Y e^{-\frac{Y^2}{2}} \sum_{k=0}^{\infty} A_k L_k^{(1/2)}\left(\frac{Y^2}{2}\right). \tag{32}$$

Neglecting quantum effects in the dielectric function probably does not impact the solution in a major way and, of course, we retain this physics everywhere else in the QLB equation. With only a few tens of parameters, namely, the A_k, to be determined numerically, this analytic representation of the dielectric function is very convenient. We can, for example, use it to determine the dispersion relation of waves in non-equilibrium plasmas. We will see how this form is also useful in the numerical solution of the QLB equation, despite the presence of the confluent hypergeometric function.

Note that we can, if we wish, simplify the problem even further by considering only static screening and setting $Y=0$ in the dielectric function. The result is

$$\epsilon_{\text{static}}(X) = 1 + \frac{\eta^2}{X^2} \sum_{k=0}^{\infty} A_k. \tag{33}$$

In this form, we no longer have dynamical screening effects, but the static screening length is still calculated from the distribution.

V. COEFFICIENTS

The coefficients of Equation (18) are computed by multiplying the equation by a Laguerre polynomial and integrating over velocity. The details of this are given in Appendix B. The result is

$$C_{lk}^n = -C_0 \frac{n!}{\Gamma(n+3/2)} \int_0^{\infty} dX \frac{e^{-X^2}}{X^3} \int_{-\infty}^{\infty} dY \frac{e^{-Y^2}}{|\epsilon(X,Y)|^2} P_{lk}^n(X,Y), \tag{34}$$

where the prefactor is

$$C_0 \equiv \frac{n_e \beta^{3/2} \sqrt{\pi} e^4}{\sqrt{m_e}}. \tag{35}$$

The functions $P_{lk}^n(X,Y)$ are polynomials in X and Y defined by

$$P_{lk}^n(X,Y) = \frac{1}{2}\left[q_{lk}^n(X,Y) + q_{lk}^n(X,-Y)\right], \tag{36}$$

where

$$q_{lk}^n(X,Y) \equiv \sum_{j=0}^{\min(l,n)} L_{n-j}^{(-\frac{1}{2})}(Y_-^2) \times [L_k^{(-\frac{1}{2})}(Y_+^2) L_{l-j}^{(-\frac{1}{2})}(Y_-^2) - L_k^{(-\frac{1}{2})}(Y_-^2) L_{l-j}^{(-\frac{1}{2})}(Y_+^2)], \tag{37}$$

so P is just the even part of q in Y. It is not really necessary to take the even part explicitly because the integration over Y filters out the odd powers, but we do it to make the following analysis more clear. Note also that the symmetry of q in X and Y means that P_{lk}^n contains only even powers of X and it turns out that X^2Y^2 is the lowest power for all n, l, k. To facilitate our approximations, we use the decomposition

$$P_{lk}^n(X,Y) = X^2Y^2G_{lk}^n(Y) + R_{lk}^n(X,Y), \tag{38}$$

where $G_{lk}^n(Y)$ is a polynomial in Y and $R_{lk}^n(X,Y)$ contains terms only of order X^4 and higher. Now, the X integration in (34) would of course be divergent as $X \to 0$ were it not for the dielectric function. However, only the first term in (38) would actually diverge. Our decomposition is therefore a separation into the term that needs the dielectric function for convergence, and the rest of the integrand that does not. From here on, we will keep the dielectric function only where it is actually needed for convergence and set it to 1 elsewhere. This approximation can be justified as follows.

Physical parameters, such as mass and density, enter into the coefficients in (34), and therefore the equation, in two places: the prefactor C_0 and the dimensionless ratio η. The constant C_0 only sets the overall time scale of the problem, and the two solutions with the same η and initial distribution but different C_0 will be identical up to time rescaling. Therefore, the only really important quantity is η, and varying things like the particle mass, the number density, and the final equilibrium temperature only matters to the extent that we are changing η. As such, as previously mentioned, we stick to electrons at 10^{25}cm^{-3}. The latter choice makes η similar with the coupling constant, Γ, as we vary the temperature. As we will show later, the expansion of C_{lk}^n in η is

$$C_{lk}^n = a_0 + \sum_{i=0}^{\infty} b_i \eta^{2i} \ln \eta + \sum_{i=1}^{\infty} a_i \eta^{2i}. \tag{39}$$

The term proportional to b_0 is the Coulomb logarithm and a_0 is the order-unity term; if η is small, we may be justified in neglecting the rest. And if η is smaller still, the logarithmic term will dominate a_0 and the Landau equation is fine. However, for arbitrary non-equilibrium initial conditions, the

$\mathrm{O}(1)$ terms depend on the distribution and must be computed before we can be sure they can be neglected, making the definition of "small" for η highly problem-dependent. Our strategy of keeping the dielectric function only where it is necessary for convergence is equivalent with computing a_0 and b_0 and dropping the rest. Thus, we neglect terms that are $\mathrm{O}(\eta^2 \log \eta)$ and higher, which does not make a great difference in many cases. For example, at a density of $10^{25}\mathrm{cm}^{-3}$ at $T = 1000\,\mathrm{eV}$, $\eta \approx 0.05$, and $\eta^2 \ln\eta \approx -7.5 \times 10^{-3}$, compared with the term $\ln\eta \approx -3$ and other $\mathrm{O}(1)$ terms that we are going to keep. We discuss below some situations where one might need the higher-order terms, but we will not be concerned about computing them in this paper. In any case, it is a straightforward generalization to include them (see Section X), but, of course, this becomes more computationally expensive.

Under this approximation, the integrals over R^n_{lk}, as they do not contain the A_k, can be precomputed. We define the coefficients

$$B^n_{lk} \equiv \frac{n!}{\Gamma(n+3/2)}\int_0^\infty dX \frac{e^{-X^2}}{X^3}\int_{-\infty}^{\infty} dY e^{-Y^2} R^n_{lk}(X,Y). \tag{40}$$

Another set we will need is

$$S^n_{lk} \equiv \frac{n!}{\Gamma(n+3/2)}\int_{-\infty}^{\infty} dY e^{-Y^2} Y^2 G^n_{lk}(Y)\,. \tag{41}$$

Even though these coefficients can be precomputed, doing so is not completely trivial. As n, l and k become large, R^n_{lk} becomes higher-order in X and Y. For example, at $(n,l,k) = (40,40,40)$, the most difficult case, $R^n_{lk}(X,Y)$ is order 234 in X and Y. If we wish to use a numerical integration scheme for this, we need to evaluate $R^n_{lk}(X,Y)$ at the quadrature points, which can prove to be tricky with such high order polynomials. There is probably an optimal solution to this problem, but we resort to brute force. We use the CLN arbitrary precision library[28] for C++ and we decompose the polynomial $R^n_{lk}(X,Y)$ into its powers,

$$R^n_{lk}(X,Y) = \sum_{ij} a^n_{lkij} X^{2i} Y^{2j}, \tag{42}$$

where the sums over i and j start at $i = j = 2$. Using the exact integrals

$$\int_0^\infty e^{-X^2} X^{2i-3} dX = \frac{1}{2}\Gamma(i-1) \tag{43}$$

and

$$\int_{-\infty}^{\infty} e^{-Y^2} Y^{2j} dY = \Gamma\left(\frac{1}{2}+j\right), \tag{44}$$

where $\Gamma(x)$ is the gamma function, combined with the decomposition (42) allows us to evaluate (40) so long as we have sufficient precision; we keep 60 digits for this purpose. Of course, we do not need this many when we solve the actual differential equation, so in the end we keep the resulting B^n_{lk} only to double precision. The constants S^n_{lk} can be handled in the same way but in Equation (57) we give the exact solution for these.

Now we are left with the problem of evaluating

$$I^n_{lk} \equiv \int_0^\infty dX \frac{e^{-X^2}}{X}\int_{-\infty}^{\infty} dY \frac{e^{-Y^2}}{|\epsilon(X,Y)|^2} Y^2 G^n_{lk}(Y), \tag{45}$$

which must be computed on the fly. The strategy is to compute the X integral exactly, which would hardly be possible if we were not using the classical dielectric function. The remaining one-dimensional integral over Y will contain a tangle of special functions, but the integrand is smooth and can easily be handled with a straightforward Gaussian quadrature. The steps required to reduce (45) are given in Appendix C, with the result

$$I^n_{lk} = -\frac{\Gamma(n+3/2)}{n!}\left[S^n_{lk}\left(\frac{\gamma_E}{2}+\ln\eta\right)+F^n_{lk}\right], \tag{46}$$

where

$$F^n_{lk} \equiv \frac{n!}{\Gamma(n+3/2)}\frac{1}{2}\int_{-\infty}^{\infty} dY e^{-Y^2} Y^2 G^n_{lk}(Y) F(Y), \tag{47}$$

and

$$F(Y) \equiv \frac{1}{2}\ln\left[w_r^2(Y)+w_i^2(Y)\right] + \frac{w_r(Y)}{w_i(Y)}\arctan\left[w_r(Y), w_i(Y)\right], \tag{48}$$

where $\arctan(x,y)$ is the quadrant-correct version of $\tan^{-1} y/x$, producing an angle in the range $(-\pi,\pi]$. The integrand in (45) is well-behaved, without any of the sharp peaks that typically characterize dielectric function integrands, and we avoid the need for any pole-correcting integration techniques.[16,17] The same strategy was used by Williams and DeWitt[5] for conductivity calculations in a two-component plasma in equilibrium. Although a very different problem from ours, it involves the same collision operator and the same kinds of integrals (compare their Equation (73) with our (48)). In fact, this method would be useful for other problems as well, such as temperature equilibration.[9,10]

Although we now have a one-dimensional integral, we are still faced with the task of evaluating it at every time step. The factor e^{-Y^2} in the integrand strongly suggests we use Gauss-Hermite quadrature. Actually, because the integrand is even, we make the substitution $x = Y^2$ and use a closely related Gauss-Laguerre scheme. We then have

$$\begin{aligned}\int_{-\infty}^{\infty} dY e^{-Y^2} Y^2 G^n_{lk}(Y) F(Y) &= \int_0^\infty dx x^{1/2} e^{-x} G^n_{lk}(\sqrt{x}) F(\sqrt{x}) \\ &\approx \sum_{i=1}^{N_p} W_i G^n_{lk}(\sqrt{x_i}) F(\sqrt{x_i}),\end{aligned} \tag{49}$$

where x_i are the abscissa points, the zeros of $L^{(1/2)}_{N_p}(x)$, and W_i are the weights, given by

$$W_j = \frac{x_j \Gamma(N_p+1/2)}{N_p!(N_p+1/2)\left[L^{(1/2)}_{N_p-1}(x_j)\right]^2}. \tag{50}$$

We choose $N_p = 200$ to ensure that we have an accurate integration even for the largest n, l, and k. This number can probably be varied to optimize performance, and it may not always be necessary to include every term in (50), especially for the smaller (n, l, k). We do not explore this particular performance issue too closely, but, as we will show, this scheme is more than sufficiently accurate for our purposes and enables a numerical solution of the Lenard-Balescu equation. What is more, because our integration is now simply a sum over quadrature points, the values of the confluent hypergeometric functions and Laguerre polynomials that appear in $w_r(\sqrt{x})$ and $w_i(\sqrt{x})$ can be precomputed at x_i and never need to be evaluated during the solution. We will also precalculate the points $G^n_{lk}(\sqrt{x_i})$. This is somewhat tricky because, exactly like R^n_{lk}, as (n, l, k) become large, the order of this polynomial also becomes large and we require high precision to evaluate it. To do this, we use a strategy similar to the one we employed to calculate B^n_{lk}. First, we compute the quadrature points, x_i, to high precision using Mathematica. Then, we decompose $G^n_{lk}(Y)$ as in (42) and evaluate each power of Y at $\sqrt{x_i}$ to 60 digits and sum these results to get $G^n_{lk}(\sqrt{x_i})$. As before, we keep these values only to double precision, so no arbitrary precision library is needed in the actual solver.

To demonstrate the accuracy of the quadrature scheme, we will compute F^n_{lk} for $(n, l, k) = (40, 40, 40)$, the most difficult case. Of course, we must also specify a distribution, and for this, we use a two-temperature plasma in which half the particles are at temperature T_1 and the other half at T_2, which is shown in the top panel of Figure 1. The coefficients for this distribution are calculated in Section VII and are characterized by the single parameter $\gamma \equiv T_1/T$, where T is the final temperature, which we choose to be 0.2, the limit of our resolution ability. Using Mathematica's adaptive numerical integration with 100-digit precision, we find that

$$\int_{-\infty}^{\infty} dY e^{-Y^2} Y^2 G^{40}_{40,40}(Y) F(Y) \approx 2.1563096073. \qquad (51)$$

The calculation takes several minutes but is accurate to the number of digits presented. For this integral, our double precision quadrature code gives 2.1563096082, correct to eight decimal places, far more than we need, and is just a sum over 200 points. Of course, the accuracy depends on the distribution and we may not always achieve this level. For example, consider $\gamma = 0.08$, which is badly under-resolved when $n_{\max} = 40$. However, the Laguerre series is positive everywhere, as shown in the bottom panel of Figure 1, and thus is an acceptable distribution. Done with adaptive integration in high precision, the integral (51) is 2.9349261394. With our Gaussian quadrature scheme, we find 2.9323983879, which is not disastrous but not nearly as accurate as in the previous example, probably due to the oscillations in the distribution. Adding more quadrature points would probably improve the accuracy, but computing the integral to three figures is sufficient for our purposes.

We have a fast and accurate method for integrating over the dielectric function, but the price we pay for this is that we must keep a huge number of precalculated values; if we want to use $n_{\max} = 40$ coefficients in the polynomial expansion and $N_p = 200$ quadrature points, the file containing the $G^n_{lk}(\sqrt{x_i})$ is 287 megabytes. While this is manageable enough, when we consider that the number of coefficients needed grows as $n^3_{\max}$ and that we will need more quadrature points as we increase (n, l, k), it is clear that this can quickly grow out of control. However, the present approach is surely the brute-force method to compute C^n_{lk}, and there are likely better ways to do this. For example, let us define V^j_{lk} to be the integral over triple products of Laguerre polynomials

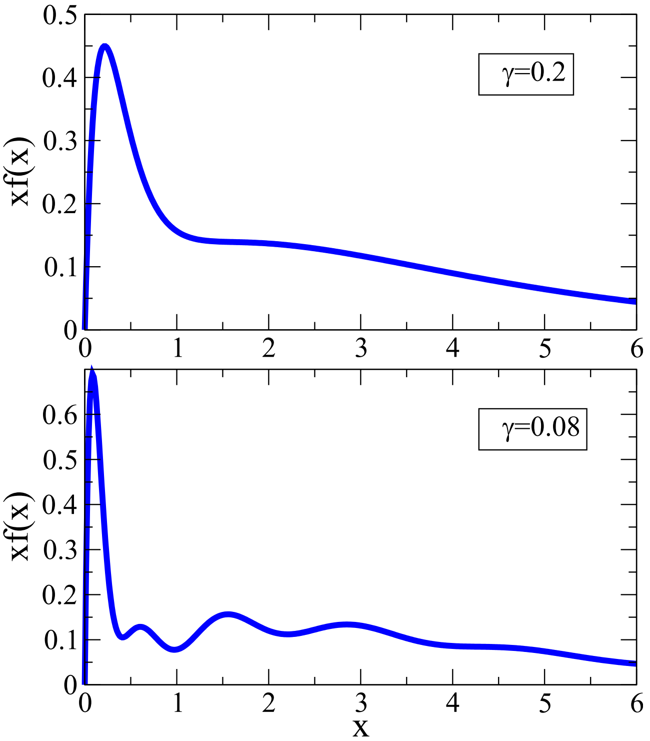


FIG. 1. Two-temperature initial distribution using $n_{\max} = 40$ polynomials with $\gamma = 0.2$ (top) and $\gamma = 0.08$ (bottom). The $\gamma = 0.08$ case is under-resolved with $n_{\max} = 40$, and the bumps are mostly a result of this, but the distribution is positive everywhere and properly normalized and is therefore suitable as an initial condition. The variable $x = \beta m_e v^2/2$ and $f(x)$ is divided by $n_e(\beta m_e/2\pi)^{3/2}$.

$$V^j_{lk} \equiv \int_{-\infty}^{\infty} dY \int_0^{\infty} dX \frac{e^{-Y^2}}{|\epsilon(X, Y)|^2} \frac{e^{-X^2}}{X^3} \times L_j^{(-1/2)}(Y_-^2) L_k^{(-1/2)}\left(Y_+^2\right) L_l^{(-1/2)}\left(Y_-^2\right). \qquad (52)$$

Equation (34) for the coefficients can then be written as

$$C^n_{lk} = -C_0 \frac{n!}{\Gamma(n+3/2)} \sum_{j=0}^{\min(l,n)} \left(V^{n-j}_{l-j,k} - V^{n-j}_{k,l-j}\right). \qquad (53)$$

V^j_{lk} satisfies a recurrence formula that can possibly be exploited to facilitate computation of C^n_{lk} without needing huge files of precomputed data. In Appendix E, we derive this formula and show that it has an exact solution. It may well be that such a method is superior once certain mathematical issues are resolved.

Putting the pieces of the present method together, we find for the coefficients

$$C_{lk}^{n} = C_0\left[S_{lk}^{n}\left(\frac{\gamma_E}{2}+\ln\eta\right) - B_{lk}^{n} + F_{lk}^{n}(\{A\})\right], \tag{54}$$

which, as mentioned, neglects terms that are $O(\eta^2 \ln\eta)$. However, we have derived the Coulomb logarithm, rather than imposing it, along with all $O(1)$ terms that arise from the quantum Lenard-Balescu equation. These are by far the dominant contributions to the equation for the situations we will consider. On the other hand, as we can see from Equation (C6) in Appendix C, the expansion of the incomplete gamma function that leads to (54) is not in η^2 but in $\eta^2 w(Y)$. It is conceivable that some distributions might make $|w(Y)|$ comparable to η^{-2} and then we would not be justified in discarding these terms. For example, if we have a two-temperature initial condition, we can make $|w(0)|$ as large as we want by increasing the temperature separation. We will not encounter such an extreme situation here but, as mentioned, we discuss how to restore these terms in Section X.

VI. SPECIAL CASES

By neglecting different terms in (54), we can find solutions to various kinetic equations. These are listed here.

A. Landau equation

If we drop the quantum diffraction terms B_{lk}^{n} and the screening terms F_{lk}^{n}, we are left with the coefficients of the Landau equation

$$C_{lk}^{n} = C_0 S_{lk}^{n}\left(\frac{\gamma_E}{2}+\ln\eta\right), \tag{55}$$

$$= -C_0 S_{lk}^{n} \log\Lambda. \tag{56}$$

In the first line, we write the Coulomb logarithm informed by the Lenard-Balescu equation, but one is free to make any choice of $\log\Lambda$ one wishes; the integrals in this equation are divergent and the form of Λ results from the choice of cutoffs. The coefficients C_{lk}^{n} are all precomputed so nothing needs to be calculated on the fly. This makes the solution extremely cheap compared with the full QLB equation. Another fact worth mentioning is that the coefficients S_{lk}^{n} actually have a closed form in terms of hypergeometric functions

$$\begin{aligned} S_{lk}^{n} = {} & \frac{\sqrt{\pi}}{2\sqrt{2}}\frac{n!}{\Gamma(n+3/2)}\frac{1}{4^k}\frac{1}{n!l!k!}\Gamma\left(\frac{3}{2}-n+k+l\right) \\ & \times \Gamma\left(-\frac{1}{2}+n+k-l\right)\{16\,{}_3\tilde{F}_2(1,-l,n;b_1^1,b_2^1;1) \\ & -[(2k-2l+2n)^2-1]\,{}_3\tilde{F}_2(1,-l,n;b_1^2,b_2^2;1)\}, \end{aligned} \tag{57}$$

where

$$b_1^1 = -\frac{1}{4}+\frac{k-l-n}{2}, \tag{58}$$

$$b_2^1 = \frac{1}{4}+\frac{k-l-n}{2}, \tag{59}$$

$$b_1^2 = \frac{3}{4}+\frac{k-l-n}{2}, \tag{60}$$

$$b_2^2 = \frac{5}{4}+\frac{k-l-n}{2}, \tag{61}$$

and ${}_3\tilde{F}_2(a_1,a_2,a_3;b_1,b_2;z)$ is a regularized hypergeometric function. The latter is defined by

$${}_3\tilde{F}_2(a_1,a_2,a_3;b_1,b_2;z) \equiv \frac{{}_3F_2(a_1,a_2,a_3;b_1,b_2;z)}{\Gamma(b_1)\Gamma(b_2)}, \tag{62}$$

where ${}_pF_q(a_1\ldots,a_p;b_1,\ldots,b_q;z)$ is the generalized hypergeometric function. These expressions may not seem terribly convenient. However, Mathematica, and probably other similar programs, quickly evaluates them and easily handles the differential equations too. This prescription therefore provides a fast and convenient way to solve the single-component Landau equation. The derivation of (57) is given in Appendix D.

B. Non-degenerate quantum Landau equation

Setting $F_{lk}^{n}=0$ in (54) neglects the dielectric function, but we still have the quantum wave effects embodied in the B_{lk}^{n} and we end up with the coefficients for what we call the non-degenerate quantum Landau equation. The reason for this ungainly term is that "quantum Landau equation" is already in use[29,30] for a kinetic equation that accounts for quantum statistics but no other quantum effects, which is sort of the complement of our equation. The integrals in this equation are divergent as $X\to 0$ but converge as $X\to\infty$, meaning that we need only a lower cutoff. This will, of course, generally be chosen to be the equilibrium Debye length, or $X_c=\eta$ in our dimensionless variables. The resulting coefficients are

$$C_{lk}^{n} = C_0\left[S_{lk}^{n}\left(\frac{\gamma_E}{2}+\ln\eta\right) - B_{lk}^{n}\right], \tag{63}$$

which can once again all be precalculated.

C. Classical Lenard-Balescu equation

Finally, we can set $B_{lk}^{n}=0$ but retain F_{lk}^{n}. The coefficients are then

$$C_{lk}^{n} = C_0\left[S_{lk}^{n}\left(\frac{\gamma_E}{2}+\ln\eta\right) + F_{lk}^{n}(\{A\})\right]. \tag{64}$$

These correspond to the classical Lenard-Balescu equation, in which we cure the $X\to\infty$ divergence by introducing a cutoff in wavenumber at the inverse of the thermal deBroglie wavelength, λ_Q in Equation (25), or $X_c=1$ in the dimensionless units. Of course, one can instead cut the integral off at the Landau length to keep everything classical.

D. Quantum Lenard-Balescu with static screening

The dielectric function that neglects dynamical screening is given in Equation (33). This corresponds to the choices

$$w_r(Y) = \sum_{k=0}^{\infty} A_k, \qquad w_i(Y) = 0. \tag{65}$$

Using these in Equation (48), we find

$$F(Y) = \ln \sum_{k=0}^{\infty} A_k + 1, \tag{66}$$

leading to the coefficients

$$C_{lk}^n = C_0 \left\{ \frac{1}{2} S_{lk}^n \left[\gamma_E + 1 + \ln\left(\eta^2 \sum A_k\right) \right] - B_{lk}^n \right\}. \tag{67}$$

We can see that static screening contributes an additional constant (i.e., one) and modifies the Coulomb logarithm by the sum over A_k. This provides a correction to the Debye length and is trivial to compute.

VII. INITIAL CONDITIONS

To test our algorithm, we consider the relaxation to equilibrium of various initial distributions. For general u, the coefficients $A_n(0)$ for a given $f(v,0)$ are

$$A_n(0) = \frac{2\pi^{3/2}u^{3/2}}{n_e} \frac{n!}{\Gamma(n+3/2)} \times \int_0^{\infty} e^{-\frac{\beta m v^2}{2}(u-1)} v^2 f(v,0) L_n^{(\frac{1}{2})}\left(u \frac{m_e \beta v^2}{2}\right) dv, \tag{68}$$

which is an easy consequence of the orthogonality property of Laguerre polynomials.

A. Two-temperature plasma

Here, we will consider the case of a two-temperature one-component plasma. A number density n_1 have temperature T_1 and n_2 have T_2 so,

$$f(v,0) = \left(\frac{m_e}{2\pi}\right)^{\frac{3}{2}} \left[n_1 \beta_1^{3/2} \exp\left(-\frac{m_e \beta_1 v^2}{2}\right) + n_2 \beta_2^{3/2} \exp\left(-\frac{m_e \beta_2 v^2}{2}\right) \right]. \tag{69}$$

We define the fractions $\xi \equiv n_1/n_e$, $\xi_2 \equiv n_2/n_e$, $\gamma \equiv T_1/T$, and $\gamma_2 \equiv T_2/T$. By conservation of particles and energy, we have

$$\xi_2 = 1 - \xi, \tag{70}$$

$$\gamma_2 = \frac{1 - \xi\gamma}{1 - \xi}. \tag{71}$$

We make the choice $\gamma < 1$, so that $\gamma_2 > 1$. Carrying out the integration (68), we find

$$A_n(0) = \frac{\xi(1-\gamma)^n u^{3/2}}{(\gamma u - \gamma + 1)^{n+3/2}} + \frac{(1-\xi)(1-\gamma_2)^n u^{3/2}}{(\gamma_2 u - \gamma_2 + 1)^{n+3/2}}. \tag{72}$$

For $u = 1$, which is what we use exclusively here, the condition (16) means that γ_2 must be less than 2 or the expansion does not converge, which is also clear enough in (72). This is, of course, a purely mathematical requirement and it leads to the constraint between ξ and γ,

$$\gamma > \frac{2\xi - 1}{\xi}. \tag{73}$$

If we needed to break this we would choose a different value of u, such as 2. However, (73) is no constraint if $\xi \leq 1/2$ so if, for example, we have an equal number of particles of each temperature, then γ can be chosen arbitrarily in the range $[0, 1]$ with $\gamma_2 < 2$ enforced by conservation of energy. This is the first situation we will consider.

With $\xi = 1/2$, the coefficients $A_n(0)$ for $\gamma = 0.2$ and $u = 1$ are shown in Figure 2. It is a mathematical peculiarity that, for this situation, every odd coefficient is zero. In an equilibration problem, the distribution becomes more Maxwellian with time, and thus, the initial condition is probably the most difficult thing to resolve. In other words, one will not need more polynomials at a later time than at the beginning. To get an idea of the number needed for a two-temperature system, we consider the $n_{\max}$ at which we first have $A_{n_{\max}} < \delta$. For this problem,

$$N = \frac{\log \delta}{\log(1-\gamma)}. \tag{74}$$

Shown in Figure 3 is a plot of $N(\gamma)$ using $\delta = 10^{-3}$. This is a somewhat arbitrary choice, to be sure, but it provides a useful rule of thumb. From this plot, it is clear that for $n = 40$, which is our maximum, one would not want to go far below $\gamma = 0.2$. We can, of course, invert (74) to estimate the minimum γ for a given N

$$\gamma = 1 - \delta^{1/N}. \tag{75}$$

We should point out that although the initial A_n depend only on the ratio of the initial to the final temperature, their

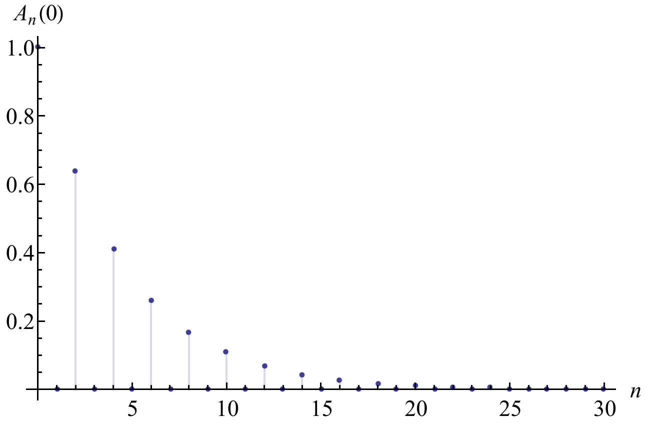


FIG. 2. The coefficients, $A_n(0)$, for the two-temperature initial condition with $\gamma = 0.2$.

subsequent values will depend on the absolute temperature through the dependence of the coefficients on η and the prefactor.

B. Gaussian distribution

We consider here the initial distribution

$$f(v) = Be^{-\frac{(v-v_0)^2}{2\sigma^2}}, \tag{76}$$

for demonstration purposes, not because we have a particular application in mind. The amplitude, B, and variance, σ^2, can be related to the number density and the energy by first defining the integrals

$$I_1(\bar{v}_0) \equiv \int_0^\infty v^2 e^{-(v-\bar{v}_0)^2} dv, \tag{77}$$

$$I_2(\bar{v}_0) \equiv \int_0^\infty v^4 e^{-(v-\bar{v}_0)^2} dv, \tag{78}$$

with $\bar{v}_0 \equiv v_0/\sqrt{2}\sigma$. We then have

$$\sigma^2 = \frac{3I_1(\bar{v}_0)}{2I_2(\bar{v}_0)\beta m_e}, \tag{79}$$

$$B = \frac{n_e}{4\pi(\sqrt{2}\sigma)^3 I_1(\bar{v}_0)}. \tag{80}$$

The integrals (77) and (78) can be expressed in terms of special functions, but they are easily evaluated numerically for a given $\bar{v}_0$. This parameter is the only one on which the A_n depend. For general u, these are given by

$$\begin{aligned} A_n = {} & \sqrt{\frac{\pi}{2}}\frac{[I_2(\bar{v}_0)]^{3/2}}{3^{3/2}[I_1(\bar{v}_0)]^{5/2}}\frac{n!}{\Gamma(n+3/2)} \\ & \times \int_0^\infty e^{-\left(1-\frac{1}{u}\right)x-\left(\sqrt{2I_2(\bar{v}_0)/[3I_1(\bar{v}_0)u]}\,x^{1/2}-\bar{v}_0\right)^2} \\ & \times L_n^{\left(\frac{1}{2}\right)}(x)x^{1/2}dx. \end{aligned} \tag{81}$$

Although the integral can be written exactly in terms of Hermite polynomials using formula 7.374.9 of Gradshteyn and Ryzhik,[31] we just solve it numerically. The parameter $\bar{v}_0$ determines the number of polynomials needed to resolve the distribution. Being limited to 40 polynomials, we find that we can choose $\bar{v}_0$ no larger than 2. Exactly what this means in terms of absolute velocity depends on the values of the other parameters, such as n_e and β.

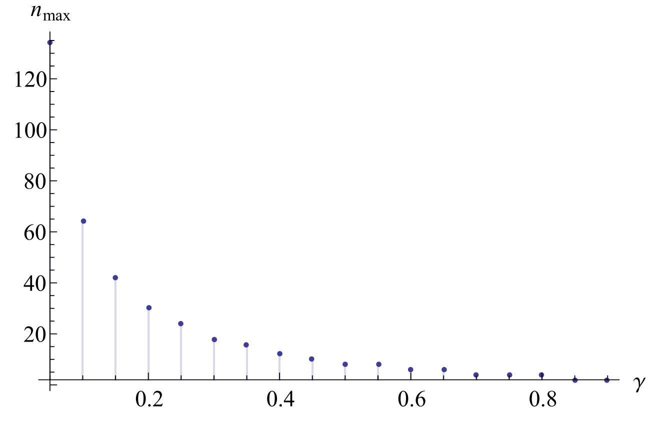


FIG. 3. The number of polynomials needed such that $A_{n_{\max}} < 10^{-3}$ as a function of γ for a two-temperature plasma. The left-most point corresponds to $\gamma = 0.01$. In this paper, we have $n_{\max} = 40$, and thus, we begin to have resolution problems when $\gamma < 0.2$.

VIII. DIFFERENTIAL EQUATIONS

The ordinary differential equations do not turn out to be very difficult to solve. We use the fifth-order Runge-Kutta scheme with adaptive time step implemented in the Boost library,[32] which easily handles the problem. For the Landau equation, and any of the others for which the coefficients can be precalculated, the solution is found more or less instantaneously using $n_{\max} = 40$ polynomials. For the Lenard-Balescu equation, the story is different and a solution can take several hours, but the bulk of the work is in the computation of the C^n_{lk} with the equation itself not being any more difficult than the other cases. This can be easily sped up with parallel computation; each processor computes every C^n_{lk} for a different range of n. The results are then shared, and the equation can be solved on a single processor. This scheme scales essentially perfectly with the number of processors, and in practice, we generally assign one n to each processor.

IX. RESULTS

A. Comparison with Fokker-Planck solution

The first thing we wish to do is check that our approach is sound by comparing our solution to the Landau equation with the result of a more traditional discretized solution to the Fokker-Planck equation. Data for this were provided by David Michta using a code he developed to study thermonuclear burn.[33] This approach uses discretization in velocity that is designed to ensure conservation of particles[11] and energy,[12] non-trivial problems in discretization schemes. The situation we considered was a two-temperature one-component plasma of electrons at a density of $2 \times 10^{25}\,\mathrm{cm}^{-3}$. Half the particles are Maxwellian at 500 eV, half are at 1500 eV, and we use a Coulomb logarithm $\log \Lambda = 1$. In Figure 4, we plot the distribution at the initial time and at two later times for both our solution using 20 polynomials and the Fokker-Planck result. The two solutions are completely indistinguishable from one another, indicating that, at least as far as the coefficients S^n_{lk} and the Landau equation go, our computations are correct.

B. Two-temperature plasma

Shown in Figure 5 is the numerical solution of the quantum Lenard-Balescu equation for the two-temperature plasma with $\gamma = 0.2$, $n_e = 1 \times 10^{25}\,\mathrm{cm}^{-3}$ and $T = 1000\,\mathrm{eV}$. All the coefficients except for A_0, which is fixed at 1, approach zero as $t \to \infty$, exactly as expected. The even coefficients fall monotonically while the odd coefficients, which start at zero, all become negative (except A_1 of course) before reaching a minimum and decaying back to zero. The distribution itself is shown in Figure 6. At 1000 eV, the solutions of the Landau

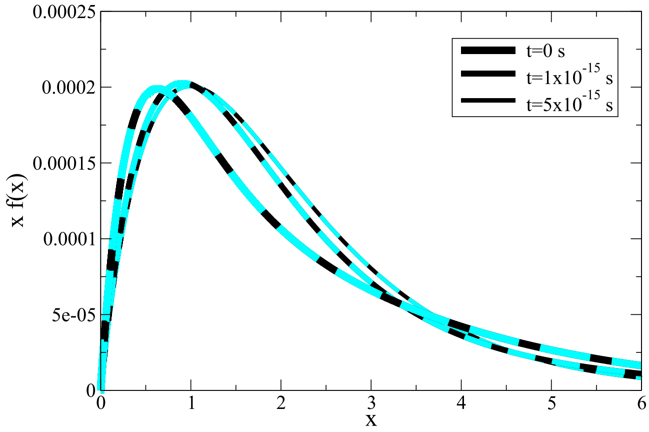


FIG. 4. Comparison between distributions calculated with our spectral solution of the Landau equation (solid black) and a discretized Fokker-Planck solution (dashed cyan). The two are indistinguishable for all times, three of which are shown.

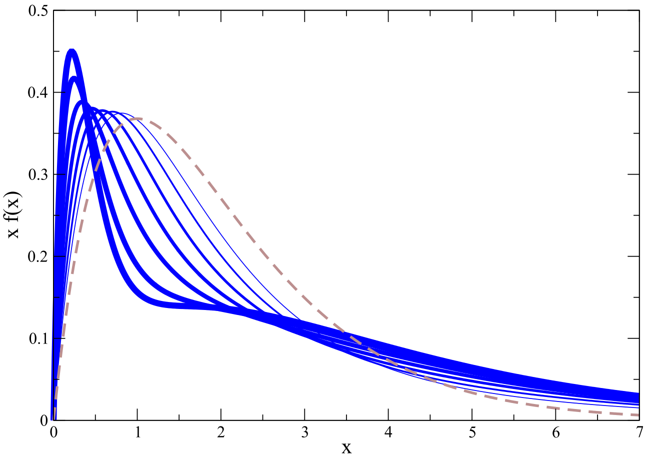


FIG. 6. Evolution of the distribution function under the quantum Lenard-Balescu equation for a two-temperature initial condition with $\gamma = 0.2$ and $T = 1000$ eV. The thickest solid line is $t = 0$, and the dashed line is equilibrium. The variable $x = \beta m_e v^2/2$ and $f(x)$ is divided by $n_e(\beta m_e/2\pi)^{3/2}$.

and quantum Landau equations are essentially the same as Figures 5 and 6, indicating that the order-unity terms are not playing much role. This is not completely obvious since $\ln\eta$ is only around -3. As we reduce the magnitude of $\ln\eta$, which we do by turning down T, we can begin to see slight differences between the Landau and Lenard-Balescu solutions, although almost no difference is ever in evidence between the Landau and quantum Landau equations. Shown in Figures 7 and 8 are the solutions of the Landau and quantum Lenard-Balescu equations for $T = 600$ eV, so $\eta \approx -2.3$; the evolution of the coefficients is noticeably different in the two cases. However, a comparison for the distribution itself is shown in Figure 9, and the differences between the Landau and quantum Lenard Balescu equations are modest at these conditions to say, the least. We cannot turn the temperature down much further without having numerical problems in the solver, an indication that our neglect of higher-order terms in η^2 is becoming problematic. However, even at $T = 600$ eV, the low-temperature electrons are at 120 eV and 5×10^{24}cm^{-3} and are becoming degenerate ($\theta \approx 1.1$). Thus, for this particular type of initial condition, our physical assumptions break down before we see any real advantage to carrying out the expensive integration over the dielectric function. On the other hand, we stress that any conclusions about where the Lenard-Balescu and Landau solutions become different are highly dependent on the initial distribution and we should not overestimate the generality of this particular example. It is certainly the case that by separating the temperature more widely, which we cannot do with only 40 polynomials, we would find ever greater divergence in the two solutions. In Section IX C, we find an initial distribution for which the two solutions are different.

Our method also allows a detailed view of the dielectric function in the random phase approximation, something that would not be easy to obtain with a discretization method. Figures 10 and 11 show the time evolution of the real and imaginary parts of the free-particle response function for the two-temperature initial condition with $\gamma = 0.2$, $T = 1000$ eV,

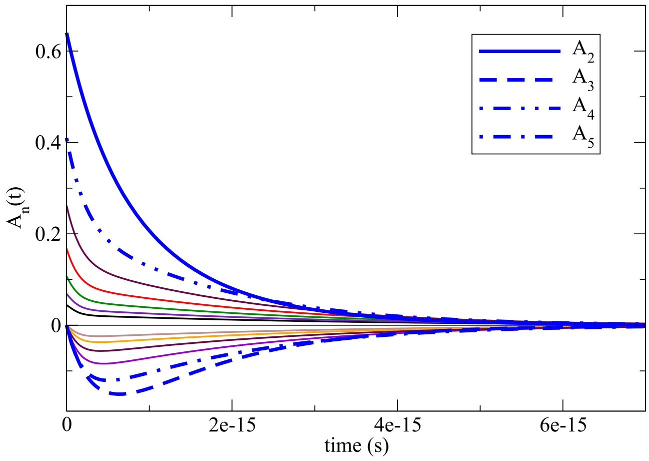


FIG. 5. Evolution of the expansion coefficients for the quantum Lenard-Balescu equation for a two-temperature initial condition with $\gamma = 0.2$ and $T = 1000$ eV. The even coefficients, which all start out positive, decay monotonically to zero with the exception of A_0 which is fixed at 1. The odd coefficients start at zero and, aside from A_1, become negative and then decay to zero. The first four non-trivial coefficients are labelled and the rest up to $n = 14$ are shown in various colors. The distribution itself is shown in Figure 6.

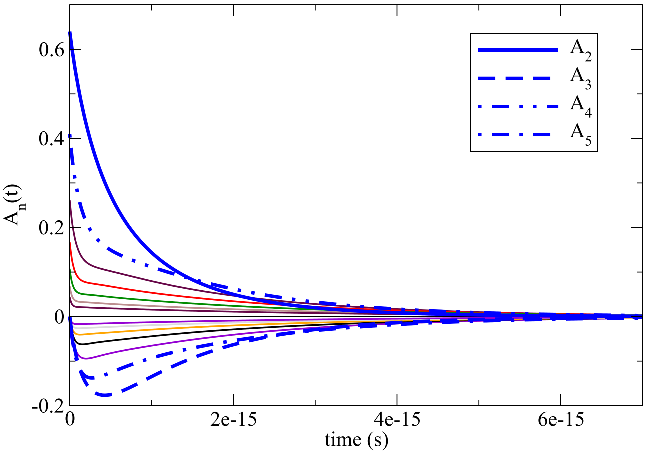


FIG. 7. Evolution of expansion coefficients by the Landau equation for $\gamma = 0.2$ and $T = 600$ eV. The first four non-trivial coefficients are labelled, and the rest up to $n = 14$ are shown in various colors.

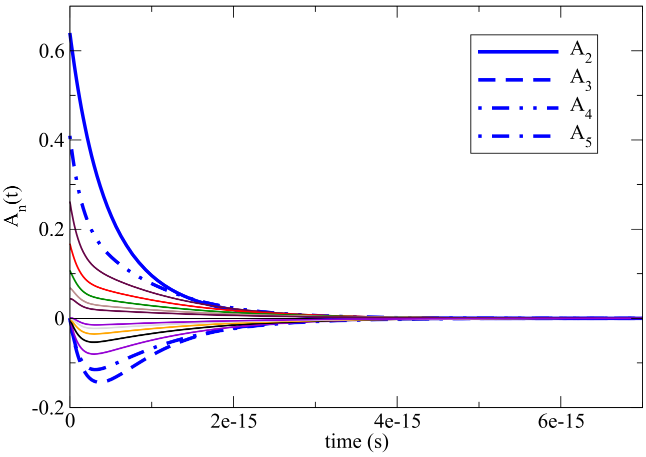


FIG. 8. Evolution of expansion coefficients by the quantum Lenard-Balescu equation for $\gamma = 0.2$ and $T = 600\,\mathrm{eV}$. The first four non-trivial coefficients are labelled and the rest up to $n = 14$ are shown in various colors.

and $n_e = 1 \times 10^{25}\mathrm{cm}^{-3}$. These are easily obtained from the coefficients and Equations (31) and (32).

C. Under-resolved two-temperature plasma

Here, we use the two-temperature initial condition but choose $\gamma = 0.08$, which is much too small for 40 polynomials. However, as Figure 1 shows, even though this distribution is badly under-resolved, it is still positive everywhere and thus constitutes a viable initial condition. To solve this problem, we keep the first 34 polynomials for the initial condition and set the remaining seven to zero. This way we still maintain a positive distribution, but we have a few modes above our resolved range to ensure we have sufficient resolution for the subsequent evolution. We find that for $T = 1000\,\mathrm{eV}$, there are modest but clear differences between the Landau and QLB evolutions, as shown in Figure 12. At 600 eV, the two solutions are very different, as shown in Figure 13. The Landau equation more quickly smoothes out the ripples in the distribution than QLB, and we have two very distinct approaches to equilibrium.

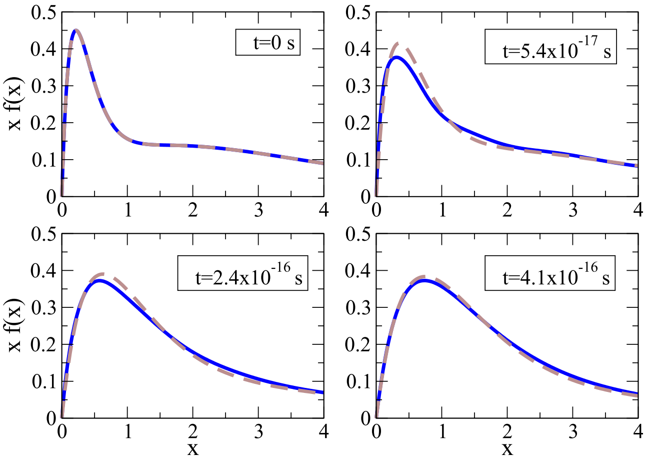


FIG. 9. Comparison between the evolution of the quantum Lenard-Balescu (solid blue) and Landau (dashed brown) equations for a two-temperature initial condition with $\gamma = 0.2$ and $T = 600\,\mathrm{eV}$. The difference between the two is minimal at these conditions. The variable $x = \beta m_e v^2/2$ and $f(x)$ is divided by $n_e(\beta m_e/2\pi)^{3/2}$.

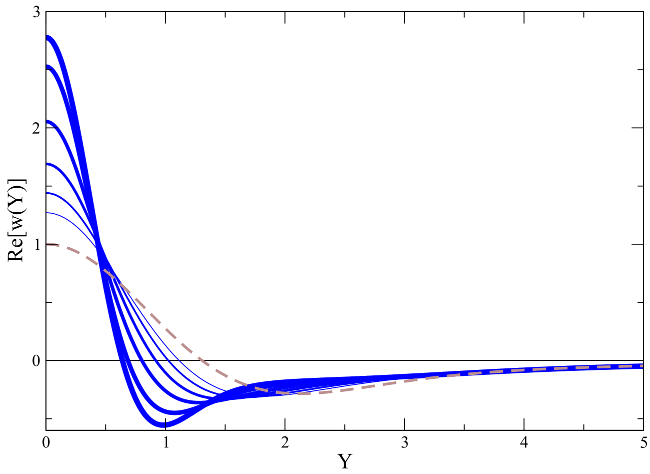


FIG. 10. Evolution of the real part of the free-particle response function under the quantum Lenard-Balescu equation for a two-temperature initial condition with $\gamma = 0.2$ and $T = 1000\,\mathrm{eV}$. The thickest solid line is $t = 0$ and the dashed line is equilibrium. The relationship to the response function in k and ω space is $\chi(k, \omega) = -n_e\beta w(Y)$, where Y is given by (20).

D. Gaussian distribution

We solve for the relaxation of the Gaussian initial condition described in Section VII B with $T = 1000\,\mathrm{eV}$, $v_0 = 2$, and $n_e = 1.0 \times 10^{25}\mathrm{cm}^{-3}$. The evolution of the coefficients is shown in Figure 14, while that of the distribution itself is in Figure 15. As in the two-temperature case, there is not much difference between the Landau and quantum Lenard-Balescu equations at these conditions. And once again, upon making η smaller, our physical and numerical approximations break down before we see any interesting differences.

X. GENERALIZATIONS AND VARIATIONS

A. Beyond order unity

All the calculations we have done here have the logarithmic and order unity terms. To get all the higher order terms is a straightforward generalization. Consider the integral in Equation (34)

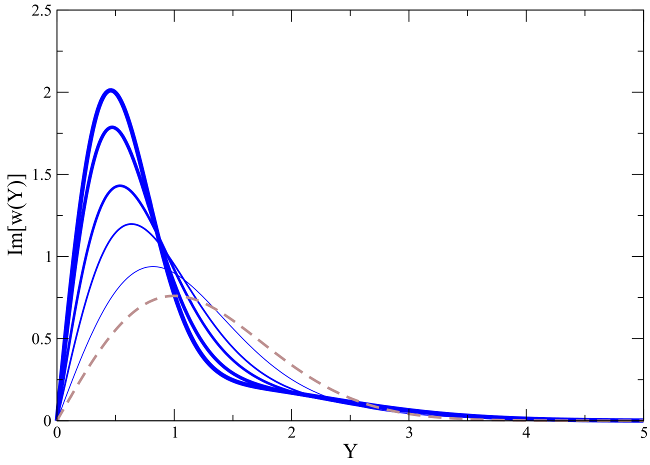


FIG. 11. Evolution of the imaginary part of the free-particle response function under the quantum Lenard-Balescu equation for a two-temperature initial condition with $\gamma = 0.2$ and $T = 1000\,\mathrm{eV}$. The thickest solid line is $t = 0$, and the dashed line is equilibrium. The relationship to the response function in k and ω space is $\chi(k, \omega) = -n_e\beta w(Y)$ where Y is given by (20).

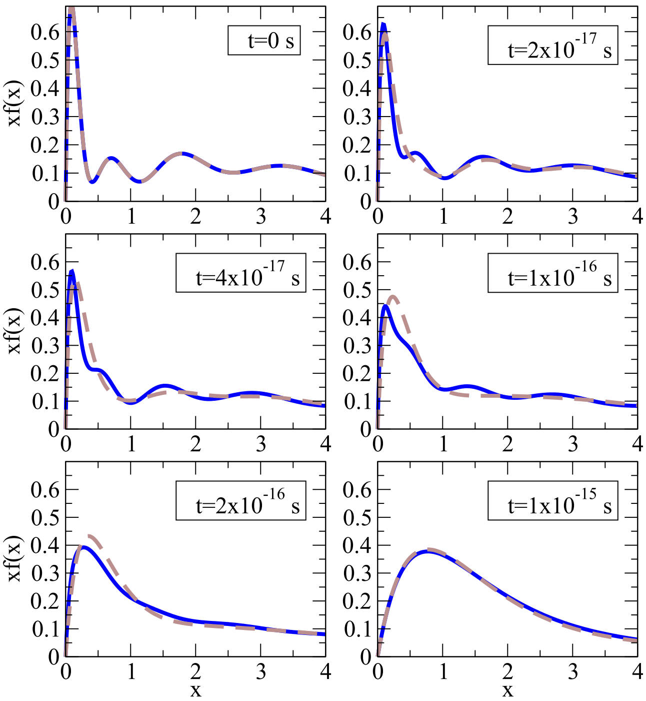


FIG. 12. Comparison between the quantum Lenard-Balescu (solid) and Landau (dashed) equations for a two-temperature initial condition with $\gamma = 0.08$ and $T = 1000\,\text{eV}$. We do not have enough polynomials to fully resolve this distribution, but the series expansion still constitutes a valid initial condition. The variable $x = \beta m_e v^2/2$ and $f(x)$ is divided by $n_e(\beta m_e/2\pi)^{3/2}$.

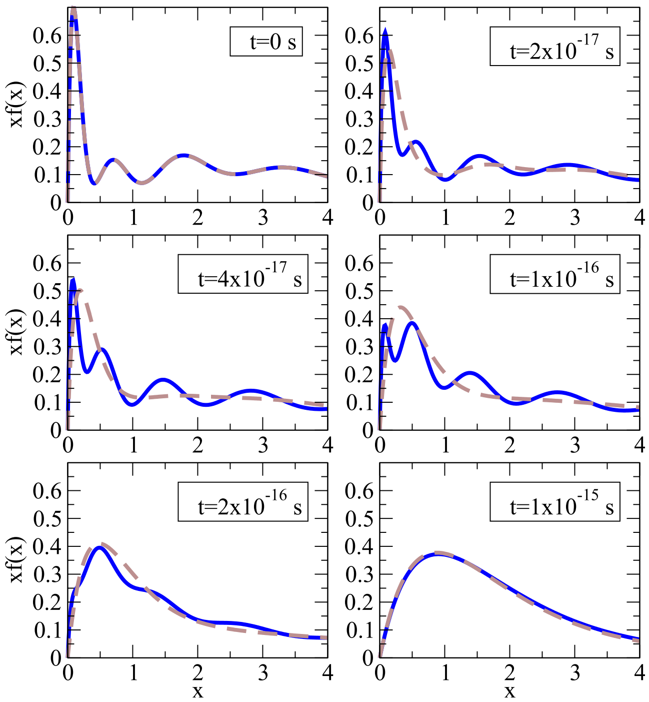


FIG. 13. Comparison between the quantum Lenard-Balescu (solid) and Landau (dashed) equations for a two-temperature initial condition with $\gamma = 0.08$ and $T = 600\,\text{eV}$. We do not have enough polynomials to fully resolve this distribution but the series expansion still constitutes a valid initial condition. The variable $x = \beta m_e v^2/2$ and $f(x)$ is divided by $n_e(\beta m_e/2\pi)^{3/2}$.

$$T^n_{lk} \equiv \int_0^\infty dX \frac{e^{-X^2}}{X^3} \int_{-\infty}^{\infty} dY \frac{e^{-Y^2}}{|\epsilon(X,Y)|^2} P^n_{lk}(X,Y). \tag{82}$$

We can decompose the polynomial into powers of X

$$P^n_{lk}(X,Y) = \sum_{p=1}^{M} X^{2p} G^{np}_{lk}(Y) \tag{83}$$

and keep the dielectric function everywhere, using the formulas in Appendix C (with $p = s+1$) to do the X-integrals for every p rather than just $p = 1$ as we have done. Using the decomposition (83) and the results of Appendix C, the integral (82) can be written

$$T^n_{lk} = -\frac{1}{2}\sum_{p=0}^{M} p!\,\eta^{2p-2} J_p(\eta), \tag{84}$$

where

$$iJ_p(\eta) \equiv \int_{-\infty}^{\infty} \frac{[w(Y)]^p}{w_i(Y)} e^{-Y^2} e^{\eta^2 w(Y)} G^{np}_{lk}(Y)\,\Gamma\big(-p, \eta^2 w(Y)\big)\,dY. \tag{85}$$

Although not clear by inspection, the real part of the integrand in (85) is odd for all p so the integral is always imaginary. Calculating these Y-integrals numerically by the method we used in this work would require that we precompute $G^{np}_{lk}(Y)$ at the quadrature points for all (n, l, k) and all p up to M, which varies depending on the polynomial. This would lead to a large quantity of precomputed data, but it is possible in principle. Alternatively, we can keep the dielectric function for $1 \le p \le p_{\max}$ but set it to 1 for $p > p_{\max}$, which would allow us to keep a prescribed number of powers of η. If the method of Appendix E, or something like it, proves feasible, then we could use it to compute all the coefficients after putting our efforts in computing the coefficients for $l = 0$ alone. This would probably be the ideal solution if it is possible.

To examine where higher powers of η might be needed, we compute $T^n_{lk}(\eta)$ for $(n, l, k) = (4, 4, 4)$ using the formula (84). We compute the Y-integrals with Mathematica's adaptive numerical integration for the case of a two-temperature plasma with $\gamma = 0.5$ and 20 polynomials. This calculation discards no powers of η, and we compare it with the approximation used in the solution of the QLB equation

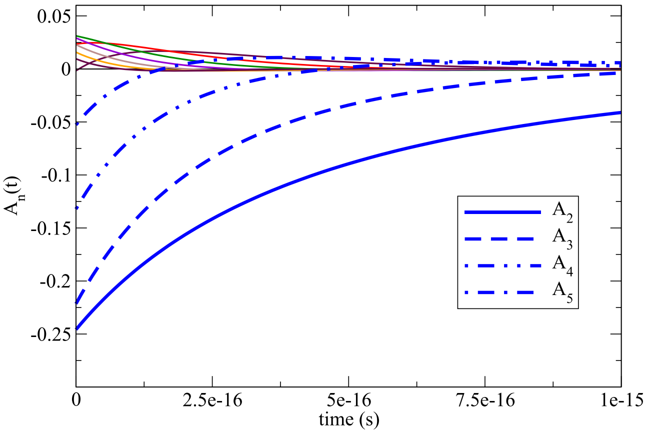


FIG. 14. Evolution of the first few A_n for the Gaussian initial distribution, Equation (76). The first four non-trivial coefficients are labelled, and the rest up to $n = 14$ are shown in various colors.

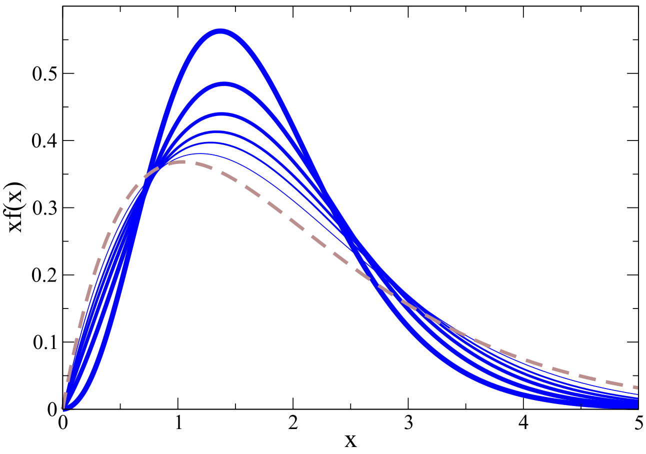


FIG. 15. Evolution of Gaussian initial distribution, Equation (76) with $\bar{v}_0 = 2$ and $T = 1000\,$eV. The thickest line is $t = 0$, and the dashed line is equilibrium. The variable $x = \beta m_e v^2/2$ and $f(x)$ is divided by $n_e(\beta m_e/2\pi)^{3/2}$.

$$T^n_{lk} \approx -\frac{\Gamma(n+3/2)}{n!}\left[S^n_{lk}\left(\frac{\gamma_E}{2} + \ln\eta\right) - B^n_{lk} + F^n_{lk}\right]. \tag{86}$$

The result is shown in the short-dashed red curve in Figure 16. Our approximation is very accurate until $\eta \sim 0.1$, and then higher-order terms become necessary. We stress again that this conclusion is highly dependent upon the distribution, but among the ones we are able to resolve with 40 polynomials, this gives a reasonable idea of where the approximations start to break down. We also compare with the result of setting either B^n_{lk}, F^n_{lk}, or both, to zero. The dashed-dotted purple line indicates that the Landau approximation, where both these terms are set to zero, is worst. Keeping B^n_{lk} but not F^n_{lk} gives the long-dashed green curve, which is a marked improvement, but keeping both is clearly best and is very accurate when $\eta < 0.1$. Table I gives the actual values for $\eta = 0.001$. Now, for this distribution at these conditions, there is not much difference between the solutions of the Landau and QLB equations, so the discrepancies in Figure 16 apparently do not have a noticeable effect.

B. Multiple species

The generalization to multiple species is straightforward, and the techniques for the evaluation of the dielectric function integrals will work in that case too. If we have, say, electrons and protons, we would need two sets of coefficients A^e_n and A^p_n corresponding to expansion (8) for $f_e(v,t)$ and $f_p(v,t)$. Of course, there would also be collision operators for e–e, p–p, and e–p interactions. If we choose $u = 1$, conservation of particles is given by $A^e_0 = A^p_0 = 1$ and energy by $A^e_1 = -A^p_1$.

C. Velocity anisotropy and spatial inhomogeneity

To treat the most general Wigner distributions, $f(\mathbf{r},\mathbf{v},t)$, we can likewise generalize the expansion (8) to be

$$f^{\rm eq}(\mathbf{r},\mathbf{v},t)\sum_{nlm} A_{nlm}(\mathbf{r},t)Y_{lm}(\theta,\phi)L_n^{\left(\frac{1}{2}\right)}\left(\frac{u\beta m_e v^2}{2}\right), \tag{87}$$

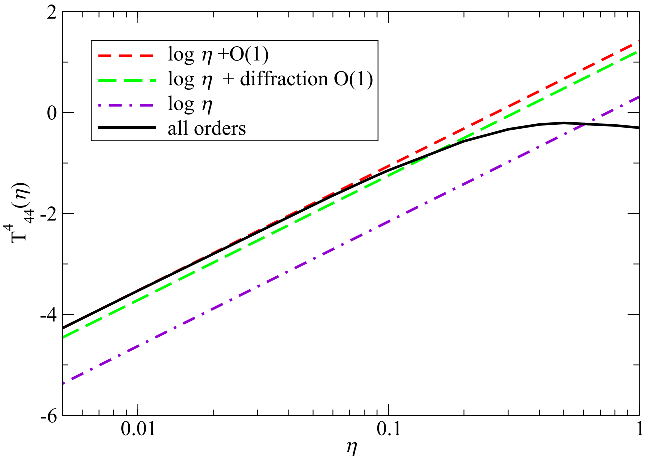


FIG. 16. The integral $T^4_{44}(\eta)$ defined in Equation (82). The solid black curve is the integration carried out to all orders in η; short dashed red is the approximation (54) used in our solution of the quantum Lenard-Balescu equation; long dashed green uses (54) with $F^n_{lk} = 0$; dashed-dotted purple uses (54) with neither B^n_{lk} or F^n_{lk}. Our approximation is very accurate until $\eta \sim 0.1$ and is much better than using the Coulomb logarithm alone.

where $Y_{lm}(\theta,\phi)$ are spherical harmonics. The $A_{nlm}(\mathbf{r},t)$ now satisfy partial differential equations, and the conservation laws are given by integrals of the coefficients over space, which must be respected by the solution method. The results of Appendixes A and B must also be generalized to include the spherical harmonics. This does not appear to be particularly easy, and we may find ourselves lacking the convenient closed formulas we obtained in the isotropic case, but it is surely not impossible.

D. Alternative expansions

There are other related expansions that can be used to solve these equations. First, we explore the possibility of choosing different values of u. As mentioned in Section III, we have been content here to use $u = 1$ because of the trivial conservation properties (12) and (13), but we cannot represent every possible distribution this way. This situation can be rectified by keeping u in Equation (8). The essential form of such an expansion is

$$f(x) = e^{-x/u}\sum_{n=0}^{\infty} A_n L_n^{\left(\frac{1}{2}\right)}(x), \tag{88}$$

that is, the argument of the exponential is a factor of u smaller than that of the Laguerre polynomial. For this expansion to converge, $e^{x/u}f(x)$ must be square-integrable with respect to the Laguerre weight, so

TABLE I. Comparison of the various approximations for calculating $T^4_{44}(\eta)$ defined in Equation (82). Quantum Landau is Equation (86) with $F^n_{lk} = 0$ and Landau is (86) with $B^n_{lk} = F^n_{lk} = 0$.

Approximation	$T^4_{44}(0.001)$
All orders	−5.99681
Equation (86)	−5.99678
Quantum Landau	−6.1849
Landau	−7.09642

$$\int_0^\infty x^{1/2} e^{x(-1+2/u)} |f(x)|^2 dx < \infty. \tag{89}$$

If $u = 2$, this condition is very mild, certainly far less stringent than (16). The cost of this is that we complicate the collision integrals even further, and conservation is no longer automatic. In Appendixes A and B, we indicate the modifications needed for computing the response function and collision integrals with general u. The condition for conservation of particles becomes

$$\frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} A_n \frac{\Gamma(n+3/2)}{n!} (1-u)^n = 1 \tag{90}$$

and for conservation of energy, we have

$$\frac{2}{3\sqrt{\pi}} \sum_{n=0}^{\infty} A_n \frac{\Gamma(n+3/2)}{n!} (3 - 3u - 2nu)(1-u)^n = 1. \tag{91}$$

Thus, if $u = 2$, conservation is no longer automatic but can be lost if we have an insufficient number of coefficients. Of course, one may question whether it is particularly valuable to maintain conservation of particles and energy even if we have insufficient resolution, so this may not be a very important consideration. Also, even though $u = 1$ was sufficient for our purposes because the system does not evolve to a state that violates (16) if it is initially satisfied, when we generalize this method to multiple particle species and situations in which there is an external force, this may no longer be the case. Therefore, in more practical applications, it may be that $u = 2$, or at least $u > 1$, is more appropriate.

An expansion based on a completely different set of orthogonal functions may prove useful. Rather than the orthogonality condition (11), one might be tempted to try polynomials that satisfy

$$\int_0^\infty x^\nu e^{-x^2} P_n(x) P_m(x) dx = M_n \delta_{nm}, \tag{92}$$

where M_n are a set of constants. The distribution expansion is then

$$f(v,t) = n_e \left(\frac{m_e \beta}{2\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta m_e v^2}{2}} \sum_{n=0}^{\infty} A_n(t) P_n\left(\sqrt{m_e\beta/2}v\right). \tag{93}$$

In recent works,[34,35] it was suggested that these polynomials may be a more efficient way to represent distribution functions for certain applications. It does seem to be the case[36] that, compared with the Laguerre polynomials, one needs fewer of them to fully resolve some distributions (in our case, $\nu = 2$ is the natural choice). However, these polynomials are "non-classical,"[37] and there are no closed forms either for the polynomials themselves or for any of their properties such as coefficients of the recurrence relation and the normalization constants, M_n. The polynomials must be generated by the Gram-Schmidt procedure, and all required quantities, such as the real and imaginary parts of the response function and the polynomial $G^n_{lk}(Y)$, would need to be computed numerically at the quadrature points without the aid of any of the exact formulas upon which we have relied. Of course, such closed forms are not really necessary and it may be worth exploring this issue further.

XI. DISCUSSION

We have used a spectral expansion to solve the quantum Lenard-Balescu equation for a one-component Coulomb system. To demonstrate the technique, we have computed the relaxation to equilibrium of various initial distributions including variants of a two-temperature plasma. We have found that including the full dynamical dielectric function makes little difference for these problems and we do just as well if we use the computationally cheaper static screening. This is in general agreement with the findings of Dolinsky for the classical Lenard-Balescu equation; he was not able to find an initial condition for which there was any difference between the Lenard-Balescu and Landau/Fokker-Planck systems. Nevertheless, this conclusion about the relevance of the dielectric function cannot be true in general. As we pointed out, even for a one-component plasma divided into two temperatures, with a large enough temperature ratio, significant differences should be expected between the two kinetic solutions. For the moment, we do not have the resolution to thoroughly study this effect, but we began to see hints of it in our under-resolved $\gamma = 0.08$ solutions. When we generalize to multiple species and anisotropic distributions, the dielectric function may become more important. For example, one can make a significant error in calculations of thermal conductivity for an electron-proton system by using static over dynamical screening in the collision integrals. This difference will also be present in time-dependent solutions.

Compared to the many advances for the Boltzmann and Landau/Fokker-Planck equations,[22–24,26] for which it is now possible to find solutions in multiple dimensions of velocity and space, our velocity-isotropic and 0D spatial solutions may not seem terribly impressive. However, we have shown that it is possible to solve the quantum Lenard-Balescu equation including faithful integrations over the dielectric function. What is more, our method provides an analytic representation of the response and dielectric functions. It is also readily generalizable to multiple space and velocity dimensions and we hope that this will be the subject of future work.

ACKNOWLEDGMENTS

We are grateful to David Michta for providing data for comparison with the Landau solution and to Lorin Benedict, Michael Murillo, Antoine Cerfon, and Cory Hauck for useful discussions. Part of this research was performed while the authors were visiting the Institute for Pure and Applied Mathematics (IPAM), which is supported by the National Science Foundation. Susana Serna was supported by Spanish MINECO Grant No. MTM2014-56218-C2-2-P. This work was performed under the auspices of the U.S. Department of Energy at the Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344.

APPENDIX A: DERIVATION OF THE DIELECTRIC FUNCTION

Here, we compute the response function by computing the integral (5) using our series expansion (8). First, we note that the response function is of the form

$$\chi(k,\omega) = Z(k,\omega_-) - Z(k,\omega_+), \tag{A1}$$

where

$$Z(k,\omega) \equiv \lim_{\eta\to 0^+} \int d^3\mathbf{v}\, \frac{f(\mathbf{v})}{\hbar\omega - \hbar\mathbf{v}\cdot\mathbf{k} + i\eta} \tag{A2}$$

and

$$\omega_\pm \equiv \omega \pm \frac{\hbar k^2}{2m_e}. \tag{A3}$$

Equation (A1) can be easily found by making the substitution $\mathbf{u} = \mathbf{v} + \hbar\mathbf{k}/m_e$ in the second term of (5). Inserting the expansion (8) into (A2), we choose $\mathbf{k}$ to point in the z-direction and integrate in the cylindrical coordinates $(v_\perp, \phi, v_z)$. We then have

$$Z(k,\omega) = \frac{n_e}{\hbar}\left(\frac{\beta m_e}{2\pi}\right)^{\frac{1}{2}} \frac{1}{k}\sum A_k \times \int_{-\infty}^{\infty}\int_0^{\infty} \frac{e^{-x-\beta m_e v_z^2/2} L_k^{\left(\frac{1}{2}\right)}\left(x + \frac{\beta m_e v_z^2}{2}\right)}{\omega/k - v_z + i\eta}\, dx dv_z, \tag{A4}$$

where we have made the substitution $x = m\beta v_\perp^2/2$. We now use the identity

$$\int_0^\infty e^{-x} L_k^{\left(\frac{1}{2}\right)}(x+y)dx = L_k^{\left(-\frac{1}{2}\right)}(y), \tag{A5}$$

which can be derived from the Laguerre sum formula

$$L_n^{(\alpha_1+\alpha_2+1)}(x+y) = \sum_{i=0}^{n} L_i^{(\alpha_1)}(x) L_{n-i}^{(\alpha_2)}(y), \tag{A6}$$

by choosing $\alpha_1 = 0$ and $\alpha_2 = -1/2$. The integral we are left with is

$$Z(k,\omega) = \frac{n_e}{\hbar}\left(\frac{\beta m_e}{2\pi}\right)^{\frac{1}{2}} \frac{1}{k}\sum A_m J_m, \tag{A7}$$

where

$$J_m \equiv \int_{-\infty}^{\infty} \frac{e^{-\beta m_e v_z^2/2} L_m^{\left(-\frac{1}{2}\right)}\left(\frac{\beta m_e v_z^2}{2}\right)}{\omega/k - v_z + i\eta}\, dv_z, \tag{A8}$$

which, with the help of the substitution $x^2 = \beta m_e v_z^2/2$, can be written as

$$J_m = \int_{-\infty}^{\infty} \frac{e^{-x^2} L_m^{\left(-\frac{1}{2}\right)}(x^2)}{Y/\sqrt{2} - x + i\eta}\, dx. \tag{A9}$$

As usual, the imaginary part is easily found from the Sokhotski-Plemelj theorem

$$\lim_{\epsilon\to 0^+} \int_{-\infty}^{\infty} \frac{f(x)}{x \pm i\epsilon} = \mp i\pi f(0) + P\int_{-\infty}^{\infty} \frac{f(x)}{x}\, dx, \tag{A10}$$

where P denotes principal value integration. Thus

$$\operatorname{Im} J_m = -\pi e^{-Y^2/2} L_m^{\left(-\frac{1}{2}\right)}\left(\frac{Y^2}{2}\right). \tag{A11}$$

To find the real part, we will not directly attempt the principal value integral but will instead use the standard trick

$$\frac{1}{\frac{Y}{\sqrt{2}} - x + i\eta} = -i\int_0^\infty e^{i\left(Y/\sqrt{2} - x + i\eta\right)t} dt \tag{A12}$$

and the identity

$$L_m^{\left(-\frac{1}{2}\right)}(x^2) = \frac{(-1)^m}{m! 2^{2m}} H_{2m}(x), \tag{A13}$$

where $H_n(x)$ are Hermite polynomials, to write

$$J_m = -i\frac{(-1)^m}{m! 2^{2m}} \times \int_0^\infty \int_{-\infty}^{\infty} e^{-x^2} e^{i\left(\frac{Y}{\sqrt{2}} - x + i\eta\right)t} H_{2m}(x) dx dt. \tag{A14}$$

Because $H_{2m}(x)$ and e^{-x^2} are even, we have

$$J_m = -i\frac{(-1)^m}{m! 2^{2m}} \times 2\int_0^\infty \int_0^\infty e^{-x^2} e^{i\left(\frac{Y}{\sqrt{2}}\right)t} \cos(xt) H_{2m}(x) dx dt. \tag{A15}$$

Using 7.388.3 of Gradshteyn and Ryzhik,[31] we find for the x-integral

$$\int_0^\infty e^{-x^2} \cos(xt) H_{2m}(x) dx = (-1)^m \frac{\sqrt{\pi}}{2} t^{2m} \exp\left(-\frac{t^2}{4}\right), \tag{A16}$$

so

$$\mathrm{Re} J_m = \frac{\sqrt{\pi}}{m! 2^{2m}} \int_0^\infty \sin\left(\frac{Y}{\sqrt{2}} t\right) t^{2m} e^{-t^2/4} dt. \tag{A17}$$

Again consulting Gradshteyn and Ryzhik, this time 3.952.7, and using Kummer's transformation for the confluent hypergeometric function, $M(a,b;z)$, we arrive at the real part of $Z(k,\omega)$

$$\mathrm{Re} Z = n_e \frac{\beta m_e}{\hbar} \frac{\omega}{k^2} \sum_m A_m M\left(1+m, \frac{3}{2}; -\frac{Y^2}{2}\right). \tag{A18}$$

Putting together our previous results, the imaginary part is

$$\mathrm{Im} Z = -\sqrt{\frac{\pi}{2}} n_e \frac{\sqrt{\beta m_e}}{\hbar k} \sum_m A_m e^{-\frac{Y^2}{2}} L_m^{\left(-\frac{1}{2}\right)}\left(\frac{Y^2}{2}\right). \tag{A19}$$

Using (A1) and the definitions $Y_\pm^2 \equiv \beta m_e \omega_\pm^2/2k^2$ and (19), we arrive at (4). The calculation for arbitrary u is the same except that one expands $H_{2m}(\sqrt{u}x)$ by means of a multiplication theorem for Hermite polynomials.

APPENDIX B: REDUCTION OF THE COLLISION INTEGRALS

The purpose of this Appendix is to derive a simplification of the quantum Lenard-Balescu collision integral using the polynomial expansion in Equation (8). The calculation is only for $u = 1$, which we use exclusively in this paper. The generalization to arbitrary u is straightforward but results in more complicated formulas.

We multiply the left- and right-hand sides of the kinetic equation (1) by $L_m^{(\frac{1}{2})}(\beta m_e v^2/2)$ and integrate over $d^3\mathbf{v}$. The left-hand side becomes

$$\frac{2n_e}{\sqrt{\pi}}\frac{\Gamma(n+3/2)}{n!}\frac{dA_n}{dt}. \tag{B1}$$

The right-hand side is, of course, the real problem; it is the nine-fold integral

$$\int C_{QLB}(f) L_n^{(\frac{1}{2})}\left(\frac{m_e\beta v^2}{2}\right) d^3\mathbf{v}, \tag{B2}$$

which we will reduce to two. Beginning with $\mathbf{v}'$, we make use of the convenient definitions

$$\omega \equiv \mathbf{k}\cdot\mathbf{v} + \frac{\hbar k^2}{2m_e}, \tag{B3}$$

$$\omega_\pm \equiv \omega \pm \frac{\hbar k^2}{2m_e}, \tag{B4}$$

to write

$$\int d^3\mathbf{v}'\delta(\omega_+ - \mathbf{k}\cdot\mathbf{v}')f(\mathbf{v}') = n_e\left(\frac{\beta m_e}{2\pi}\right)^{\frac{1}{2}}\frac{1}{k}\sum_{k=0}^{\infty}A_k(t)L_k^{(-\frac{1}{2})}\left(Y_+^2\right)e^{-Y_+^2} \tag{B5}$$

and

$$\int d^3\mathbf{v}'\delta(\omega_+ - \mathbf{k}\cdot\mathbf{v}')f(\mathbf{v}' - \hbar\mathbf{k}/m_e) = n_e\left(\frac{\beta m_e}{2\pi}\right)^{\frac{1}{2}}\frac{1}{k}\sum_{k=0}^{\infty}A_k(t)L_k^{(-\frac{1}{2})}(Y_-^2)e^{-Y_-^2}, \tag{B6}$$

where

$$Y_\pm^2 \equiv \frac{\beta m_e \omega_\pm^2}{2k^2}. \tag{B7}$$

To derive (B5), we take $\mathbf{k}$ as the z-direction and integrate in the cylindrical coordinates $(v_\perp, \phi, v_z')$. We find then

$$\begin{aligned}\int_{-\infty}^{\infty}\int_0^{2\pi}\int_0^{\infty}&\delta(\omega_+ - kv_z')e^{-\frac{\beta m_e\left(v\prime_\perp{}^2 + v_z\prime^2\right)}{2}}\\ &\times L_n^{(\frac{1}{2})}\left(\frac{\beta m_e v_\perp^2}{2} + \frac{\beta m_e v_z\prime^2}{2}\right)v_\perp dv_\perp d\phi dv_z'\\ &= \frac{2\pi}{\beta m_e}\frac{1}{k}e^{-Y_+^2}\int_0^{\infty}e^{-x}L_n^{(\frac{1}{2})}\left(x + Y_+^2\right)dx,\end{aligned} \tag{B8}$$

where we have made the substitution $x = \beta m_e v_\perp^2/2$. Combining (A5) and (B8) with the prefactors and series in $f(\mathbf{v}')$, we find (B5). Equation (B6) is found in exactly the same way after making the substitution $\mathbf{u} = \mathbf{v} - \hbar\mathbf{k}/m_e$.

We aim in the end to have an integration over k and ω. For the $\mathbf{v}$ integration, we again take $\mathbf{k}$ to point in the z-direction, and then we have $v_z = \omega_-/k$ and $dv_z = d\omega/k$. We will therefore employ cylindrical coordinates $(v_\perp, \phi, v_z)$ and integrate over $v_\perp$ and ϕ, leaving v_z as the ω-integral. Because (B5) and (B6) depend only on ω and k, they will play no further role in the integration. The pieces we do need are

$$\int d^3\mathbf{v} f(\mathbf{v})L_n^{(\frac{1}{2})}\left(\frac{m_e\beta v^2}{2}\right) \tag{B9}$$

and

$$\int d^3\mathbf{v} f(\mathbf{v} + \hbar\mathbf{k}/m_e)L_n^{(\frac{1}{2})}\left(\frac{m_e\beta v^2}{2}\right), \tag{B10}$$

integrated over $v_\perp$ and ϕ, for which we find

$$\begin{aligned}\int_0^{2\pi}\int_0^{\infty}&f(\mathbf{v})L_n^{(\frac{1}{2})}\left(\frac{m_e\beta v^2}{2}\right)v_\perp dv_\perp d\phi\\ &= n_e\left(\frac{m_e\beta}{2\pi}\right)^{3/2}2\pi\sum A_l(t)\int_0^{\infty}dv_\perp v_\perp e^{-\frac{\beta m_e\left(v_\perp^2 + v_z^2\right)}{2}}\\ &\quad\times L_n^{(\frac{1}{2})}\left(\frac{\beta m_e v_\perp^2}{2} + Y_-^2\right)L_l^{(\frac{1}{2})}\left(\frac{\beta m_e v_\perp^2}{2} + Y_-^2\right),\\ &= n_e\left(\frac{m_e\beta}{2\pi}\right)^{1/2}e^{-Y_-^2}\sum A_l(t)\\ &\quad\times\int_0^{\infty}e^{-x}L_n^{(\frac{1}{2})}(x + Y_-^2)L_l^{(\frac{1}{2})}(x + Y_-^2)dx,\end{aligned} \tag{B11}$$

and, similarly

$$\begin{aligned}\int_0^{2\pi}\int_0^{\infty}&f(\mathbf{v} + \hbar\mathbf{k}/m_e)L_n^{(\frac{1}{2})}\left(\frac{m_e\beta v^2}{2}\right)v_\perp dv_\perp d\phi\\ &= n_e\left(\frac{m_e\beta}{2\pi}\right)^{1/2}e^{-Y_+^2}\sum A_l(t)\\ &\quad\times\int_0^{\infty}e^{-x}L_n^{(\frac{1}{2})}(x + Y_-^2)L_l^{(\frac{1}{2})}\left(x + Y_+^2\right)dx.\end{aligned} \tag{B12}$$

To handle the last integrals in (B11) and (B12), we use the identity

$$\int_0^{\infty}e^{-x}L_n^{(\frac{1}{2})}(x + y)L_l^{(\frac{1}{2})}(x + z)dx = \sum_{i=0}^{\min(l,n)}L_{n-i}^{(-\frac{1}{2})}(y)L_{l-i}^{(-\frac{1}{2})}(z), \tag{B13}$$

which can be easily derived by again using (A6) with $\alpha_1 = 0$ and $\alpha_2 = -1/2$. At this point, only the magnitude of k is left in the integrand, so $\int d^3\mathbf{k} \to 4\pi \int k^2 dk$. To get the final form of the integrand, we multiply (B5) by (B11), subtract the product of (B6) and (B12), integrate over $4\pi \int_{-\infty}^{\infty} dv_z \int_0^{\infty} k^2 dk$, and include the dimensional prefactor in (2). The resulting expression for the coefficients is

$$C^n_{lk} = -\frac{\beta m_e n_e}{4\pi^{3/2}\hbar^2} \frac{n!}{\Gamma(n+3/2)} \int_{-\infty}^{\infty} \int_0^{\infty} e^{-\left(Y_-^2 + Y_+^2\right)} \times \frac{|\phi(k)|^2}{|\epsilon(k,\omega)|^2} q^n_{lk}(X,Y) dk d\omega, \tag{B14}$$

where $q^n_{lk}(X,Y)$ is given in Equation (37). It is a straightforward matter to use the dimensionless variables X and Y defined in (19) and (20) along with the Coulomb potential to arrive at (34).

APPENDIX C: EXACT INTEGRATION OVER THE DIELECTRIC FUNCTION

Here, we will simplify Equation (46) by an exact integration over X. In fact, we will solve the more general case

$$I_X(Y,s) \equiv \int_0^{\infty} dX X^{2s-1} \frac{e^{-X^2}}{|\epsilon(X,Y)|^2} \tag{C1}$$

for which the X-integral in (46) is the special case $s = 0$. We do this because including terms in the collision integrals greater than O(1) requires integrals for which $s > 0$, a generalization we may wish to consider in the future. These are also no more difficult than the $s = 0$ case.

To begin, we make use of the identity

$$\frac{1}{|\epsilon(X,Y)|^2} = \frac{1}{2i\mathrm{Im}(\epsilon)} \left(\frac{1}{\epsilon *} - \frac{1}{\epsilon}\right). \tag{C2}$$

Using the classical dielectric function in Equation (30), the integral becomes

$$I_X(Y,s) = \frac{1}{2i\eta^2 w_i(Y)} \times \int_0^{\infty} \left[\frac{X^{2s+3}e^{-X^2}}{X^2 + \eta^2 w^*(Y)} - \frac{X^{2s+3}e^{-X^2}}{X^2 + \eta^2 w(Y)}\right] dX. \tag{C3}$$

Glancing at (31) and (32), it is clear that $w_i(Y) = -w_i(-Y)$ and $w_r(Y) = w_r(-Y)$ so that $w^*(Y) = w(-Y)$. The first term of (C3) can be written as

$$\frac{1}{2i\eta^2 w_i(Y)} \int_0^{\infty} \frac{X^{2s+3}e^{-X^2}}{X^2 + \eta^2 w(-Y)} dX. \tag{C4}$$

If in the Y-integration we make the substitution $Y \to -Y$, we find that the first and second terms of (C3) are actually equal and opposite when integrating over Y and we can set

$$I_X(Y,s) = -\frac{1}{i\eta^2 w_i(Y)} \int_0^{\infty} \frac{X^{2s+3}e^{-X^2}}{X^2 + \eta^2 w(Y)} dX. \tag{C5}$$

This integral can be evaluated in terms of special functions

$$\int_0^{\infty} \frac{X^{2s+3}e^{-X^2}}{X^2 + \eta^2 w(Y)} dX = \frac{1}{2} e^{\eta^2 w(Y)} \times \left[\eta^2 w(Y)\right]^{1+s} (1+s)!\Gamma\left(-1-s, \eta^2 w(Y)\right) \tag{C6}$$

where $\Gamma(t,z)$ is the incomplete gamma function, defined by

$$\Gamma(t,z) = \int_z^{\infty} u^{t-1} e^{-u} du. \tag{C7}$$

This function has the series

$$\Gamma(-1-s,z) = \frac{(-1)^{1+s}}{(1+s)!} \left[\psi_{s+2} - \gamma_E - \log z\right] - \frac{1}{z^{s+1}} \sum_{k=0,k\neq s+1}^{\infty} \frac{(-z)^k}{(k-s-1)k!}. \tag{C8}$$

where $\gamma_E \approx 0.57721566$ is Euler's constant and ψ_{s+2} are constants appearing in the digamma function at integer arguments. The first few of these are

$$\psi_2 = 1,$$
$$\psi_3 = 3/2,$$
$$\psi_4 = 11/6,$$
$$\psi_5 = 25/12\,.$$

The order-unity terms arising from dynamical screening come only from $s = 0$ and we can now isolate these using (C8) in (C6) and expanding $\exp[\eta^2 w(Y)]$. Doing this, we find (46).

APPENDIX D: EXACT EXPRESSIONS FOR LANDAU COEFFICIENTS

Here, we derive Equation (57), the closed expression for S^n_{lk}, and the coefficients of the Landau equation. We begin by setting $\epsilon(k,\omega) = 1$ and using the variables $g \equiv k\sqrt{m_e\beta/2}$ and $z \equiv v_z\sqrt{m_e\beta/2}$, to write the coefficients for the quantum Landau equation C^n_{lk}

$$C^n_{lk} = -\bar{C} \sum_{j=0}^{\min(\ell,n)} \int_{g_0}^{\infty} dg \int_{-\infty}^{\infty} dz I_{QL}(g,z;\hbar), \tag{D1}$$

where the integrand is

$$I_{QL}(g,z;\hbar) \equiv \frac{1}{\hbar^2}\frac{1}{g^3} \exp\left(-z^2 - (z + \hbar g/m_e)^2\right) \times L_{n-j}^{\left(-\frac{1}{2}\right)}(z^2) \left[L_k^{\left(-\frac{1}{2}\right)}\left((z+\hbar g/m_e)^2\right) L_{l-j}^{\left(-\frac{1}{2}\right)}(z^2) - L_{\ell-j}^{\left(-\frac{1}{2}\right)}\left((z+\hbar g/m_e)^2\right) L_k^{\left(-\frac{1}{2}\right)}(z^2)\right], \tag{D2}$$

the prefactor is now

$$\bar{C} = \frac{\sqrt{\pi} n_e (2m_e\beta)^{3/2} e^4\, n!}{\Gamma(n+3/2)} \tag{D3}$$

and g_0 is the small-k cutoff. To obtain the classical version of this expression, we take the $\hbar \to 0$ limit

$$I_L(g,z) = \lim_{\hbar \to 0} I_{QL}(g,z;\hbar). \tag{D4}$$

To do this, we must expand to second-order in $\hbar$, and using some simple identities of Laguerre polynomials, this yields

$$I_L(g,z) = \frac{2e^{-2z^2} L_{n-j}^{\left(-\frac{1}{2}\right)}(z^2)}{m_e^2 g} \times \Bigg[(\ell - j + 1) L_k^{\left(-\frac{1}{2}\right)}(z^2) L_{\ell-j+1}^{\left(-\frac{1}{2}\right)}(z^2) - (k+1) L_{\ell-j}^{\left(-\frac{1}{2}\right)}(z^2) L_k^{\left(-\frac{1}{2}\right)}(z^2)\Bigg]. \tag{D5}$$

The z-integration of this expression can be done by taking advantage of the relationship between Laguerre and Hermite polynomials in Equation (A13), and Titchmarsh's identity

$$\int_{-\infty}^{\infty} dz\, e^{-2z^2} H_a(z) H_b(z) H_c(z) = \frac{2^{(a+b+c-1)/2}}{\pi} \times \Gamma\left(\frac{a+b-c+1}{2}\right)\Gamma\left(\frac{a-b+c+1}{2}\right) \times \Gamma\left(\frac{-a+b+c+1}{2}\right), \tag{D6}$$

when $a+b+c$ is even and the integral is zero otherwise. The coefficients then become

$$C_{lk}^n = -\bar{C}\frac{\sqrt{2}\log\Lambda}{m_e^2 \pi} \sum_{j=0}^{\min(\ell,n)} (-1)^{n+k+\ell} \times \frac{2^{-(n+k+\ell-2j+1)}}{(n-j)!k!(\ell-j)!} \Gamma\left(-n+k+\ell+\frac{3}{2}\right) \times \left[\Gamma\left(n+k-\ell-\frac{1}{2}\right)\Gamma\left(n-k+\ell-2j+\frac{3}{2}\right) - \Gamma\left(n+k-\ell+\frac{3}{2}\right)\Gamma\left(n-k+\ell-2j-\frac{1}{2}\right)\right]. \tag{D7}$$

At this point, we confess that we simply evaluated the above sum over j in Mathematica, which returns an analytical form involving the regularized hypergeometric function, ${}_3\tilde{F}_2$. Comparing this expression with Equation (56) to get the correct numerical constants, we find Equation (57) for the S_{lk}^n. We do not yet know how to derive this formula legitimately, but we performed many checks between the analytic expression and the numerically determined S_{lk}^n to ensure that (57) is indeed correct.

APPENDIX E: RECURRENCE FORMULAS FOR COEFFICIENTS

In this Appendix, we derive and solve the recurrence formula for the triple product integrals V_{lk}^j defined in Equation (52). This will be based on the formula for Laguerre polynomials

$$L_n^{\left(-\frac{1}{2}\right)}(x) = \left(2 - \frac{3/2 + x}{n}\right) L_{n-1}^{\left(-\frac{1}{2}\right)}(x) - \left(1 - \frac{3}{2n}\right) L_{n-2}^{\left(-\frac{1}{2}\right)}(x). \tag{E1}$$

Note that in the definition of V, two of the three Laguerre polynomials have the same argument. The strategy is to exploit this fact to find the recurrence formula for the family of integrands

$$v_{lk}^j(X,Y) \equiv g(X,Y) L_j^{\left(-\frac{1}{2}\right)}(Y_-^2) L_k^{\left(-\frac{1}{2}\right)}\left(Y_+^2\right) L_l^{\left(-\frac{1}{2}\right)}(Y_-^2), \tag{E2}$$

where $g(X,Y)$ is an arbitrary function of X and Y. We begin by using the Laguerre formula to show

$$L_j^{\left(-\frac{1}{2}\right)}(Y_-^2) = \left(2 - \frac{3}{2j}\right) L_{j-1}^{\left(-\frac{1}{2}\right)}(Y_-^2) - \left(1 - \frac{3}{2j}\right) L_{j-2}^{\left(-\frac{1}{2}\right)}(Y_-^2) - \frac{1}{j} Y_-^2 L_{j-1}^{\left(-\frac{1}{2}\right)}(Y_-^2). \tag{E3}$$

Therefore,

$$v_{lk}^j(X,Y) = \left(2 - \frac{3}{2j}\right) v_{lk}^{j-1}(X,Y) - \left(1 - \frac{3}{2j}\right) v_{lk}^{j-2}(X,Y) - \frac{1}{j} Y_-^2 v_{lk}^{j-1}(X,Y). \tag{E4}$$

We have a factor of Y_-^2 that will cause problems when we integrate. We can rid ourselves of it by using the recurrence formula again in the rearranged form

$$Y_-^2 L_l^{\left(-\frac{1}{2}\right)}(Y_-^2) = \left(2l + \frac{1}{2}\right) L_l^{\left(-\frac{1}{2}\right)}(Y_-^2) - \left(l - \frac{1}{2}\right) L_{l-1}^{\left(-\frac{1}{2}\right)}(Y_-^2) - (l+1) L_{l+1}^{\left(-\frac{1}{2}\right)}(Y_-^2) \tag{E5}$$

to get

$$Y_-^2 v_{lk}^{j-1}(X,Y) = \left(2l + \frac{1}{2}\right) v_{lk}^{j-1}(X,Y) - \left(l - \frac{1}{2}\right) v_{l-1,k}^{j-1}(X,Y) - (l+1) v_{l+1,k}^{j-1}(X,Y), \tag{E6}$$

which we can plug into (E4) to get a recurrence formula for the integrand free of any additional factors of X or Y. Putting everything together, we find

$$V_{lk}^j = \left[2 - \frac{2(l+1)}{j}\right] V_{lk}^{j-1} - \left(1 - \frac{3}{2j}\right) V_{lk}^{j-2} + \frac{l - \frac{1}{2}}{j} V_{l-1,k}^{j-1} + \frac{l+1}{j} V_{l+1,k}^{j-1}. \tag{E7}$$

Note that k does not participate in this recurrence formula. This system is comprised of $n_{\max}+1$ discrete boundary-value problems in j and l corresponding to each possible value of k. For example, we can use our numerical integration techniques to find V_{lk}^j for $j = 0$ and $j = n_{\max}$ for all values of k. In principle, the recurrence formula then generates the rest. Note that although this is a three-point recurrence, it can be used to compute $j = 1$ as long as we adhere to the convention that when $j > 0$, we set negative-index polynomials to zero when they arise. This convention allows computation of the Laguerre polynomials themselves for $j > 0$ and therefore also works for our coefficients. Another possibility is that rather than specify the values on the boundary of the cube, we instead compute the coefficients only for $j = 0$. We now need to compute these up to $2n_{\max}$. However, in practice, it appears that this procedure is unstable and the initial errors become out of control after a few iterations. We will not investigate this issue further here, but we point out that (E7) can be solved exactly. To do this, rewrite (E7) in the more symmetric form

$$\begin{aligned} jV^{j}_{l-1,k} - 2jV^{j-1}_{l-1,k} + \left(j-\frac{3}{2}\right)V^{j-2}_{l-1,k} \\ = lV^{j-1}_{lk} - 2jV^{j-1}_{l-1,k} + \left(j-\frac{3}{2}\right)V^{j-1}_{l-2,k}, \end{aligned} \tag{E8}$$

so that $l-1$ is on the left and $j-1$ on the right. Looking for a separable solution of the form

$$V^{j}_{lk} = U_j W_l \tag{E9}$$

we find the equations

$$\begin{aligned} jU_j - 2jU_{j-1} + \left(j-\frac{3}{2}\right)U_{j-2} &= \lambda U_{j-1}, \\ lW_l - 2lW_{l-1} + \left(l-\frac{3}{2}\right)W_{l-2} &= \lambda W_{l-1}, \end{aligned} \tag{E10}$$

where λ is an arbitrary constant. We can immediately recognize these as the recurrence formulas for the associated Laguerre polynomials with $\alpha = -1/2$ and $x = -3/2 - \lambda$. The most general solution to (E7) is a superposition of various values of λ each with a different amplitude, μ

$$V^{j}_{lk} = \sum_i \mu_i^{(k)} L_j^{\left(-\frac{1}{2}\right)}\left(-\frac{3}{2} - \lambda_i^{(k)}\right) L_l^{\left(-\frac{1}{2}\right)}\left(-\frac{3}{2} - \lambda_i^{(k)}\right). \tag{E11}$$

We require enough constants μ and λ that can satisfy the boundary conditions. It is not clear how best to apportion and calculate these but we have several options. This requires further investigation, but it seems clear that if such an approach can work, it would greatly improve the algorithm presented in the paper.

[1] M. N. Rosenbluth, W. M. MacDonald, and D. L. Judd, Phys. Rev. **107**, 1 (1957).

[2] R. S. Cohen, L. Spitzer, and P. McRoutly, Phys. Rev. **80**, 230 (1950).

[3] L. Spitzer and R. Härm, Phys. Rev. **89**, 977 (1953).

[4] D. O. Gericke, S. Kosse, M. Schlanges, and M. Bonitz, Phys. Rev. B **59**, 10639 (1999).

[5] R. H. Williams and H. E. DeWitt, Phys. Fluids **12**, 2326 (1969).

[6] H. D. Whitley, C. R. Scullard, L. X. Benedict, J. I. Castor, A. Randles, J. N. Glosli, D. F. Richards, M. P. Desjarlais, and F. R. Graziani, Contrib. Plasma Phys. **55**, 192 (2015).

[7] F. Morales, M. K. Kilimann, R. Redmer, M. Schlanges, and F. Bialas, Contrib. Plasma Phys. **29**, 425 (1989).

[8] S. Ichimaru and S. Tanaka, Phys. Rev. A **32**, 1790 (1985).

[9] J. Daligault and G. Dimonte, Phys. Rev. E **79**, 056403 (2009).

[10] L. X. Benedict, M. P. Surh, J. I. Castor, S. A. Khairallah, H. D. Whitley, D. F. Richards, J. N. Glosli, M. S. Murillo, C. R. Scullard, P. E. Grabowski, D. Michta, and F. R. Graziani, Phys. Rev. E **86**, 046406 (2012).

[11] J. S. Chang and G. Cooper, J Comput. Phys. **6**, 1 (1970).

[12] E. M. Epperlein, J Comput. Phys. **112**, 291 (1994).

[13] H. W. Wyld and D. Pines, Phys. Rev. **127**, 1851 (1962).

[14] A. Lenard, Ann. Phys. **10**, 390 (1960).

[15] R. Balescu, Phys. Fluids **3**, 52 (1960).

[16] J. Vorberger and D. O. Gericke, Phys. Plasmas **16**, 082702 (2009).

[17] D. A. Chapman, J. Vorberger, and D. O. Gericke, Phys. Rev. E **88**, 013102 (2013).

[18] R. Binder, D. Scott, A. E. Paul, M. Lindberg, K. Henneberger, and S. W. Koch, Phys. Rev. B **45**, 1107 (1992).

[19] M. Bonitz, *Quantum Kinetic Theory* (Springer, 2016).

[20] A. Dolinsky, Phys. Fluids **8**, 436 (1965).

[21] P. Ricci and G. Lapenta, Phys. Plasmas **9**, 430 (2002).

[22] M. Tzoufras, A. Bell, P. Norreys, and F. Tsung, J. Comput. Phys. **230**, 6475 (2011).

[23] W. T. Taitano and L. Chacón, J. Comput. Phys. **284**, 718 (2015).

[24] W. Taitano, L. Chacón, A. Simakov, and K. Molvig, J. Comput. Phys. **297**, 357 (2015).

[25] I. M. Gamba and J. R. Haack, J. Comput. Phys. **270**, 40 (2014).

[26] J. R. Haack and I. M. Gamba, AIP Conf. Proc. **1501**, 334 (2012).

[27] A. V. Bobylev and I. F. Potapenko, J. Comput. Phys. **246**, 123 (2013).

[28] B. Haible, http://www.ginac.de/CLN/ for CLN—Class library for numbers.

[29] J. Hu, S. Jin, and B. Yan, Commun. Comput. Phys. **12**, 1541 (2012).

[30] J. Daligault, Phys. Plasmas **23**, 032706 (2016).

[31] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic Press Inc., 1965).

[32] B. Schling, *The Boost C++ Libraries* (XML Press, 2011).

[33] D. Michta, F. Graziani, T. Luu, and J. Pruet, Phys. Plasmas **17**, 012707 (2010).

[34] M. Landreman and D. R. Ernst, J. Comput. Phys. **243**, 130 (2013).

[35] J. Wilkening, A. J. Cerfon, and M. Landreman, J. Comput. Phys. **294**, 58 (2015).

[36] P. Baddoo, B. Denny, J. Oehrlein, and M. J. Munoz-Lopez, "Exact formulas for a set of orthogonal polynomials" (unpublished).

[37] L. Miranian, J. Phys. A: Math. Gen. **38**, 6379 (2005).

Imperial College London
Department of Physics

ACCRETION DISCS AND PLANET FORMATION AROUND YOUNG STARS

Marija Jankovic

Submitted in part fulfilment of the requirements
for the degree of Doctor of Philosophy at
Imperial College London, June 2020

Declaration of Originality

This thesis is my own work, except where specifically indicated in the text. Some parts of this thesis are based on work completed in collaboration with others, and some have also been published as journal articles:

- Chapter 2 is based on work that has been completed in collaboration with S. Mohanty, J. Tan and J. E. Owen, and is published as
  Subhanjoy Mohanty, Marija R. Jankovic, Jonathan C. Tan, and James E. Owen, Inside-out Planet Formation. V. Structure of the Inner Disk as Implied by the MRI, Astrophysical Journal, 861, 144 (2018)
- Chapters 3 and 4 are based on work that has been completed in collaboration with J. E. Owen and S. Mohanty, and is published as
  Marija R. Jankovic, James E. Owen, and Subhanjoy Mohanty, Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc, Monthly Notices of the Royal Astronomical Society, 484, 2296–2308 (2019)
- Chapters 5 and 6 are based on work that has been completed in collaboration with S. Mohanty, J. E. Owen and J. Tan.

Marija Jankovic (2020)

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND).

Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Among the extrasolar planets discovered so far, the most abundant are the close-in super-Earths. These are planets with sizes between that of the Earth and Neptune, and orbits typically smaller than Mercury's. In this thesis, I study the innermost regions of accretion discs surrounding young stars, and if and how close-in super-Earths can form at such short orbital periods.

I start by discussing a simple model of the inner disc structure coupled to a detailed prescription of disc accretion due to the magneto-rotational instability (MRI). I use the inferred structure of the gas to show that the MRI leads to accumulation of dust in the inner disc, as necessary for the formation of solid planet cores. Next, assuming that solid cores do form in the inner disc, I investigate the accretion and evolution of planetary atmospheres. I show that, despite the MRI-accreting inner disc being gas-poor, the predicted planet atmospheres are at least as large as observed. Finally, I present an improved model of the inner disc that accounts for disc heating due to accretion and stellar irradiation, vertical energy transport, dust opacities, and dust effects on disc ionization. The optically-thick inner disc is weakly affected by stellar irradiation, and also convectively unstable. Dust controls the ionization state of the inner disc, and thus the onset of the MRI. I show that sustained dust accumulation can occur in the inner disc, without suppressing the MRI. If planets form in the inner disc, larger gas accretion rates (and thus earlier times in the disc lifetime) are favoured.

The work in this thesis advances our knowledge of the planet-forming environment at short orbital distances and supports the hypothesis that super-Earths could form near their present orbits. This work also identifies impediments to planet formation in the inner disc which require further study.

Acknowledgements

I gratefully acknowledge generous support from the President's PhD scholarship of the Imperial College London, the Dositeja stipend from the Fund for Young Talents of the Serbian Ministry for Youth and Sport, ERC-STG-2019 grant (PEVAP) and the Dr Francis John Warner Prize.

I would like to express my deepest gratitude to my supervisor, Subhanjoy Mohanty, for sharing with me his deep knowledge and passion for astrophysics, his guidance, and the meticulous comments that helped improve this thesis. I am most appreciative of his patience and exceptional support. Furthermore, I am extremely grateful to James Owen, for countless insightful suggestions, stimulating discussions, and constructive comments. I am most thankful for his kindness and warm encouragement. I must also thank Jonathan Tan whose ideas and expertise were instrumental in developing this work.

I would like to thank Thomas Haworth for his persistent help, valuable advice and moral support. I also thank Richard Booth, Steven Desch, Eve Lee, Colin McNally, Neal Turner, Lauren Weiss, and Zhaohuan Zhu for helpful discussions.

I have truly enjoyed my time at Imperial. I would like to express my gratitude to the members (past and present) of the Imperial Astrophysics group, for the friendly and engaging working environment.

Special thanks to my family and friends for their continued support and encouragement. Finally, and above all, thank you, Vanja, for your love and support.

Contents

Abstract **3**

Acknowledgements **4**

1 Introduction **9**

1.1 Extrasolar planets . . . 9

1.2 Protoplanetary discs . . . 12

1.2.1 Thin circumstellar discs . . . 15

1.2.2 Viscous accretion discs . . . 16

1.2.3 Source of viscosity . . . 18

1.2.4 Evolution of dust . . . 21

1.3 Formation of close-in super-Earths . . . 23

1.3.1 Migration scenario . . . 24

1.3.2 Formation in the inner disc . . . 25

1.4 Thesis outline . . . 27

2 MRI-accreting inner disc **28**

2.1 Introduction . . . 28

2.2 Methods . . . 29

2.2.1 Standard α-disc model . . . 29

2.2.2 MRI-driven viscosity parameter α . . . 30

2.2.3 Self-consistent α-disc model . . . 35

2.3 Results . . . 36

2.3.1 Fiducial model . . . 36

2.3.2 Varying model parameters . . . 47

2.4 Discussion and conclusions . . . 54

3 Early stages of planet formation in the inner disc **56**

3.1 Introduction . . . 56

3.2 Methods . . . 57

3.2.1 Gas disc model . . . 57

3.2.2 Dust evolution model . . . 58

3.2.3 Numerical methods . . . 60

3.3 Results . . . 61

3.4 Implications for planetesimal formation . . . 64

3.5 Discussion and conclusions . . . 67

4 Atmospheres of planets formed in the inner disc 69
4.1 Introduction 69
4.2 Methods 70
4.2.1 Accretion of planetary atmospheres 70
4.2.2 Photoevaporation of planetary atmospheres 71
4.3 Results 72
4.3.1 Accretion of planetary atmospheres 72
4.3.2 Photoevaporation of planetary atmospheres 74
4.4 Comparison to observations 74
4.5 Discussion and conclusions 79

5 Improved model of the MRI-accreting inner disc 81
5.1 Introduction 81
5.2 Methods 82
5.2.1 The disc model 82
5.2.2 Opacities 85
5.2.3 Viscosity 87
5.2.4 Ionization 88
5.2.5 Numerical methods 92
5.3 Results 95
5.3.1 Disc thermal structure and the MRI 96
5.3.2 Disc chemical structure and the MRI 102
5.4 Discussion 109
5.4.1 Effects of dust 109
5.4.2 Importance of stellar irradiation 110
5.4.3 Convective instability in the inner disc 111
5.4.4 Energy transport by turbulent elements 112
5.4.5 Ambipolar diffusion in the strong-coupling regime 112
5.5 Conclusions 113

6 Dependence on dust, disc and stellar parameters 115
6.1 Introduction 115
6.2 Methods 116
6.3 Results 118
6.3.1 Dust-to-gas ratio and dust size 119
6.3.2 Gas accretion rate, stellar mass and dead-zone viscosity 124
6.4 Location of the pressure maximum 125
6.5 Discussion 129
6.5.1 Dust growth 129
6.5.2 Dust accumulation 133
6.6 Conclusions 135

7 Summary and Outlook 137
7.1 Summary 137
7.2 Outlook 140

Bibliography 142

List of Figures

1.1 Extrasolar planets with known mass or radius . . . 11
1.2 Mass-radius relationship for a set of extrasolar planets . . . 13
1.3 Radius distribution for short-period planets . . . 14
1.4 Illustration of the onset of the magneto-rotational instability . . . 19
1.5 Illustration of the MRI active and dead zones . . . 20
1.6 Illustration of a dust trap . . . 22
1.7 Illustration of key transitions in protoplanetary discs . . . 23

2.1 Fractional ionization in the inner disc . . . 37
2.2 Ambipolar, Ohmic and Hall resistivities . . . 38
2.3 Relative importance of the Ohmic, Hall and ambipolar resistivities . . . 39
2.4 The MRI-active, dead, zombie and Hall zones for the fiducial model . . . 40
2.5 Midplane fractional ionization and MRI criteria . . . 41
2.6 Magnetic field strength for the fiducial model . . . 41
2.7 Vertically-averaged viscosity parameter for the fiducial model . . . 42
2.8 Disc structure for the fiducial model . . . 43
2.9 Surface density profile for various accretion rates . . . 44
2.10 Rate of change of the surface density with the accretion rate . . . 45
2.11 Rates of change of viscosity parameter and sound speed with the accretion rate . 46
2.12 Rosseland-mean opacity . . . 46
2.13 Comparison of the recombination timescale and the dynamical timescale . . . 47
2.14 MRI zones for different dead-zone viscosity parameters . . . 48
2.15 Magnetic field strength and viscosity parameter for different dead-zone viscosity parameters . . . 49
2.16 Disc structure for different dead-zone viscosity parameters . . . 49
2.17 Radial location of the pressure maximum and the dead-zone viscosity parameter 49
2.18 MRI zones for different accretion rates . . . 50
2.19 Relative importance of the resistivities for different accretion rates . . . 51
2.20 Magnetic field strength and viscosity parameter for different accretion rates . . . 51
2.21 Disc structure for different accretion rates . . . 52
2.22 Radial location of the pressure maximum and the accretion rate . . . 52
2.23 MRI zones for $M_* = 0.1\,\mathrm{M}_\odot$ . . . 53
2.24 Magnetic field strength and viscosity parameter for $M_* = 0.1\,\mathrm{M}_\odot$ . . . 53
2.25 Disc structure for $M_* = 0.1\,\mathrm{M}_\odot$ . . . 54

3.1 Gas disc structure . . . 59
3.2 Dust-to-gas ratio . . . 61
3.3 Dust particle size limits . . . 63
3.4 Criteria for streaming and gravitational instabilities . . . 66

4.1 Envelope mass fraction of accreted atmospheres 72
4.2 Envelope mass fractions of the atmospheres after accounting for photoevaporation 75
4.3 Mass-radius relationship for sub-Neptune planets 76
4.4 Sub-Neptune planets with measured masses and radii and predictions from models 78

5.1 Opacities for micron grains 86
5.2 Ionization fraction for various effective dust-to-gas ratios 91
5.3 Comparison of temperature in models of varying complexity 97
5.4 Radiative and convective zones in the inner disc 97
5.5 Ratio of the total irradiation heating to the total viscous dissipation 98
5.6 Disc structure for models with and without irradiation 99
5.7 Comparison of the local viscosity parameter in models of varying complexity . . 100
5.8 Comparison of the viscosity parameter and magnetic field strength in models of varying complexity 101
5.9 Comparison of models with various thermal and non-thermal sources of ionization 103
5.10 Disc structure for the model with thermal ionization and dust effects 104
5.11 Disc structure for the model with all sources of ionization 106
5.12 Local viscosity parameter for the model with all sources of ionization 106
5.13 Ionization rates due to stellar X-rays and cosmic rays 107
5.14 Degeneracy in the vertical disc structure 108
5.15 Comparison of the recombination timescale and the dynamical timescale 113

6.1 Opacities for various dust grain sizes 117
6.2 Opacities as a function of dust grain size 118
6.3 Results for the thermally-ionized disc for various dust properties 120
6.4 Results for the full model for various dust properties 122
6.5 Viscosity parameter and ionization levels for a high dust-to-gas ratio 124
6.6 Results for the full model for various disc and stellar parameters 126
6.7 Radial location of the pressure bump as a function of maximum dust grain size . 128
6.8 Radial location of the pressure bump, grain size and Stokes number in the fragmentation limit 132

1 INTRODUCTION

The past few decades have seen unprecedented advances in planetary science, starting with the first discovery of a planet orbiting a main sequence star other than the Sun in 1995. Thousands more extrasolar planets have been discovered since then. It emerged that most stars host planetary systems that are quite dissimilar to the Solar system. By far the most prevalent planets seem to be of size between that of the Earth and Neptune, in orbits typically smaller than Mercury's. The abundance of these so-called close-in super-Earths has opened many new questions in the theory of how planets form.

All planets are believed to have formed in discs of gas and dust that surround young, newly-formed stars, the so-called protoplanetary discs. Our understanding of protoplanetary discs has also greatly advanced due to modern observational facilities. However, observations show us that planets are incredibly common at small orbital radii, whereas the structure of protoplanetary discs at such small radii is difficult to observe due to large distances to these objects. In this thesis I aim to bridge this gap by building theoretical models of the structure of the innermost regions of protoplanetary discs, and examining how they set the initial conditions for planet formation.

This chapter starts with a brief overview of the currently known extrasolar planet population and properties of the close-in super-Earths. Then, I introduce the basic concepts in the theory of protoplanetary discs, focusing on the special conditions that arise in the innermost regions of these objects. Last, I discuss our current understanding of how the super-Earths might have formed, and outline the new work presented in this thesis.

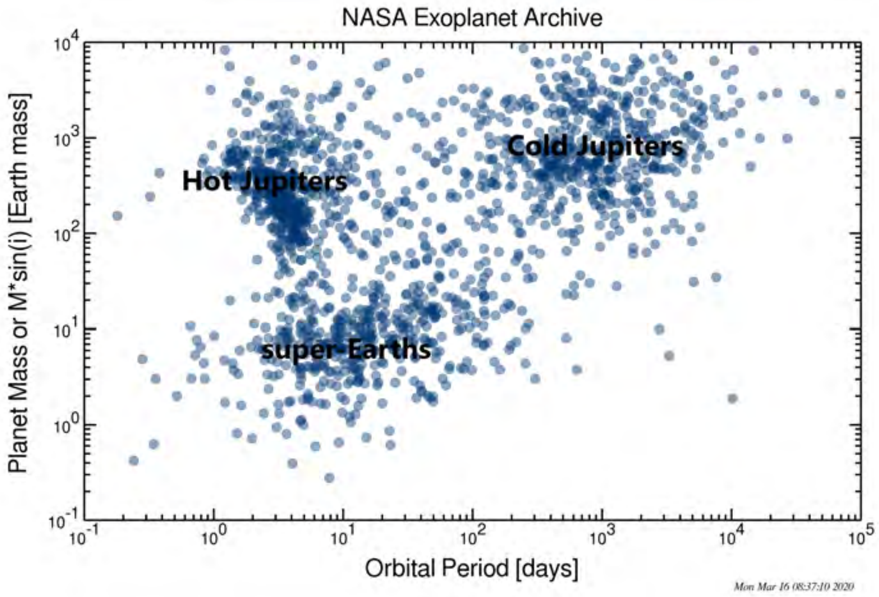
1.1. Extrasolar planets

Two methods are responsible for the vast majority of extrasolar planet (or exoplanet) discoveries so far: radial velocity and transit photometry. A planet-hosting star orbits around their joint centre of mass, periodically moving forwards and backwards relative to an observer. The radial velocity method detects this motion by measuring the Doppler shift in the light emitted by the star. This is a function of planet mass and the system inclination relative to the observer, and the radial velocity method obtains a lower limit on the planet mass. Transit photometry measures a decrease in the light received from the star when a planet transits over the stellar disc. It yields the planet radius. Therefore, in general, we may only know either the (minimum) mass or the radius of a planet. Where possible, planets are detected with one method and confirmed with the other, providing both. A radial-velocity mass of a transiting planet is also its actual mass, since inclination is known. Additionally, in some cases it is possible to obtain masses of planets through transit photometry only. In multi-planet systems planets interact gravitationally and this can cause small transit timing variations (TTVs). Precise measurements coupled to dynamical modelling can then be used to infer planet masses (Agol et al. 2005; Holman and Murray 2005).

Unfortunately, both mass and radius are known for only a small number of exoplanets. Therefore, Fig. 1.1 shows planet (minimum) mass as a function of orbital period in one panel, and planet radius in the other, for sets of planets for which these quantities have been determined (and also includes a small portion of planets discovered through other, less common detection methods). Both panels illustrate the diversity of the extrasolar planets discovered so far. Three groupings of planets are evident in the planet mass-period plot: short-period (Hot) and long-period (Cold) Jupiter-sized giants, and short-period super-Earths. Neither Hot Jupiters nor super-Earths have analogues in the Solar system.

These plots are, naturally, highly biased by the sensitivity of the detection methods. Such biases have been well characterized for the most successful planet finder to date, NASA's *Kepler* mission (Borucki et al. 2010). The main biases introduced by the transit photometry survey by *Kepler* are incompleteness due to the requirement that a planet's orbit is aligned with the observer's line of sight, false positives (i.e., non-planet transits), incompleteness due to the transit-detecting algorithm missing planet transits in the data and uncertainties in stellar parameters (e.g. Fressin et al. 2013; Dressing and Charbonneau 2013; 2015). By removing these biases it is possible to calculate the occurrence rates, i.e., the average number of planets of certain properties per star. Such analyses have shown that planets around other stars are very common. In an early study based on the first 16 months of *Kepler* data Fressin et al. (2013) estimated that there are 0.7 planets per star with orbital periods up to 85 days, the vast majority of which (0.65 planets per star) are planets with radii in the range $0.8 - 4\,\mathrm{R}_{\oplus}$, i.e. super-Earths. This is roughly consistent with estimates from radial velocity surveys that the occurrence rate of Hot Jupiters is in the range 0.1%-1% and the occurrence rate of Cold Jupiters of the order of several 1% (Mayor et al. 2011; Santerne et al. 2016). Unsurprisingly, the actual detected population (shown in Fig. 1.1) is highly biased towards giant planets.

Estimates based on the full four-year *Kepler* data set yield even higher occurrence rates for the super-Earths. Focusing on M dwarf stars, Dressing and Charbonneau (2015) estimated 2.5 planets per star in the super-Earth size range and orbital periods up to 200 days. More recent studies estimated that for orbital periods of $0.5 - 256$ days there are 3.5 super-Earths per star for FGK type stars and between 4.2 and 8.4 super-Earths per star for M dwarfs (Hsu et al. 2019; 2020). As can be inferred from these high occurrence rates, the super-Earths readily appear in multi-planet systems (e.g. Borucki et al. 2011). It is estimated that at least 42% of all Sun-like stars have planet systems with at least 7 planets (Mulders et al. 2018) that are



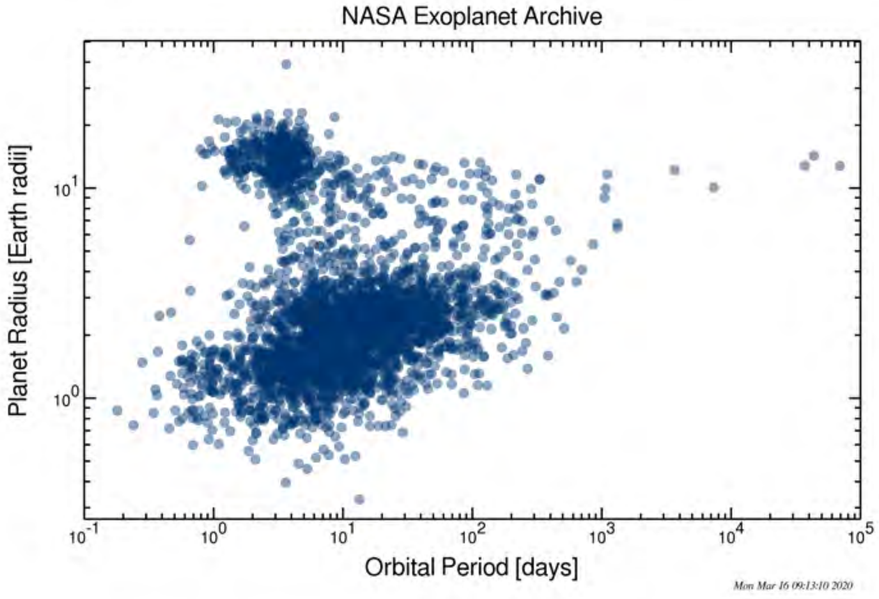


Figure 1.1: Extrasolar planets with known mass (or the lower limit on mass, see text; top) and with known radius (bottom) as functions of orbital period. Note the three groupings of planets in the mass-period plot: super-Earths, Hot Jupiters and Cold Jupiters. Obtained from NASA Exoplanet Archive. Courtesy NASA/JPL-Caltech.

nearly coplanar.

Evidently, the close-in super-Earths are incredibly common around low-mass stars. What are these planets like? Some indications of the structure and composition of super-Earths are gained from those planets for which both radius and mass have been determined. Fig. 1.2 shows the mass-radius relationship for the observed planets and also theoretical predictions for planets made purely of iron, silicates or water ice. Importantly, many of the planets have radii well above the pure water ice radius at a given mass, indicating that they must contain a significant volume of gas on top of a solid core, most likely made of hydrogen and helium (as found by e.g. Marcy et al. 2014; Weiss and Marcy 2014; Rogers 2015). A small amount of mass is needed in a H/He envelope to make a large difference in planet radius. Assuming that the super-Earths have rocky cores, Wolfgang and Lopez (2015) estimate that only about 1% to 10% percent of planet mass is needed in the envelope to explain the radii of the super-Earths.

In the presence of a gaseous envelope, determination of the core composition becomes a degenerate problem, e.g., it cannot be determined how much, if any, water ice is present in the core. The bottom panel of Fig. 1.2 zooms into a region of the radius-mass plane showing small planets. There appears to be a transition radius of about $1.5\,\mathrm{R}_\oplus$, above which most of the planets are of density so low that they must contain large amounts of volatiles (Rogers 2015). Most of the planets smaller than about $1.5\,\mathrm{R}_\oplus$ appear consistent with an iron-rock (Earth-like) composition (Dressing et al. 2015), and those that do not, have large mass uncertainties (Jontof-Hutter 2019).

Fortunately, some advances can be gained even without the mass measurements, using the radius distribution alone. Fig. 1.3 shows a de-biased radius distribution of the close-in planets based on the *Kepler* survey. It shows the relative scarcity of giant close-in planets relative to the super-Earths. It also reveals that the super-Earths are, in fact, made up of two populations. This gap in the radius distribution of close-in super-Earths had been predicted by models of atmospheric evolution (Owen and Wu 2013). These models predict that all close-in super-Earths were initially formed with large H/He envelopes; however, the ones that are least massive and closest to the star lost their envelopes due to photoevaporation by high-energy stellar flux. The location of the gap in these models is sensitive to the composition of the solid planet core, and the observed gap is consistent with an Earth-like (i.e., rock-iron) composition (Owen and Wu 2017; Van Eylen et al. 2018; Wu 2019).

The close-in super-Earths stand in stark contrast to the inner region of the Solar system, as they typically have orbital periods shorter than Mercury's, while many of them are more massive than all of the terrestrial planets combined. Given their occurrence rate around low-mass stars, explaining how these planets form is undoubtedly a vital task.

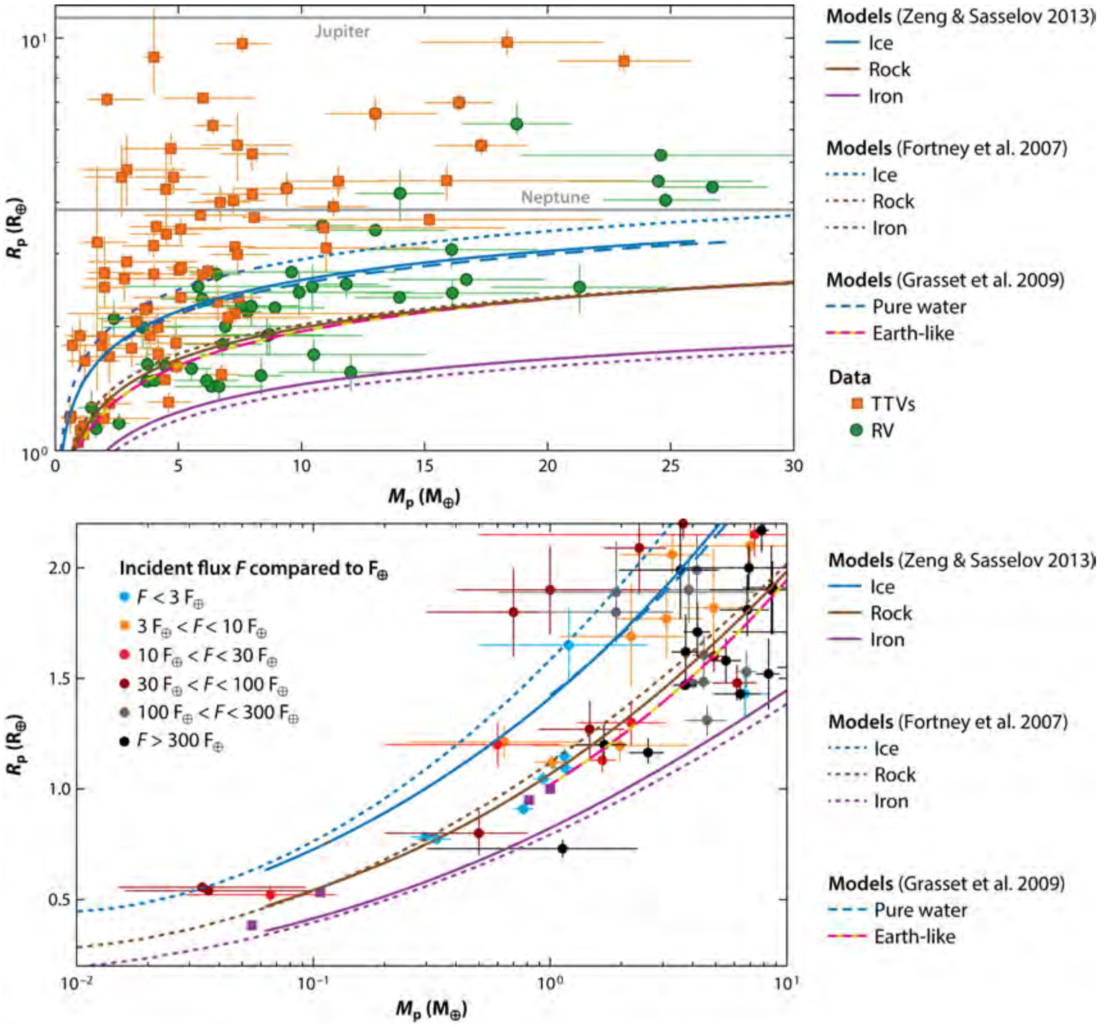


Figure 1.2: Planet radius as a function of planet mass for a set of extrasolar planets with theoretical predictions for pure iron, silicate rock and water ice planets (top) and the same figure zoomed into small planets (bottom). In the top panel, orange squares show planets whose masses are determined using TTVs and green circles planets whose masses are determined using the radial velocity method. In the bottom panel, planet data are colour-coded according to the incident flux, as indicated in plot legend. In both panels, solid and dashed lines show theoretical predictions, as indicated in plot legend (Zeng and Sasselov 2013; Fortney et al. 2007; Grasset et al. 2009). Republished with permission of Annual Reviews, Inc., from Jontof-Hutter (2019); permission conveyed through Copyright Clearance Center, Inc.

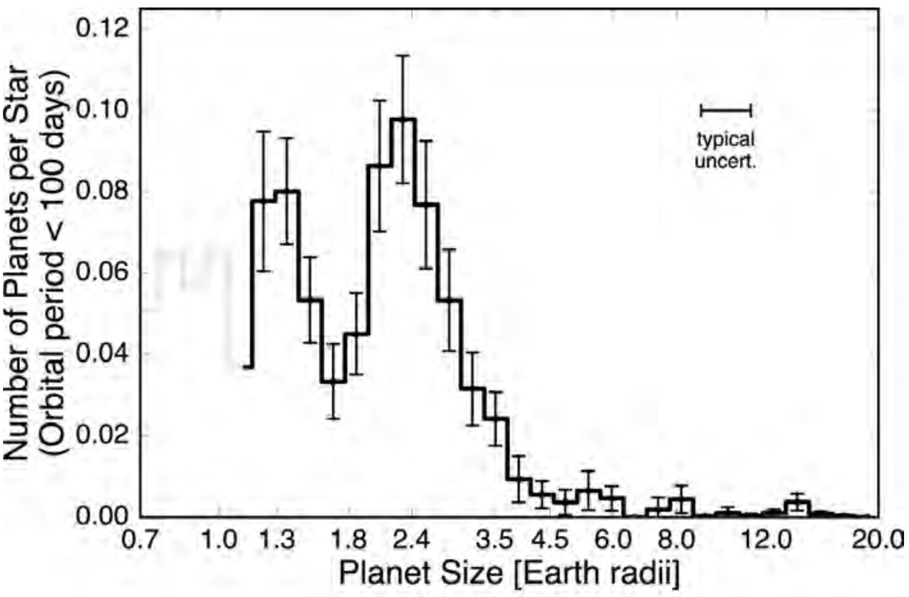


Figure 1.3: Completeness-corrected radius distribution for planets with orbital periods shorter than 100 days (solid black line) including the small radii for which completeness is low (solid grey line) and the uncorrected radius distribution (dotted grey line). The uncertainties in the occurrence rates are based on simulated planet populations. The median uncertainty in radius is shown in the upper right. Adapted from Fulton et al. (2017). ©AAS. Reproduced with permission.

1.2. Protoplanetary discs

Stars form when parts of cold dense clouds of interstellar matter (ISM) gravitationally collapse. Since these clouds are slowly rotating, their gravitational contraction combined with conservation of angular momentum also results in the formation of a disc around the newly-formed star (e.g. Terebey et al. 1984). It is believed that planets form in these discs, and hence they are commonly referred to as protoplanetary discs.

Since they form out of ISM, protoplanetary discs are expected to start out with 99% of their mass in gas and 1% of their mass in solid dust grains (Mathis et al. 1977). The dust grains absorb stellar light in the optical waveband and re-emit it in the infrared. The resulting infrared excess in the spectral distribution (SED) of young stellar objects has long been detected (e.g. Mendoza 1966, 1968) and its dusty disc origin understood (e.g. Adams et al. 1987). Inferred disc masses are in the range of $10^{-3} - 10^{-1}\,\mathrm{M}_{\odot}$ (Beckwith et al. 1990; Beckwith and Sargent 1991; Andrews and Williams 2005; 2007). The disc-like geometry of these objects was first confirmed with Hubble Space Telescope images of light scattered off dust grains (O'dell and Wen 1994). Dust thermal emission was resolved using millimetre interferometry, revealing dust disc sizes of few hundreds of AU (Dutrey et al. 1996). The more abundant gas component of these discs is more difficult to detect, since it is largely made up of relatively cold molecular hydrogen. Nevertheless, observations of molecular lines of trace species such as CO yield insight into the kinematics of these objects, which is to first order determined by simple Keplerian rotation around the central star (e.g. Sargent and Beckwith 1987; Simon et al. 2000).

Protoplanetary discs accrete onto the star. Material flows from the inner disc edge onto the star along magnetic field lines, emitting broad atomic emission lines (Muzerolle et al. 1998; 2001). As the material falls onto the stellar surface it creates shocks and emits UV continuum radiation (Gullbring et al. 1998). Both effects can be used to infer the mass accretion rate onto the star, which is typically in the range of $10^{-10} - 10^{-6}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ (Hartmann et al. 1998). While these measurements only constrain accretion within a few stellar radii from the star, the obtained accretion rates are found to be correlated with the observed total disc dust mass, implying that the disc evolves as a whole (Manara et al. 2016).

Eventually, protoplanetary discs disperse. The vast majority of the youngest stars host discs, but very few stars older than a few Myr have one (Haisch et al. 2001; Mamajek 2009). Importantly, the disc lifetime sets the timescale for planet formation. Any planet with a H/He atmosphere accreted from the protoplanetary disc, including the close-in super-Earths, had to have formed while the gaseous disc was still around. Furthermore, accretion in protoplanetary discs is detected up to somewhat smaller, but comparable stellar ages as the infrared excess (Fedele et al. 2010). This means that accretion and its effects on disc structure are important during planet formation.

In the remainder of this section I will discuss the basics of the structure and evolution of protoplanetary discs, starting with some basic notions about the structure of thin circumstellar discs. I will discuss the basic theory of how accretion controls the disc evolution and the processes driving accretion, focusing on the short orbital distances at which the super-Earths are observed. Last, I will discuss theoretical expectations and observational evidence of dust growth and evolution in protoplanetary discs.

1.2.1. Thin circumstellar discs

Throughout this work I will consider axisymmetric discs, i.e. I will assume that rotational shear azimuthally smooths all disturbances. Although there are certainly some exceptions to this among the observed discs, the majority are indeed axisymmetric. Protoplanetary discs are also observed to be geometrically thin, i.e., their height is much smaller than their radius. It is thus commonly assumed that in the time necessary for the disc to evolve over large radial distances, the disc vertical structure settles into hydrostatic equilibrium. As most discs are observed to be much less massive than their host star, I will also neglect the disc self-gravity. Then, in the vertical direction, thermal pressure balances stellar gravity,

$$\frac{dP}{dz} = -\rho \frac{GM_* z}{(r^2 + z^2)^{3/2}}, \tag{1.1}$$

where P is gas thermal pressure, ρ is disc density and M_* stellar mass. For a thin disc in which $z \ll r$, adopting the ideal gas law, the vertical disc structure is determined by

$$\frac{dP}{dz} = -\frac{P\Omega_{\rm K}^2 z}{c_{\rm s}^2}, \tag{1.2}$$

where $c_{\rm s}$ is the speed of sound, and $\Omega_{\rm K}$ is the Keplerian angular velocity.

The disc angular velocity Ω is not strictly Keplerian. Even in a disc that is only slowly evolving in the radial direction ($u_r \ll r\Omega$), Ω is affected by the radial gas thermal pressure gradient,

$$r\Omega^2 = r\Omega_{\rm K}^2 + \frac{1}{\rho}\frac{dP}{dr}. \tag{1.3}$$

This term can be safely neglected when considering gas disc evolution, so that $\Omega = \Omega_{\rm K}$. However, it is hugely important in the evolution of dust in circumstellar discs, as discussed further below. Here I have also assumed that the disc angular velocity Ω is constant with height, since the disc is thin.

1.2.2. Viscous accretion discs

In order for disc material to accrete onto the star it has to lose angular momentum. How exactly this happens in protoplanetary discs is an unsolved problem. One possibility is that a shear stress redistributes angular momentum throughout the disc. Consider a differentially rotating disc as a series of infinitesimal rings and a Newtonian shear stress of the form

$$W_{r\phi} = \rho\nu r\frac{\partial\Omega}{\partial r}, \tag{1.4}$$

where ν is kinematic viscosity. In a Keplerian disc such shear stress between two rings will slow down the inner ring and speed up the outer. Material slowing down in the gravitational field of the star will flow inwards, accreting, and matter speeding up will move outwards, carrying angular momentum away. In protoplanetary discs molecular viscosity is negligible. However, as discussed further below, other processes can produce the same behaviour.

The theory of evolution of viscously accreting discs was set out by Shakura and Sunyaev (1973) and Lynden-Bell and Pringle (1974). From conservation of angular momentum it follows that

$$\rho u_r \frac{1}{r}\frac{\partial(r^2\Omega)}{\partial r} = (\nabla W)_\phi \tag{1.5}$$

in a thin axisymmetric disc. For the viscous stress given by eq. (1.4),

$$\rho u_r \frac{1}{r}\frac{\partial(r^2\Omega)}{\partial r} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\rho\nu r\frac{\partial\Omega}{\partial r}\right). \tag{1.6}$$

Here viscosity ν can be a function of both height above disc midplane and of radius. The equation of mass continuity is

$$\frac{\partial \rho}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}\left(r\rho u_r\right) = 0. \tag{1.7}$$

Vertically integrating eq. (1.7) and (1.6) yields

$$\frac{\partial \Sigma}{\partial t} - \frac{1}{2\pi r}\frac{\partial \dot{M}}{\partial r} = 0, \tag{1.8}$$

$$-\frac{1}{2\pi}\dot{M}\frac{\partial}{\partial r}\left(r^2\Omega\right) = \frac{\partial}{\partial r}\left(r^2\Sigma\bar{\nu}r\frac{\partial \Omega}{\partial r}\right), \tag{1.9}$$

where $\dot{M} = -2\pi r\int_{-\infty}^{\infty} dz\rho u_r$ is the inwards gas accretion rate, $\Sigma = \int_{-\infty}^{\infty} dz\rho$ is the gas surface density and $\bar{\nu} = \int_{-\infty}^{\infty} dz\rho\nu/\Sigma$ is the vertically-averaged viscosity.

Therefore, in a Keplerian disc, the gas accretion rate is given by

$$\dot{M} = \frac{6\pi}{r\Omega}\frac{\partial}{\partial r}\left(r^2\Omega\bar{\nu}\Sigma\right). \tag{1.10}$$

Substituting eq. (1.10) into the mass continuity equation yields

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{r}\frac{\partial}{\partial r}\left(r^{1/2}\frac{\partial}{\partial r}\left(r^{1/2}\bar{\nu}\Sigma\right)\right), \tag{1.11}$$

i.e., that time evolution of the gas surface density is governed by what is essentially a diffusion equation. Time-dependent solutions to this equation are given by Lynden-Bell and Pringle (1974) for some special cases of viscosity. For example, they show that for constant viscosity or viscosity that is a power-law function of radius, gas flows inwards onto the star at the inner disc edge, and viscously spreads at the outer edge. In this work I will primarily consider steady-state solutions, in which the gas accretion rate $\dot{M}$ is radially constant. Radially integrating eq. (1.9) from the inner disc edge, $r_{\rm in}$, to an arbitrary radius r,

$$-\frac{1}{2\pi}\dot{M}(r^2\Omega - r_{\rm in}^2\Omega_{\rm in}) = r^2\Sigma\bar{\nu}r\frac{\partial \Omega}{\partial r}\Big|_{r_{\rm in}}^{r}. \tag{1.12}$$

Here, the right-hand side requires specifying a boundary condition at the inner disc edge. In reality this depends on the complicated nature of gas flows and the magnetic field between the inner disc edge and stellar surface (Bouvier et al. 2007). A simple way to specify this boundary condition is to assume that the boundary layer between the disc and the star is thin ($r_{\rm in} \sim R_*$, e.g. Frank et al. 2002). Within this thin layer the angular velocity of material has to decrease from its Keplerian value in the disc to the rotation velocity of the star, and so at some point in the boundary layer $\partial\Omega/\partial r = 0$. Therefore, in steady state,

$$\dot{M} = 3\pi\bar{\nu}\Sigma f_r^{-1}, \tag{1.13}$$

where $f_r = 1 - \sqrt{R_*/r}$.

Finally, to obtain the full description of the system, one needs to consider the energy balance. If the only source of heat is dissipation due to the viscous stress, then at any point in the disc the viscous dissipation rate and the cooling rate are balanced, i.e.

$$\frac{9}{4}\rho\nu\Omega^2 = \nabla.\mathbf{F}, \tag{1.14}$$

where $\mathbf{F}$ is the energy transport flux, e.g. radiative flux. In a thin disc one can expect the heat to escape vertically much more easily than in the radial direction, so that $F \approx F_z$. Furthermore, at disc midplane $F = 0$ due to symmetry. Then, vertically integrating eq. (1.14) from disc midplane to disc surface yields

$$F_{\mathrm{tot}} = \frac{9}{8}\Sigma\bar{\nu}\Omega^2 = \frac{3}{8\pi}\Omega^2\dot{M}f_r, \tag{1.15}$$

where F_{tot} is the total heat flux that escapes through one side of the disc and the second equality is obtained by substituting $\Sigma\bar{\nu}$ from eq. (1.13). Therefore, the total viscous dissipation at a given radius depends only on the gas accretion rate $\dot{M}$ and stellar parameters.

Further progress can only be made by specifying the energy transport flux F and the viscosity ν. In radiative energy transport, flux F is a function of disc structure and disc radiative properties, i.e. absorptivity and emissivity of the material. This inter-dependency will be considered with varying degrees of complexity in this thesis, and so I defer further discussion to relevant chapters. As for the viscosity ν, it was already noted that it must arise from pseudo-viscous processes in the disc. Its origin is discussed in the next section.

1.2.3. Source of viscosity

Shakura and Sunyaev (1973) argued that viscosity in accretion discs comes from turbulence. Essentially, correlated fluctuations in radial and azimuthal velocities due to turbulent motions yield a shear stress analogous to the viscous stress given by eq. (1.4). If a magnetic field is present, the same is true of fluctuations in the radial and azimuthal components of the field. In both cases it is convenient to assume that the shear stress is proportional to the gas pressure, with a dimensionless proportionality constant α. In the case of turbulent velocity fluctuations, for example, α is a square of the ratio of the turbulent velocity to the speed of sound (e.g. Balbus and Hawley 1998). Then, since supersonic turbulence would lead to shocks and quickly dissipate, Shakura and Sunyaev argue that $\alpha < 1$, placing a constraint on the magnitude of shear stress in discs. Furthermore, shear stress is given by eq. (1.4) if viscosity ν is given by

$$\nu = \alpha\frac{c_{\mathrm{s}}^2}{\Omega}, \tag{1.16}$$

connecting the Shakura and Sunyaev α parameter to the above formulation of a viscous accretion disc.

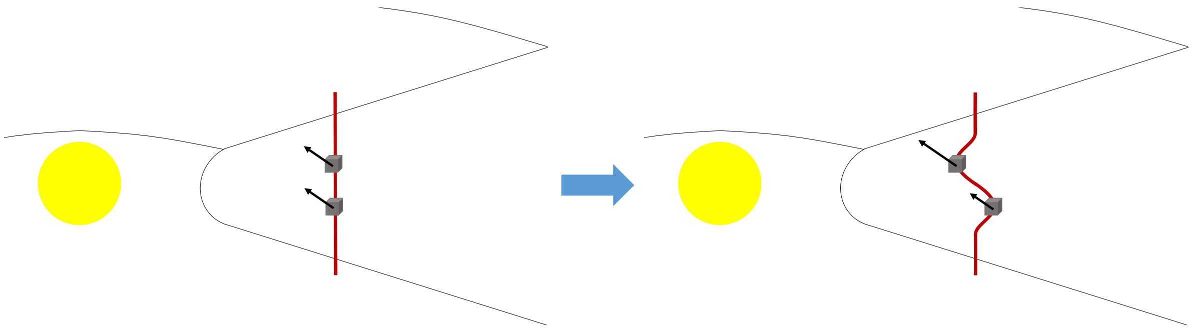


Figure 1.4: Illustration of the onset of the magneto-rotational instability. Two fluid elements (grey cubes), coupled to the magnetic field (red line), are initially at the same orbital radius (left), rotating around the star at the same azimuthal velocity (black arrows). Following a perturbation in the radial direction, the inner element has a higher azimuthal velocity than the outer, and the magnetic field line is stretched (right). This results in the growth of the perturbation as discussed in Section 1.2.3.

What instability or instabilities may give rise to turbulence in protoplanetary discs? First, consider a well ionized disc, so that the gas is perfectly coupled to the magnetic field and the induction equation is

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}), \tag{1.17}$$

where **B** is the magnetic field strength. In this so-called ideal magnetohydrodynamic (MHD) regime, a rotating disc threaded by a magnetic field is unstable to small perturbations if $d\Omega/dr < 0$, i.e. if the angular velocity decreases outwards (Balbus and Hawley 1991). The onset of the magneto-rotational instability (MRI) is illustrated in Fig. 1.4 for an initially vertical magnetic field. A small perturbation leads to a radial displacement of two fluid elements. Due to the resulting difference in their angular velocities, with the outer element moving slower than the inner one, the fluid elements move apart azimuthally. This stretches the magnetic field line (shown in red) coupled to the fluid elements. Magnetic tension acts to reduce this stretch. However, pulling the inner fluid element back reduces its angular momentum, causing it to fall further inwards, and *vice versa* for the outer fluid element. As a result, the instability grows and at the same time transports angular momentum outwards, as needed for accretion. Furthermore, for the instability to develop, the magnetic field strength can be arbitrarily weak, but not arbitrarily strong - roughly, the magnetic pressure needs to be smaller than the gas pressure. The instability also develops quickly, with the growth timescale being of the order of the local orbital period in the linear regime.

In the non-linear regime, explored through numerical simulations, the MRI leads to MHD turbulence. Shear stress is measured in simulations by averaging over turbulent fluctuations in velocity and magnetic field strength. The viscosity parameter α due to the MRI is found to be in the range $10^{-3} - 10^{-1}$ in homogeneous local simulations threaded by a uniform magnetic field (Hawley and Balbus 1992; Hawley et al. 1995), homogeneous local zero net magnetic flux simulations (Hawley et al. 1996), and stratified local zero net flux simulations (Brandenburg et al. 1995; Stone et al. 1996). These results were confirmed more recently to be independent of spatial resolution in the simulations (Simon et al. 2009; Davis et al. 2010; Shi et al. 2010).

Therefore, if protoplanetary discs were well ionized, they would easily become turbulent due to the MRI. However, ionization levels are expected to be low in these discs and non-ideal MHD effects are expected to play a significant role in disc evolution. For the conditions relevant in protoplanetary discs, collisional coupling between neutral and charged particles remains strong. Because of this, the Lorentz force effectively acts on the neutrals too, i.e., on the fluid as a whole. However, when the fluid is weakly ionized, the coupling between the fluid and the magnetic field weakens. The induction equation becomes (Balbus 2011)

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) - \nabla \times (\eta_{\mathrm{O}}(\nabla \times \mathbf{B}) + \eta_{\mathrm{A}}(\nabla \times \mathbf{B})_{\perp} + \eta_{\mathrm{H}}(\nabla \times \mathbf{B}) \times \hat{\mathbf{B}}). \tag{1.18}$$

Here, $\hat{\mathbf{B}}$ is a unit vector and "$\perp$" denotes the component of a vector perpendicular to $\mathbf{B}$. The first term on the right-hand side is the advective term, which is the same as in the ideal case given by eq. (1.17). The second and the third terms are due to Ohmic and ambipolar diffusion, proportional to the Ohmic resistivity η_{O} and the ambipolar resistivity η_{A}, respectively. If Ohmic diffusion is important, collisions of charged particles with neutrals decouple the charged particles (and thus the fluid overall) from the magnetic field. If ambipolar diffusion is important, charged particles remain coupled to the magnetic field, but the collisions with the neutrals are not efficient enough to couple the entire fluid to the field. By decoupling the fluid from the magnetic field, Ohmic and ambipolar diffusion stabilize the disc against the MRI. Finally, the fourth term is due to the Hall effect. While it is also proportional to what is dimensionally a resistivity η_{H}, it is not a diffusive term.

The fluid ionization fraction at which Ohmic diffusion, for example, suppresses the MRI is of the order of $x \sim 10^{-13}$ (e.g. Fromang et al. 2013). However, the ionization levels in protoplanetary discs can be even lower. The main sources of ionization are typically considered to be thermal (collisional) ionization, stellar X-rays and cosmic rays. Thermal ionization only becomes important above about a 1000 K, the ionization temperature of potassium (Umebayashi and Nakano 1988). It is thus only relevant in the innermost, hot regions of protoplanetary discs. Stellar X-rays and cosmic rays can only penetrate mass columns of the order of $10\,\mathrm{g\,cm^{-2}}$ and $100\,\mathrm{g\,cm^{-2}}$, respectively (Glassgold et al. 1997; Umebayashi and Nakano 1981). Since the disc surface density is expected to be higher than such mass columns in the bulk of the disc, X-rays and cosmic rays may only be relevant in the uppermost layers, near disc surface, and in the outer regions of the disc. Fig. 1.5 illustrates the structure of the inner protoplanetary disc that emerges from this analysis. In the ionized regions the MRI is expected to be active, driving turbulence and efficient accretion. The cold dense regions, on the other hand, form the so-called dead zone, where ionization is too low to support the MRI (Gammie 1996; Jin 1996).

In the dead zone, in the absence of (the coupling with) the magnetic field, various hydrody-

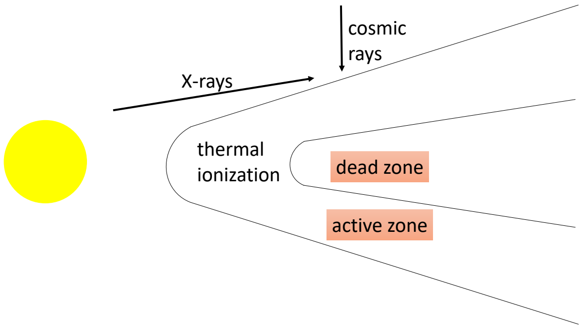


Figure 1.5: Illustration of the inner protoplanetary disc showing key ionization mechanisms and the MRI active and dead zones (Gammie 1996).

namic instabilities can drive turbulence under the right conditions (Pfeil and Klahr 2019; Lyra and Umurhan 2019). However, the resulting viscosity is expected to be far smaller than in the MRI-active zone (e.g. Lesur and Papaloizou 2010; Nelson et al. 2013; Stoll and Kley 2014). Therefore, it is expected that the viscosity, or accretion efficiency, decreases with distance from the star with a potentially sharp drop at the transition between the active and the dead zone.

Finally, given the difficulties in producing turbulence in protoplanetary discs, it must be recognised that accretion might be (at least partially) driven by non-viscous processes. In the bulk of the disc, the Hall effect can be the dominant non-ideal MHD effect. Unlike the diffusive terms, it does not merely suppress the MRI. Rather, the Hall effect can drive large scale laminar flows through the disc (although this depends on the alignment of the magnetic field with the disc rotation axis; Lesur et al. 2014). Additionally, in the presence of a magnetic field threading the disc MHD winds can be launched from the disc surface, lifting angular momentum from the disc (Suzuki and Inutsuka 2009; Suzuki et al. 2010; Moll 2012; Bai and Stone 2013; Fromang et al. 2013; Lesur et al. 2013). It appears likely that both the Hall effect and MHD winds play a significant role in the overall evolution of discs, driving gas accretion at a much larger range of radii than the MRI (Bai 2017). Nevertheless, the structure of the innermost, thermally-ionized regions of discs are still likely to be strongly affected by the MRI, and especially so the disc midplane where planets are expected to form.

1.2.4. Evolution of dust

Protoplanetary discs are expected to start out with small (sub-micron) ISM-like dust grains. Dust thermal emission offers evidence for grain growth to larger sizes. In optically thin regions of the disc the emitted flux is directly proportional to dust emissivity, which itself is a function of dust grain size (Beckwith and Sargent 1991). Comparison of the slope of the disc spectral energy distribution (SED) at millimetre wavelengths with the predicted slope of dust emissivity as a function of wavelength shows growth beyond millimetre sizes (D'Alessio et al. 2001).

Dust grain size determines the dust spatial evolution. Small dust grains are entrained with the gas, well mixed throughout the disc, slowly accreting onto the star. On the other hand, larger dust grains are only partially coupled to the gas through gas drag. As discussed above, as the gas thermal pressure decreases with distance from the star the pressure gradient effectively lowers the gravitational force, making the gas orbital velocity slightly sub-Keplerian. As partially-decoupled dust grains tend towards Keplerian orbits, they feel drag from the slower gas. This causes the dust grains to lose angular momentum and fall radially towards the star. The effect is known as radial drift and it is expected to be most drastic for centimetre-sized dust, as larger bodies become completely decoupled from the gas (Weidenschilling 1977). Indeed, observations show that discs appear significantly smaller in the dust continuum than in the gas line emission, with a steep drop in the dust continuum at disc outer edge, a signature of radial drift (Panić et al. 2009; Andrews et al. 2012; Rosenfeld et al. 2013; Trapman et al. 2020).

Additionally, for dust grains at a finite height above the disc midplane, Keplerian orbits around the central star are inclined relative to the midplane (because the plane of a Keplerian orbit must pass through the centre of the star). The gas drag tends to damp the resulting relative velocity between the dust and the gas in the vertical direction. As a result, dust grains also fall, or settle, vertically towards the disc midplane (Weidenschilling 1977). The dust is not, however, expected to settle in the midplane, as turbulent motions in the disc lift and mix the grains. The dust vertical structure is then determined as an equilibrium between settling and turbulent mixing (Dubrulle et al. 1995).

The radial drift of dust has an important consequence in discs which feature a local gas pressure maximum, see Fig. 1.6. In the region of positive pressure gradient the gas is super-Keplerian, reversing the effect of gas drag on the dust grains. Therefore, just inwards of a pressure maximum, dust grains radially drift outwards. For dust grains that encounter a pressure maximum while radially drifting inwards from the outer disc, the pressure maximum acts as a trap, where over time dust can accumulate (e.g. Haghighipour and Boss 2003; Pinilla et al. 2012). The resulting dust structure in the disc is a dust ring. Many of such dust rings have been observed in protoplanetary discs (e.g. ALMA Partnership et al. 2015; Andrews et al. 2018; Long et al. 2019), with direct kinematic evidence that the dust rings correspond to locations of gas pressure maxima (Teague et al. 2018).

Radial drift, vertical settling, turbulent mixing and, for the smallest grains, Brownian motion, also drive collisions between dust grains (Testi et al. 2014). Dust growth occurs when such collisions result in sticking. Growth by sticking easily produces particles up to millimeters in size; however, collisional velocities increase with grain size and at higher velocities collisions are more likely to result in bouncing or even fragmentation of dust grains (Blum and Münch 1993), impeding growth beyond millimetre sizes (Zsom et al. 2010; Birnstiel et al. 2011). Even if some lucky particles surmount these barriers (Windmark et al. 2012a;b; Garaud et al. 2013), they are prone to fast radial drift towards the star. Nevertheless, in many discs at least some of the dust must somehow form large planet-size bodies (since planets exist).

Finally, it is important to note that the composition of dust grains varies spatially inside

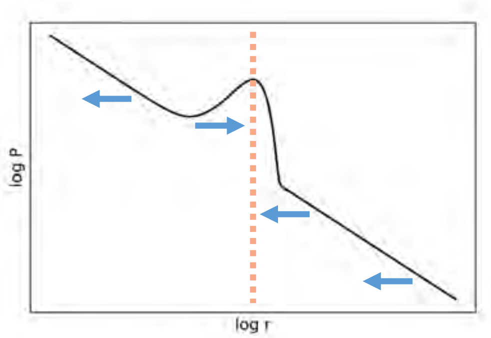


Figure 1.6: Illustration of a radial midplane gas pressure profile in a disc featuring a local maximum (indicated by a dashed line). Blue arrows show directions of radial drift of dust grains embedded in such a disc.

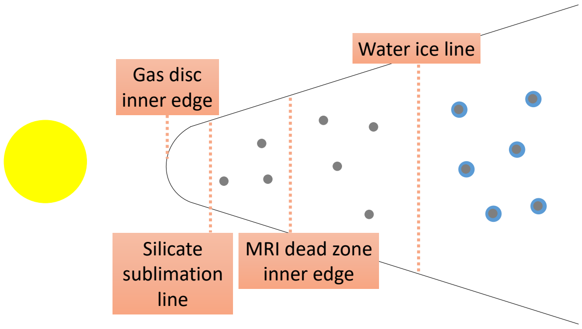


Figure 1.7: Illustration of a protoplanetary disc showing the gas disc inner edge, silicate sublimation line, MRI dead zone inner edge and water ice line.

discs. The temperature in protoplanetary discs can vary from a few tens of Kelvin in the outermost disc to a few thousand near the star, covering the sublimation points of all key materials. Water is only condensed beyond the so-called ice line, where the temperature falls below $\sim$150 K. Furthermore, from observations of disc emission in the near-infrared, very little to no dust is expected to exist inside the sublimation line of silicates where temperatures are higher than $\sim$1500 K (Hillenbrand et al. 1992; Natta et al. 2001). Naturally, spatial variations in the composition of planet-building material leave an imprint on planetary systems. For example, the terrestrial planets and the main asteroid belt in the Solar system are relatively dry, and it is thought that the Earth's entire water content was delivered by scattered water-rich asteroids formed in the outer Solar system (Morbidelli et al. 2000).

1.3. Formation of close-in super-Earths

In this chapter so far I have discussed the known exoplanet population and the structure and evolution of protoplanetary discs, inside which planets are believed to form. From the detected exoplanet population it is inferred that roughly half of all Sun-like and lower mass stars host one or more close-in super-Earth-sized planets. It follows then that protoplanetary discs must efficiently produce such planets. However, how, and even where exactly inside protoplanetary discs, the close-in super-Earths form is an open question.

The main theoretical problem in forming super-Earths is in the large amount of solid mass required to assemble these planets. Assuming that planets form near their present orbital distances (i.e. *in situ*) from locally available material, a model of a protoplanetary disc can be constructed such that contains the minimum amount of dust required to build solid planet cores. For the Solar system, such model is known as the Minimum mass Solar nebula (MMSN; Hayashi 1981). Close to the star the MMSN has sufficient amounts of dust to form the terrestrial planets. To build the close-in super-Earths, a disc would have to be significantly richer in solids at short orbital distances (Raymond et al. 2008; Chiang and Laughlin 2013). While this led some authors to conclude that the super-Earths cannot have formed *in situ* (e.g. Raymond et al. 2008), others proposed that the proto-Solar system might have been an exception rather than the rule, and that the typical protoplanetary disc is indeed more massive (Chiang and Laughlin 2013). However, a minimum mass nebula is an idealization, and growing planets cannot access all of the material available in the disc. After taking this into account, and also considering that solids comprise only about a percent of total disc mass, the required disc mass is found to be close to or above the criterion for gravitational stability of the disc (Schlichting 2014). Other problems also arise if the super-Earths form in discs that are rich in both dust and gas. For example, in a massive gaseous disc, accretion of gas onto super-Earth-sized solid cores is so efficient that the planets would easily accumulate Jupiter-sized atmospheres (Lee et al. 2014), whereas Jupiter-sized planets are much less common than the super-Earths.

There are two main ways to solve this problem, both relying on the fact that the solid-to-gas ratio need not (and almost certainly does not) stay constant throughout disc lifetime. First, when small dust grains grow into pebbles, they radially drift inwards, towards the star, due to gas drag (see Section 1.2.4). If the drift can be stopped or slowed down at short orbital periods, the pebbles accumulate there, enhancing the solid-to-gas ratio. Second, large planet-sized bodies interact with the gaseous disc gravitationally and this can cause them to move (migrate) radially (e.g. Goldreich and Tremaine 1979; 1980; Kley and Nelson 2012). It is thus possible for a planet to form at an orbital distance different to the one we observe today, where the conditions for planet formation are more favourable, following which the planet migrates to its final orbit. The latter, the so-called migration scenario, has been studied in more detail, and so I will discuss it first.

1.3.1. Migration scenario

Planets (or planet embryos) can launch spiral density waves in the disc (Goldreich and Tremaine 1979; 1980). As these waves propagate outwards through the disc, through gas that rotates slower than the planet, they add angular momentum to the gas. The opposite is true for the waves propagating inwards. To conserve angular momentum, a back-reaction torque acts on the planet. Due to a slight asymmetry, regions from which the density waves are launched exterior to the planet are closer to the planet than the ones interior to it. Consequently, the net torque acting on the planet causes it to move inwards. A planet also interacts with the material near its orbit, i.e., the material that corotates with it. The torque that acts on the planet from the corotation region is sensitive to the disc structure, and can also act to drive outwards planet migration (e.g. Paardekooper and Mellema 2006). At some locations then, such as the inner disc edge, the net torque may be zero, and migration stalls (Masset et al. 2006). It is thus possible that super-Earths form at an orbital distance different to the one we observe today, and then migrate inwards through gravitational interactions with the disc (e.g. Terquem and Papaloizou 2007; Ogihara and Ida 2009; Cossou et al. 2014).

Super-Earths might form more easily at larger orbital distances due to several factors. Further away from the star there is more material to form planets. Observations of protoplanetary discs show that the radial surface density profiles are shallow, so the amount of mass available at a given orbital distance increases radially outwards (typically $\Sigma(r) \propto r^{-1}$, e.g. Williams and Cieza 2011, although, note that there is a large spread in the observed $\Sigma(r)$ profiles, and also that such observations are limited to the outer regions of discs). From theoretical models, in the MMSN, for example, there is enough mass to form the super-Earths at a few AU from the star (Schlichting 2014). Moreover, beyond the water ice line, where the temperature is low enough for water to condense into ice (see Fig. 1.7), the amount of solids increases by a factor of 2–4 (Lodders 2003). Furthermore, the proximity of the water ice line in particular could be fertile ground for the further growth of radially drifting pebbles, overcoming the dust growth barriers discussed in Section 1.2.4 (Ida and Guillot 2016; Drazkowska and Alibert 2017; Schoonenberg and Ormel 2017; Schoonenberg et al. 2018; 2019). Additionally, larger solid bodies can grow by capturing and accreting the pebbles, and this process is also thought to be more efficient if the pebbles are icy (Morbidelli et al. 2015). Once sufficiently large bodies form near or beyond the ice line, they migrate inwards, concurrently growing further by colliding and merging with each other. This typically results in a system of super-Earth-sized planets tightly packed near the inner disc edge (Terquem and Papaloizou 2007; Ogihara and Ida 2009; Cossou et al. 2014).

There are two main arguments against the migration scenario. First, as a system of planets collectively migrates inwards, it forms a resonant chain (e.g. Cresswell and Nelson 2006; Terquem and Papaloizou 2007; Ogihara and Ida 2009; Ida and Lin 2010). Most of the observed close-in super-Earths do not lie in such resonant orbits (Burke et al. 2014). The most promising mechanism to break the resonant chains (i.e., to bring the theory in agreement with the data) is that once the gaseous disc dissipates, planetary systems become dynamically unstable and go through one last phase of scattering and collisions. If this happens, the final planetary system is no longer in resonance (Terquem and Papaloizou 2007; Ida and Lin 2010; Cossou et al. 2014; Coleman and Nelson 2016). How often this happens in real systems is, however, unclear. Matching the observed properties of super-Earth systems requires 95% of the systems to become unstable, whereas in simulated systems this number ranges from 45% to 95% depending on uncertain model parameters (Izidoro et al. 2017; 2019). Moreover, it is generally found that such a phase of giant impacts strips the planets of their atmospheres (Inamdar and Schlichting 2015; Poon et al. 2020), whereas a large number of the close-in super-Earths have significant H/He envelopes (admittedly, some authors argue that super-Earths can efficiently accrete their atmospheres even with very little gas remaining in the disc; Lee and Chiang 2016). It is important to add, however, that convergent migration which leads to planets in resonance is a feature of migration inside viscous (turbulent) discs. Predictions differ in simulations of planet migration inside inviscid (laminar) discs (e.g. those that accrete primarily via magnetohydrodynamic winds or Hall-driven radial flows, Nelson 2018). In inviscid discs, systems of super-Earths do not end up in resonances in the first place (McNally et al. 2019), and so the giant impact phase need not happen.

The second argument against the migration scenario is related to the composition of the planets. If the embryos of the super-Earths form near or outside the water ice line, the super-Earths should be water-rich (Ogihara and Ida 2009; McNeil and Nelson 2010; Izidoro et al. 2019). This is the case even if planet embryos form both inside and outside the ice line, as the water-rich embryos further out migrate inwards and collide with the purely rocky ones (Izidoro et al. 2019). In contrast, the solid cores of the super-Earths are inferred to be consistent with a rock-iron, Earth-like composition, as discussed in Section 1.1. This problem for the migration scenario seems inescapable at present.

1.3.2. Formation in the inner disc

The Earth-like composition of the super-Earths implies that they form near their present orbits, or at the very least inside the water ice line. As discussed above, formation of super-Earths at short orbital distances requires a delivery of solid material to the inner disc from the outer. It has been proposed that this delivery happens in the form of pebbles that radially drift inwards due to gas drag, and whose radial drift is stopped or slowed down at short orbital distances (Hansen and Murray 2012; Boley and Ford 2013; Chatterjee and Tan 2014, see also Kretke et al. (2009)). Radially drifting pebbles are stripped of their water ice at the water ice line, but up to a half of their material is in the form of silicates and iron (Lodders 2003), which continues its radial drift inwards.

The pebble-driven formation of the super-Earths in the inner disc has been mostly explored within the so-called inside-out planet formation scenario (IOPF; Chatterjee and Tan 2014; 2015; Hu et al. 2016; 2018). In this scenario, radially-drifting pebbles halt their drift in a local gas pressure maximum present in the inner disc (as inwards of the pressure maximum radial drift switches direction; see discussion in Section 1.2.4 and Fig. 1.6). Such a local gas pressure maximum is expected to arise in the inner disc, if the disc is in steady state and if gas accretion is driven by the MRI. As discussed in Section 1.2.3, MRI-driven viscosity (and hence accretion efficiency) is expected to decrease with increasing distance from the star as the disc transitions from being hot and thermally-ionized to cold and weakly ionized. To maintain a steady state, i.e. a constant gas accretion rate, eq. (1.13) implies that gas surface density should increase outwards in this region. The viscosity reaches a minimum at the inner edge of the MRI dead zone (Fig. 1.5). At that location then, one expects a local maximum in both gas surface density and gas pressure. The location of this pressure maximum is estimated to be at an orbital distance of a few tenths of an AU (Gammie 1996; Chatterjee and Tan 2014), roughly consistent with the orbital distances of the super-Earths.

If a local gas pressure maximum can indeed trap radially drifting pebbles, enough solids could accumulate over time to form a planet. One suggestion is that pebbles accumulate at the pressure maximum until the resulting ring of pebbles becomes gravitationally unstable and collapses into a planet-sized body (Chatterjee and Tan 2014). In IOPF, once the planet forms, it is not expected to migrate inwards, as it remains trapped inside the gas surface density maximum at which it formed (Hu et al. 2016). This surface density maximum acts as a trap for the planet in the same manner as the disc inner edge (Masset et al. 2006). Instead, the planet is expected to carve out a gap in the disc (Chatterjee and Tan 2014; Hu et al. 2016). It has been hypothesized that the material interior to the gap would then accrete onto the star, leaving the material exterior to the gap exposed to stellar X-rays. Stellar X-rays would ionize the innermost region of the remnant disc, pushing the inner edge of the MRI dead zone outwards. A second planet may then form at the new local gas pressure maximum, and the process may be repeated for the third planet etc. (i.e., the planet system is formed inside-out). However, it is not clear that super-Earths formed via IOPF would indeed be immune to migration. The MRI-induced planet trap is expected to move inwards as the disc evolves, and the planet would move with it (Coleman and Nelson 2014).

Formation of the super-Earths in the inner disc has been explored in much less detail, yet it appears necessary to explain the rocky composition of these planets. In particular, the structure of an MRI-accreting inner disc has not been modelled in detail, and it is not clear if temperatures remain high enough in the inner disc for the local gas pressure maximum to persist throughout the disc lifetime. Similarly, the hypothesis that the MRI-induced pressure maximum leads to pebble accumulation remains to be tested, and the growth of pebbles to larger bodies investigated.

1.4. Thesis outline

The aim of this thesis is to investigate the structure of the innermost regions of protoplanetary discs and to explore if these regions can support formation of the close-in super-Earths. The first question to consider is if the inner edge of the MRI dead zone, and the local gas pressure maximum, indeed occur at orbital distances at which we observe the close-in super-Earths. In Chapter 2 I present theoretical models of the inner regions of viscously accreting discs in which the viscosity is self-consistently calculated from the criteria for the onset of the MRI. Then, Chapter 3 deals with the question of whether such an MRI-accreting inner disc leads to accumulation of dust. Assuming that super-Earth-sized solid planet cores do form in the MRI-accreting inner disc, in Chapter 4 I investigate accretion of gas onto these cores, i.e. the formation of planet atmospheres. The atmospheres of the super-Earths affect greatly their radii, and coupling the accreted atmospheres with a simple model of their subsequent evolution allows comparison to the observed planet properties. The models presented in Chapter 2 adopt many assumptions about the disc structure, making the problem tractable enough for a preliminary study. These assumptions are lifted in Chapter 5, which presents a detailed model of disc physical and chemical structure. Then, in Chapter 6 I investigate how this new model depends on various disc and stellar parameters. Finally, in Chapter 7 I summarise the main findings and discuss the remaining open problems.

2 MRI-ACCRETING INNER DISC

2.1. Introduction

For the super-Earths to form near their present short orbits, the inner protoplanetary disc must become enriched in solids. This may happen via the following process. Small solids (dust grains, or pebbles) radially drift through protoplanetary discs towards the star due to gas drag. Concurrently, in steady state, the inner protoplanetary disc is expected to feature a local gas pressure maximum, at the boundary between the MRI-active and the MRI-dead zones. This pressure maximum may trap the radially-drifting pebbles, and super-Earths might then form from the accumulated pebbles (e.g. Chatterjee and Tan 2014). To investigate whether the inner disc indeed features a local gas pressure maximum, I produce models of the inner disc in which the disc accretes viscously via turbulence induced by the MRI.

The boundary between the MRI-active and the MRI-dead zone in the inner disc has been previously studied using numerical MHD simulations (Dzyurkevich et al. 2010; Flock et al. 2017). These simulations can directly probe the turbulent motions of the fluid and the evolution of the magnetic field. However, they are computationally expensive; thus, they either use a fixed, prescribed temperature profile (Dzyurkevich et al. 2010), or cannot be run for long enough to reach a steady state (Flock et al. 2017). As such, they can be used to study the formation of the pressure maximum at this boundary, but they cannot predict its location in a real, steady-state disc. Additionally, the cited studies only consider the MRI-dead zone arising due to Ohmic diffusion, whereas ambipolar diffusion can also suppress the MRI.

The formation of the pressure maximum due to the MRI can also be studied using much simpler and computationally cheaper viscous disc models (e.g. Terquem 2008; Kretke and Lin 2007; Kretke et al. 2009; Kretke and Lin 2010; Chatterjee and Tan 2014). This is achieved by parametrizing the viscosity arising from the MRI as a function of disc structure parameters (e.g. temperature, density). However, previous studies either adopt simplified parametrizations that do not consider the detailed physics of gas ionization and the coupling between the gas and the magnetic field, or they make ad hoc assumptions about the magnetic field strength.

In the study presented in this chapter, viscosity is parametrized based on MRI criteria extracted from MHD simulations (Sano and Stone 2002; Turner et al. 2007; Bai and Stone 2011; Bai 2011a), with a detailed calculation of magnetic resistivities and a physically motivated choice of the magnetic field strength at every orbital radius. Additionally, both Ohmic and ambipolar diffusion are taken into account. On the other hand, I use a simplified model for the disc structure. It is assumed that the inner disc is vertically isothermal, disc opacity is a constant, and there is only a single ionized species. Both heating by stellar irradiation and ionization of the disc by stellar X-rays are neglected. These are considerable simplifications, which will be re-considered in Chapter 5. However, the models presented here still capture the basic physics expected to set the structure of the inner disc, and the simplicity allows clear insights into the MRI-driven accretion in the inner disc and the inner disc structure.

2.2. Methods

In Section 2.2.1 I discuss the adopted model of the disc structure, Section 2.2.2 explains how a viscosity driven by the MRI is calculated, and Section 2.2.3 how the disc structure and the MRI viscosity are coupled to obtain a self-consistent model.

2.2.1. Standard α-disc model

To model the disc structure, I employ the standard Shakura-Sunyaev α-disc model. This is a model of a thin steady-state viscously accreting disc, with the viscosity parametrized using the Shakura and Sunyaev (1973) α parameter. The basics of thin viscous discs were covered in the previous Chapter. In this model it is further assumed that the heat released through viscous dissipation is transported vertically by radiation and that the disc is optically thick, i.e. that the optical depth at disc midplane $\tau_{\mathrm{mid}} = \frac{1}{2}\Sigma\kappa \gg 1$. Here, κ is the Rosseland-mean opacity, assumed to be vertically constant. Then, the disc midplane temperature is given by (e.g. Hubeny 1990)

$$\sigma T_{\mathrm{mid}}^4 = \frac{3}{8}\tau_{\mathrm{mid}} F_{\mathrm{acc}}, \tag{2.1}$$

where F_{acc} is the total heat flux leaving through one side of the disc, given by eq. (1.15).

Equation (2.1) implies that the temperature varies with optical depth, and thus varies with height above disc midplane. However, the dependence is weak, $T \propto \tau^{-1/4}$. Hence, in this model it is assumed that the disc is vertically isothermal with the temperature at all heights equal to the temperature at the midplane. The solution to the hydrostatic equilibrium, eq. (1.2), is then given by

$$P(r,z) = P_{\mathrm{mid}}(r) e^{-\frac{z^2}{H^2}}, \tag{2.2}$$

where P_{mid} is the pressure at midplane and $H = \sqrt{2}c_{\mathrm{s}}/\Omega$ is the disc pressure scale height. Adopting the ideal gas law, an analogous expression describes the vertical density (ρ) profile, with

$$\rho_{\mathrm{mid}} = \frac{\Sigma}{\sqrt{\pi}H}. \tag{2.3}$$

Gathering expressions for the midplane density and temperature, eqns. (2.3) and (2.1), total viscous dissipation, eq. (1.15), the relation between the disc surface density, vertically-averaged viscosity and accretion rate, eq. (1.13), and the parametrization of viscosity, eq. (1.16), one obtains (e.g. Hu et al. 2016)

$$\Sigma = \frac{2^{7/5}}{3^{6/5}\pi^{3/5}}(\frac{\mu m_p}{k_B})^{4/5} G^{1/5} \sigma_{SB}^{1/5} \kappa^{-1/5} \bar{\alpha}^{-4/5} M_*^{1/5} (f_r \dot{M})^{3/5} r^{-3/5}, \tag{2.4}$$

$$\rho_{\rm mid} = \frac{2^{8/5}}{3^{13/10}\pi^{9/10}}(\frac{\mu m_p}{k_B})^{6/5} G^{11/20} \sigma_{SB}^{3/10} \kappa^{-3/10} \bar{\alpha}^{-7/10} M_*^{11/20} (f_r \dot{M})^{2/5} r^{-33/20}, \tag{2.5}$$

$$T_{\rm mid} = \frac{3^{1/5}}{2^{7/5}\pi^{2/5}}(\frac{\mu m_p}{k_B})^{1/5} G^{3/10} \sigma_{SB}^{-1/5} \kappa^{1/5} \bar{\alpha}^{-1/5} M_*^{3/10} (f_r \dot{M})^{2/5} r^{-9/10}, \tag{2.6}$$

$$P_{\rm mid} = \frac{2^{1/5}}{3^{11/10}\pi^{13/10}}(\frac{\mu m_p}{k_B})^{2/5} G^{17/20} \sigma_{SB}^{1/10} \kappa^{-1/10} \bar{\alpha}^{-9/10} M_*^{17/20} (f_r \dot{M})^{4/5} r^{-51/20}, \tag{2.7}$$

where $f_r = 1-\sqrt{\frac{R_*}{r}}$. This series of expressions show simple relationships between the quantities describing the disc structure and stellar parameters, disc accretion rate $\dot{M}$, opacity κ and vertically-averaged viscosity parameter $\bar{\alpha}$. The latter is in general a pressure-weighted average, and in the case of a vertically-isothermal disc equal to a density-weighted average,

$$\bar{\alpha} \equiv \frac{\int_{-\infty}^{\infty} \alpha P dz}{\int_{-\infty}^{\infty} P dz} = \frac{\int_{-\infty}^{\infty} \alpha \rho dz}{\int_{-\infty}^{\infty} \rho dz}. \tag{2.8}$$

In principle, both the opacity κ and viscosity parameter $\bar{\alpha}$ are functions of the disc structure. For simplicity, in this work it is assumed that the opacity is constant everywhere in the disc and equals $\kappa = 10\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, consistent with typical properties of small dust grains in protoplanetary discs (e.g. Wood et al. 2002). The dependence of the viscosity parameter $\bar{\alpha}$ on the disc structure is, on the other hand, the key ingredient of the model presented here and is discussed in the next section.

2.2.2. MRI-driven viscosity parameter α

In the following, I discuss the criteria for the onset of the MRI, calculation of the ionization state of the gas, calculation of the coupling between the gas and the magnetic field (i.e., the resistivities), and finally the calculation of the viscosity parameter α due to the MRI.

2.2.2.1. Criteria

The viscosity that the MRI induces depends, first of all, on whether a certain region of the disc is unstable to the MRI. As discussed in Chapter 1, there are three effects that can modify the onset of the MRI: Ohmic diffusion, ambipolar diffusion and the Hall effect. Ohmic diffusion can suppress the MRI entirely by decoupling the magnetic field from the gas. Numerical simulations (Sano and Stone 2002; Turner et al. 2007) show that Ohmic diffusion does *not* suppress the

MRI if

$$\Lambda \equiv \frac{v_{\mathcal{A}z}^2}{\eta_{\rm O}\Omega} > 1, \tag{2.9}$$

where Λ is the Ohmic Elsasser number, η_O is the Ohmic resistivity and $v_{\mathcal{A}z}$ the vertical component of the local Alfven velocity ($v_{\mathcal{A}z} = B_z/\sqrt{4\pi\rho}$, where B_z is the vertical component of the magnetic field strength and ρ is the local gas density). This criterion expresses the requirement that for the MRI to be active, the MRI growth rate must be larger than the Ohmic diffusion rate (Fleming et al. 2000). The growth rate of the fastest-growing MRI mode is $\sim\Omega$ and the wavenumber of that mode is $k \sim \Omega/v_{\mathcal{A}z}$. From eq. (1.18), one can also see that for a given wavenumber k, the Ohmic diffusion rate is $\sim\eta_O k^2$. Then, the comparison of the two rates yields eq. (2.9). Note that the dependence on the vertical component of the Alfven velocity $v_{\mathcal{A}z}$ (and the magnetic field strength B_z), rather than the total $v_{\mathcal{A}}$ (and the total magnetic field strength B), comes from the notion that the Ohmic diffusion affects smaller lengthscale perturbations most easily, and in a thin disc those are in the vertical direction (Turner et al. 2007).

Ambipolar diffusion may also suppress the MRI. A criterion analogous to the above has been investigated (Hawley and Stone 1998) using the ambipolar Elsasser number

$$Am \equiv \frac{v_{\mathcal{A}}^2}{\eta_{\rm A}\Omega}, \tag{2.10}$$

where $\eta_{\rm A}$ is the ambipolar resistivity. However, Bai and Stone (2011) showed that when ions and neutrals in the gas are strongly coupled, behaving as a single fluid under the influence of the magnetic field, ambipolar diffusion will not suppress the MRI even if $Am < 1$, if the magnetic pressure is sufficiently weak compared to the gas pressure. That is, for a plasma β-parameter, $\beta \equiv P/P_B$, where $P_B = B^2/8\pi$ is the magnetic field pressure, the MRI will *not* be suppressed if

$$\beta > \beta_{\rm min}, \tag{2.11}$$

where the minimum value $\beta_{\rm min}$ is given by

$$\beta_{\rm min}(Am) = \left[\left(\frac{50}{Am^{1.2}}\right)^2 + \left(\frac{8}{Am^{0.3}} + 1\right)^2\right]^{1/2}. \tag{2.12}$$

For the ions and neutrals to be strongly coupled it is required that the density of ions be much lower than the density of neutrals, and that the recombination timescale $t_{\rm rcb}$ for ions be much shorter than the dynamical timescale, i.e. $t_{\rm rcb} \ll t_{\rm dyn} \sim 1/\Omega$. I assume that both of these conditions are fulfilled, and *a posteriori* check if this is true. Furthermore, note that while the Ohmic Elsasser number depends on the vertical component of the magnetic field strength B_z, the ambipolar Elsasser number depends on the total strength B. The two are not independent, however, and numerical simulations show that $B_z^2 \sim B^2/25$ Sano et al. (2004).

The Hall effect can also significantly modify the behaviour of the gas and the magnetic field (Lesur et al. 2014). However, its effects are non-diffusive, non-trivial and also dependent on the alignment between the rotation axis of the disc and the magnetic field threading the disc. Therefore, it is neglected in this study, and its importance for our results will be only *a posteriori* investigated by considering the Hall Elsasser number,

$$\chi = \frac{v_{\mathcal{A}}^2}{|\eta_{\mathrm{H}}|\Omega}. \tag{2.13}$$

In summary, if conditions given by eq. (2.9) and (2.11) are both fulfilled at some location in the disc, it is assumed that the MRI is active there. A region where the Ohmic criterion, eq. (2.9), is not fulfilled is called a dead zone (Gammie 1996). A region where the ambipolar criterion, eq. (2.11), is not fulfilled will be referred to as a zombie zone (following Mohanty et al. 2013). Note that a zombie zone is qualitatively the same as a dead zone, as in both regions the MRI is fully suppressed.

2.2.2.2. Ionization

For simplicity, in this chapter I only consider thermal (collisional) ionization. The inner edge of the dead zone is expected to occur at the orbital radius at which the thermal ionization becomes inefficient, due to the temperature decreasing with distance from the star. The temperature below which thermal ionization becomes inefficient is determined by the ionization potential of a given chemical species. Therefore, the location of the inner edge of the dead zone is determined by the lowest ionization potential among the species that are abundant in the inner disc. In terms of the lowest ionization potential, the top three species overall are rubidium, caesium and potassium; however, the abundance of potassium is orders of magnitude higher than that of rubidium or caesium (Lodders 2003). Sodium, which has the fourth lowest ionization potential, is more abundant than potassium. However, it becomes efficiently ionized at temperatures significantly higher than the one needed for potassium, and so sodium ions only become important radially closer to the star. Therefore, in this chapter, I consider thermal ionization of potassium only.

If both ionization and recombination are driven by gas-phase collisions only, the equilibrium number densities of free electrons (n_{e}), atomic ions (n_{i}) and neutral atoms (n_0) are determined by the Saha equation,

$$\frac{n_{\mathrm{e}}\, n_{\mathrm{i}}}{n_0} = \frac{1}{\lambda_{\mathrm{e}}^3} \frac{g_{\mathrm{e}}\, g_{\mathrm{i}}}{g_0} \exp\left(\frac{-\mathcal{I}}{k_{\mathrm{B}}T}\right), \tag{2.14}$$

where $\lambda_{\mathrm{e}} = \sqrt{h^2/(2\pi m_{\mathrm{e}} k_{\mathrm{B}} T)}$ is the thermal de Broglie wavelength of electrons of mass m_{e}; $g_{\mathrm{e}} = 2$, g_0 and g_{i} are the degeneracy of states for free electrons, neutrals and ions; and $\mathcal{I}$ is the ionization potential. For potassium $\mathcal{I} = 4.34\,\mathrm{eV}$, and it is assumed that g_{i}/g_0 is the same as for sodium, $g_{\mathrm{i}}/g_0 = 1/2$ (Rouse 1961). Since it is assumed that potassium is the only ionized species, charge conservation requires that $n_\mathrm{e} = n_\mathrm{i}$. Then, it also follows that $n_0 = n_\mathrm{K,tot} - n_\mathrm{e}$, where $n_\mathrm{K,tot}$ is the total number density of potassium ions and neutrals ($n_\mathrm{K,tot} = 1.97\times10^{-7} n_\mathrm{H_2}$, following Keith and Wardle 2014).

Finally, since it is assumed that free electrons can only recombine in collisions with potassium ions, the recombination rate is given by $dn_\mathrm{e}/dt = k_\mathrm{ei} n_\mathrm{e} n_\mathrm{i}$, where $k_\mathrm{ei} = 3\times10^{-11}/\sqrt{T}\,\mathrm{cm^3\,s^{-1}}$ (Ilgner and Nelson 2006). The recombination timescale is then $t_\mathrm{rcb} = n_\mathrm{e}/(dn_\mathrm{e}/dt) = 1/(k_\mathrm{ei} n_\mathrm{e})$.

2.2.2.3. Resistivities

In this work, The Ohmic, Hall and ambipolar resistivities are calculated following Wardle (2007), first expressing them as

$$\eta_\mathrm{O} = \frac{c^2}{4\pi\sigma_\mathrm{O}}, \tag{2.15}$$

$$\eta_\mathrm{H} = \frac{c^2}{4\pi\sigma_\perp}\frac{\sigma_\mathrm{H}}{\sigma_\perp}, \tag{2.16}$$

$$\eta_\mathrm{A} = \frac{c^2}{4\pi\sigma_\perp}\frac{\sigma_\mathrm{P}}{\sigma_\perp} - \eta_\mathrm{O}. \tag{2.17}$$

where σ_O, σ_H and σ_P are the Ohmic, Hall and Pederson conductivities, respectively, and $\sigma_\perp = \sqrt{\sigma_\mathrm{H}^2 + \sigma_\mathrm{P}^2}$ is the total conductivity perpendicular to the magnetic field. The conductivites are given by

$$\sigma_\mathrm{O} = \frac{ec}{B}\sum_j n_j\,|Z_j|\,\beta_j \tag{2.18}$$

$$\sigma_\mathrm{H} = \frac{ec}{B}\sum_j \frac{n_j\,Z_j}{1+\beta_j^2} \tag{2.19}$$

$$\sigma_\mathrm{P} = \frac{ec}{B}\sum_j \frac{n_j\,|Z_j|\,\beta_j}{1+\beta_j^2} \tag{2.20}$$

where the sums are over all charged species j (here, these are free electrons and potassium ions), m_j is particle mass, n_j the number density, and $Z_j e$ the charge of each species. Further, β_j is the Hall parameter, the ratio of the gyrofrequency of a particle to its collision frequency with neutral particles. It is given by

$$\beta_j = \frac{|Z_j|eB}{m_j\,c}\frac{1}{\gamma_j\rho_\mathrm{n}}, \tag{2.21}$$

where γ_j is the drag coefficient and ρ_n is the mass density of neutral particles (and $\gamma_j\rho_\mathrm{n}$ the collision frequency with neutrals). The drag coefficient γ_j is calculated as

$$\gamma_j = \frac{\langle\sigma v\rangle_j}{m_j + m_\mathrm{n}}, \tag{2.22}$$

where $m_{\mathrm{n}} = \mu m_{\mathrm{H}}$ is the mean neutral particle mass, and $\langle \sigma v \rangle_j$ the rate coefficient for momentum transfer in collisions between charged species j and neutrals. The latter are given by Wardle (1999),

$$\langle \sigma v \rangle_{\mathrm{e}} = 10^{-15} \left(\frac{128\, k_{\mathrm{B}} T}{9 \pi m_{\mathrm{e}}} \right)^{1/2} \ \mathrm{cm}^3\,\mathrm{s}^{-1}, \tag{2.23}$$

$$\langle \sigma v \rangle_i = 1.6 \times 10^{-9} \ \mathrm{cm}^3\,\mathrm{s}^{-1}, \tag{2.24}$$

for electrons and ions, respectively.

2.2.2.4. *Viscosity*

If criteria (2.9) and (2.11) are met, the MRI induces turbulence. Turbulence produces a time- and volume-averaged shear stress, measured in numerical simulations (Sano et al. 2004) to be

$$\ll W_{r\phi} \gg = -\frac{1}{2} \ll P_B \gg . \tag{2.25}$$

In the framework of viscous accretion discs, at any location in the disc the shear stress is given by eq. (1.4),

$$W_{r\phi} = -\frac{3}{2} \alpha P, \tag{2.26}$$

where I have substituted viscosity in terms of the Shakura-Sunyaev α parameter using eq. (1.16). Therefore, the MRI-driven α viscosity is

$$\alpha = \frac{1}{3} \frac{P_B}{P} = \frac{1}{3\beta}. \tag{2.27}$$

The time and volume average symbols have been dropped here for simplicity, and it is implicitly assumed that at any location in the disc α, P_B and P are averages over some finite volume and time.

Furthermore, numerical simulations show that when the MRI is suppressed by Ohmic diffusion (i.e., in the dead zone, where criterion (2.9) is not met), there is a non-zero shear stress induced by the adjacent MRI-active zone. The resulting viscosity parameter in the dead zone is of the order of $\alpha_{\mathrm{DZ}} \sim 10^{-4}$ (e.g. Dzyurkevich et al. 2010). Purely hydrodynamic instabilities may also produce a weak stress in the dead zone (e.g. Lesur and Papaloizou 2010; Nelson et al. 2013; Stoll and Kley 2014). In this work I assume that the viscosity parameter is constant in the dead zone, and explore a range of values ($\alpha_{\mathrm{DZ}} = 10^{-5} - 10^{-3}$). For simplicity, the same value is adopted for the viscosity parameter in the zombie zone (where the MRI is suppressed by ambipolar diffusion). Additionally, in this model α_{DZ} also denotes a minimum value of α. That is, even if the Ohmic and ambipolar criteria indicate that the MRI should be active, if eq. (2.27) yields a viscosity parameter lower than α_{DZ}, it is assumed that $\alpha = \alpha_{\mathrm{DZ}}$.

The criteria for the MRI and the value of the MRI-driven viscosity in the active zone all depend on the magnetic field strength. The Ohmic and ambipolar criterion are functions of the vertical component of the magnetic field strength B_z and the total strength B, respectively, and the value of α in the active zone is also a function of the total strength B. As already noted above, simulations show that $B_z^2 \sim B^2/25$ (Sano et al. 2004), and so one only needs to consider either B_z or B. Furthermore, following Bai (2011a), I assume that the magnetic field strength is vertically constant. Then, at any orbital radius, one only needs to consider a single value of B. In this model, this value is chosen at each radius such that maximizes the accretion efficiency (viscosity parameter). How exactly this is done is described further below; here I only discuss the motivation behind this choice. First note that, as discussed in Chapter 1, an arbitrarily weak magnetic field is required to start the instability. The induced magnetohydrodynamic turbulence amplifies the magnetic field strength. This cannot go on infinitely, as there is a maximum B at which the MRI may be active without being quenched by ambipolar diffusion. Choosing this maximum B would, by definition, yield no active MRI. Instead, it is assumed that the MRI amplifies the field until a point at which further increase in the magnetic field strength implies a lower accretion efficiency given the above MRI criteria, i.e., it is assumed that B is a monotonously increasing function of the accretion efficiency. Similar arguments were employed by Bai (2011a) and Mohanty et al. (2013).

Finally, in the standard α-disc model discussed above, at any given orbital radius the disc structure depends on the vertically-averaged viscosity parameter $\bar{\alpha}$, defined in eq. (2.8). The MRI-driven α varies with height above disc midplane. At a given value of the magnetic field strength, the Ohmic and ambipolar criteria depend on gas pressure, density and resistivities (which are also functions of the ionization levels). In general, for a given disc structure (temperature, pressure and density) one could evaluate the ionization levels, resistivities and the MRI criteria at a number of points above disc midplane. Then, in the MRI-active zones α follows from eq. (2.27), and in the MRI-dead and zombie zones $\alpha = \alpha_{\rm DZ}$. A pressure-weighted (or a density-weighted) average $\bar{\alpha}$ could then be evaluated numerically. However, the obtained structure of the different MRI zones allows for some simplifications. It is found that there is, generally, a dead zone around disc midplane, a zombie zone in the disc upper layers, and an active zone in between. Therefore, to calculate $\bar{\alpha}$ from a given disc structure, first I use a root-finding algorithm (specifically, the bisection method) to find the height of the dead zone, i.e., the height at which the Ohmic criterion eq. (2.9) is fulfilled ($h_{\rm DZ}$, which may be zero). The same is repeated to find the height of the boundary between the active and the zombie zone ($h_{\rm ZZ}$). The vertically-averaged MRI-driven $\bar{\alpha}$ is then given by

$$\bar{\alpha} = \frac{\int_0^{h_{\rm DZ}} \alpha_{\rm DZ} P dz + \int_{h_{\rm DZ}}^{h_{\rm ZZ}} \frac{1}{3} P_B dz + \int_{h_{\rm ZZ}}^{\infty} \alpha_{\rm DZ} P dz}{\int_0^{\infty} P dz}, \tag{2.28}$$

where I use eq. (2.27) for the value of α in the active zone. Since P_B is vertically-constant, the second term in the numerator is calculated trivially. Furthermore, since it is assumed that the disc is vertically isothermal, pressure as a function of height is given by eq. (2.2) and so the rest of the integrals are calculated using the well-known Error function.

2.2.3. Self-consistent α-disc model

The key parameters of the model presented here are the disc accretion rate $\dot{M}$, stellar mass M_*, stellar radius R_* and the value of viscosity parameter in the dead zone α_{DZ}. If these parameters are specified, a self-consistent solution for the disc structure can be obtained as follows. At a given orbital radius and an input vertically-averaged viscosity parameter $\bar{\alpha}_{\mathrm{in}}$, equations (2.4, 2.2), along with the ideal gas law, yield the temperature, pressure and density at an arbitrary height above disc midplane. Then, as described in Section 2.2.2, this inferred disc structure can be used to calculate the MRI-driven $\bar{\alpha}_{\mathrm{out}}$. At a fixed value of the magnetic field strength B, I use a root-finding method (bisection) to find the self-consistent value $\bar{\alpha}_{\mathrm{in}} = \bar{\alpha}_{\mathrm{out}}$.

At a fixed value of B, the solution appears to always be unique. However, there is an infinite number of solutions with varying B. At low enough and high enough B the only solution is the one where the MRI is completely suppressed, $\bar{\alpha}_{\mathrm{in}} = \bar{\alpha}_{\mathrm{out}} = \alpha_{\mathrm{DZ}}$. In the innermost disc, i.e., at short orbital radii, there is a range of B for which the MRI is active, and $\bar{\alpha}_{\mathrm{in}} = \bar{\alpha}_{\mathrm{out}} = \bar{\alpha}$ peaks at some value of B. This determines our chosen solution. At larger orbital radii, the MRI is suppressed at any value of B, as expected in the absence of non-thermal sources of ionization; there $\bar{\alpha} = \alpha_{\mathrm{DZ}}$. At any radius, a scan over B is performed in the range $10^{-5} - 10^{3}\,\mathrm{G}$. If there is a range of values for which an MRI-active zone exists, $\bar{\alpha}$ is maximized as a function of B within that range.

The numerical procedures described in this section have been implemented as an extension of the code originally developed by S. Mohanty.

2.3. Results

Section 2.3.1 presents an in-depth analysis of the ionization, MRI-driven viscosity and the structure of the disc, for a fiducial model with a gas accretion rate $\dot{M} = 10^{-9}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, stellar mass $M_* = 1\,\mathrm{M}_{\odot}$ and a dead-zone viscosity parameter $\alpha_{\mathrm{DZ}} = 10^{-4}$. I also analyse the stability of the obtained steady-state disc structure to surface density perturbations, and consider the validity of the strong-coupling assumption. In Section 2.3.2, I consider the effects of varying the above three model parameters. For a Solar mass star (the fiducial choice), I assume the stellar radius is $R_* = 2.33\,\mathrm{R}_{\odot}$, based on the stellar models of Baraffe et al. (1998) and a stellar age of 1 Myr. In all models the inner disc boundary is at the stellar radius and the outer boundary is at the orbital radius outwards from which the MRI is completely suppressed.

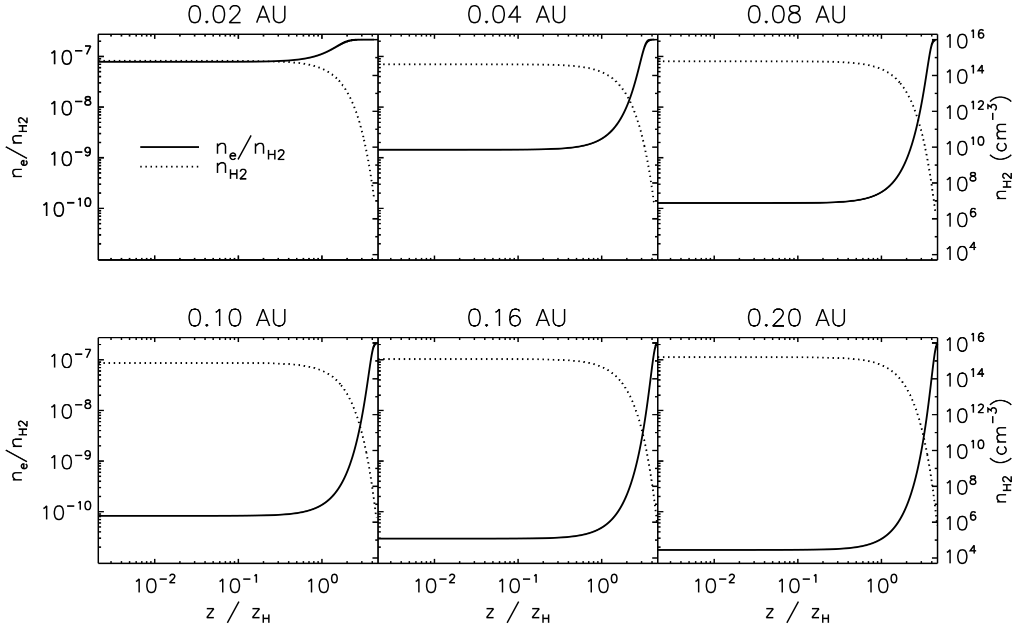


Figure 2.1: Fractional ionization ($n_\mathrm{e}/n_{\mathrm{H}_2}$) and molecular hydrogen density (n_{H_2}) as a function of height above disc midplane (in units of disc scale height z_H at each radius), at different orbital radii for the fiducial disc model.

2.3.1. Fiducial model

2.3.1.1. Ionization and resistivities

Molecular hydrogen is by far the most abundant species in discs, and so the number density of neutral particles and the total number density of particles are approximately equal to that of molecular hydrogen, $n_\mathrm{n} \approx n_\mathrm{tot} \approx n_{\mathrm{H}_2}$. Therefore, when considering the ionization levels in the disc, the relevant quantity is the ratio $n_\mathrm{e}/n_{\mathrm{H}_2}$ (where $n_\mathrm{e} = n_+$, since potassium is the only ionized species). Fig. 2.1 shows $n_\mathrm{e}/n_{\mathrm{H}_2}$ for the fiducial disc model as a function of height above disc midplane (z), at different orbital radii. For reference, also shown is the number density of molecular hydrogen (n_{H_2}). Two trends emerge. First, as expected, the fractional ionization near the disc midplane (at small heights) decreases with orbital radius (this is further discussed below). Second, at all radii $n_\mathrm{e}/n_{\mathrm{H}_2}$ increases with height above the midplane. The latter follows simply from the solution to the Saha equation: since the temperature is vertically constant, and density decreases with height, the fraction $n_\mathrm{e}/n_{\mathrm{H}_2}$ increases with height. As $n_{\mathrm{H}_2} \to 0$ towards disc surface, $n_\mathrm{e} \to n_\mathrm{K}$.

Fig. 2.2 shows the resulting ambipolar, Ohmic and Hall resistivities (η_A, η_O and $|\eta_\mathrm{H}|$, respectively). The Ohmic resitivity is inversely proportional to the fractional ionization and directly proportional to the density of neutrals (through its dependence on the Hall parameter of each charged species). As a result, the Ohmic resitivity decreases with height above disc midplane at all radii. The reverse is true for the ambipolar and Hall resistivities, which both increase with the decreasing gas density and thus with height. To compare the three resistivities, Fig. 2.3

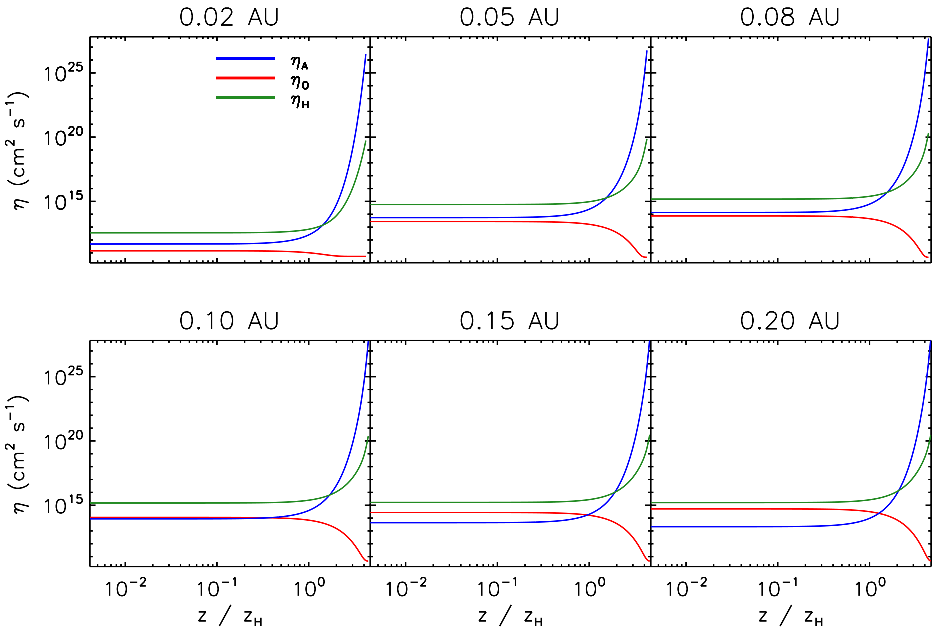


Figure 2.2: Ambipolar (η_{A}), Ohmic (η_{O}) and Hall ($|\eta_{\mathrm{H}}|$) resistivities as a function of height above disc midplane (in units of disc scale height z_H at each radius), at different orbital radii for the fiducial disc model.

shows in different colours different regions of the disc according to the relative ratios between the three resistivities. High in the disc atmosphere ambipolar resistivity dominates, and lower below, in the bulk of the disc, Hall resistivity. For this fiducial model, Ohmic resistivity is not dominant at any location in the disc.

2.3.1.2. *Active, dead and zombie zones*

Fig. 2.4 shows the regions in the disc where the MRI is active, the dead zone (where the MRI is suppressed due to Ohmic diffusion, and the Ohmic Elsasser number $\Lambda < 1$), the zombie zone (where the MRI is suppressed due to ambipolar diffusion, and the plasma parameter $\beta < \beta_{\mathrm{min}}$). Also shown is the Hall zone (where Hall Elsasser number $\chi < 1$), where the Hall effect can be expected to alter the MRI and the disc structure. This is, however, neglected, and the Hall zone is shown only as an indication of where the results presented here might be inapplicable (see discussion at the end of this section). Furthermore, the midplane values of $n_{\mathrm{e}}/n_{\mathrm{H_2}}$, β and Λ are shown as functions of radius in Fig. 2.5, and Fig. 2.6 shows the magnetic field strength B.

From the inner disc edge to the outer edge where the MRI becomes completely suppressed, three distinct regions appear as a result of the competing influences of Ohmic and ambipolar diffusion. From the inner disc edge to about 0.05 AU, the height of the active zone is limited by the ambipolar diffusion only. Due to the high ionization levels near disc midplane, as well as low densities (discussed further below), Ohmic resistivity is low in this region, and $\Lambda \gg 1$. To

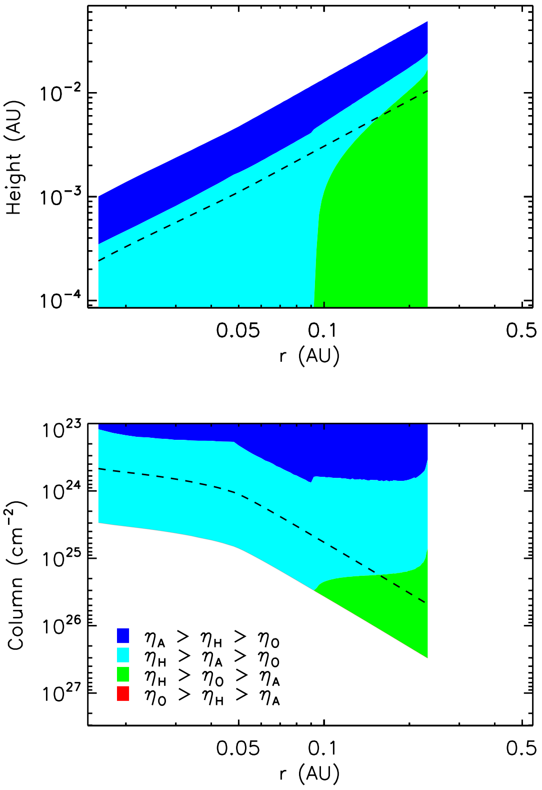


Figure 2.3: Relative importance of the Ohmic ($\eta_{\rm O}$), Hall ($|\eta_{\rm H}|$) and ambipolar ($\eta_{\rm A}$) resistivities as a function of location in the disc, for the fiducial disc model. The top panel shows vertical location in units of height above disc midplane; the bottom panel shows vertical location in units of column density measured from the disc surface. The dashed line in both panels indicates one pressure scale height.

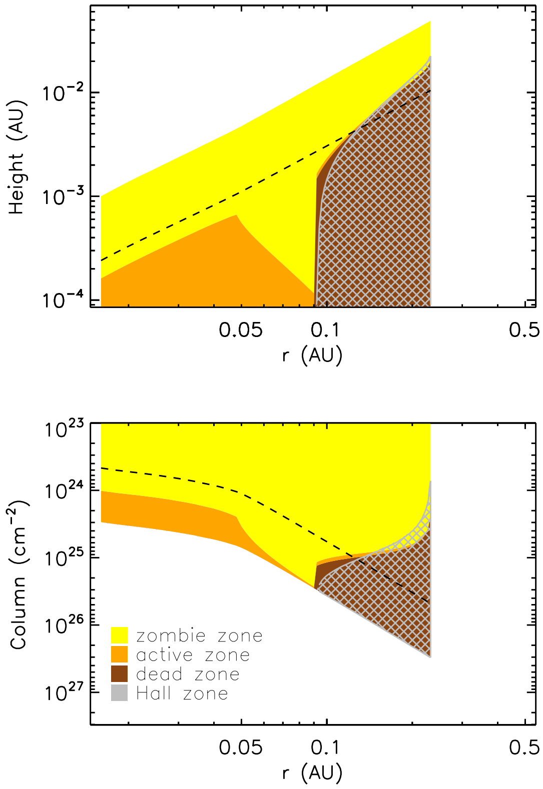


Figure 2.4: The MRI-active ($\Lambda > 1$ and $\beta > \beta_{\rm min}$), dead ($\Lambda < 1$), zombie ($\beta < \beta_{\rm min}$) and Hall ($\chi < 1$) zones as a function of disk location. The top panel shows vertical location in units of height above disc midplane; the bottom panel shows vertical location in units of column density measured from the disc surface. The dashed line in both panels indicates one pressure scale height. Note that the active zone exists at all orbital radii shown here, although beyond 0.09 AU it is a thin layer sandwiched between the dead and zombie zones

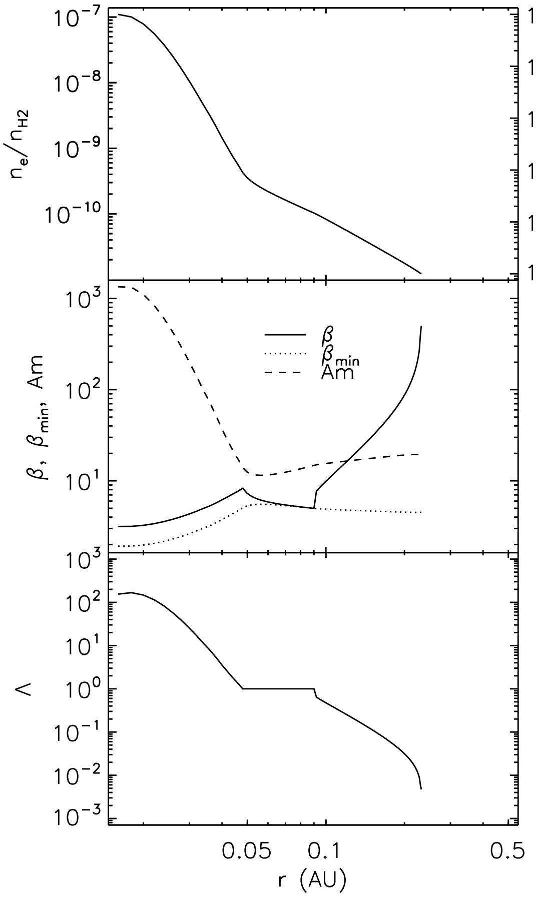


Figure 2.5: Top: Fractional abundance of free electrons as a function of radius, expressed relative to both the number density of hydrogen molecules (n_e/n_{H_2}; left axis) and number density of potassium (n_e/n_K; right axis). Middle: Plasma β parameter, minimum value β_{min} required for active MRI, and ambipolar Elsasser number Am. Bottom: Ohmic Elsasser number Λ.

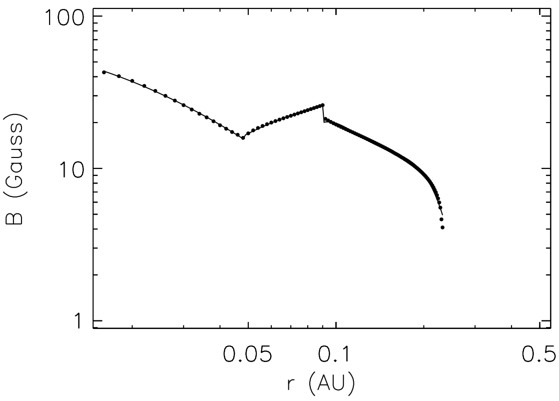


Figure 2.6: Magnetic field strength B as a function of radius, for the fiducial disc model. Filled circles are the model results, and the solid line is a combined polynomial fit to these results.

maintain the highest accretion efficiency at each radius against ambipolar diffusion, magnetic field strength decreases with radius.

From 0.05 AU to about 0.09 AU, $\Lambda = 1$. This region still features no dead zone. However, Ohmic diffusion indirectly influences the disc structure. The most efficiently accreting state in this region is that in which the magnetic field strength B increases with radius, to keep midplane Λ at unity. On the other hand, this results in an increased suppression of the MRI by ambipolar diffusion, and so the active zone diminishes in this region.

At about 0.09 AU, the solution changes qualitatively. The MRI-active zone is here a thin layer between the dead zone around midplane, and the zombie zone in the disc upper layers. As the ionization levels and the magnetic field strength drop even further, this layer becomes thinner and thinner, and at about 0.25 AU the MRI is completely suppressed.

The Hall zone mostly overlaps with the dead zone in Fig. 2.4, and also covers parts of the thin active layer. The Hall effect on accretion is dependent on the alignment between the magnetic field threading the disc and the rotation axis of the disc. It could either act to completely suppress the MRI (moving the outer boundary of the model shown here inwards), or to drive laminar accretion near disc midplane. In the latter case, the influence on the disc structure would depend on the ratio of the accretion efficiencies in the MRI-active and the Hall zone. The fact that the Hall zone overlaps with a significant part of (although not with the entire) inner disc indicates importance of including it in future studies. However, as the accretion is laminar in the Hall-dominated regions, it cannot be trivially accounted for in the viscous disc framework.

2.3.1.3. *Viscosity, disc structure and pressure maximum*

The vertically-averaged viscosity parameter $\bar{\alpha}$ that results from the discussed structure of active, dead and zombie zones is shown in Fig. 2.7 as a function of radius. At the inner disc edge, where nearly all of the potassium is ionized, $\bar{\alpha} \sim 0.08$. Radially outwards $\bar{\alpha}$ decreases until at about

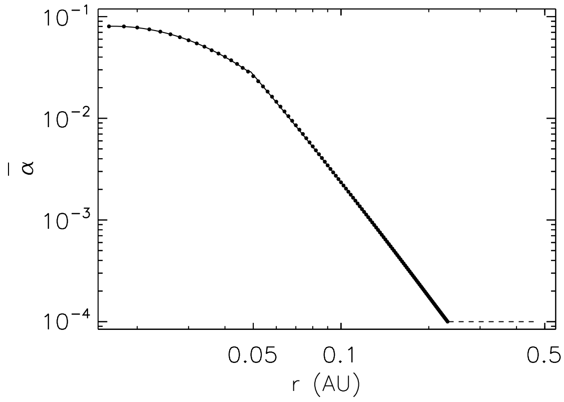


Figure 2.7: Vertically-averaged viscosity parameter $\bar{\alpha}$ as a function of radius for the fiducial disc model. Filled circles represent the model results; the solid line is a combined polynomial fit to these results. The outer boundary of the model is at ~ 0.25 AU, the radius at which $\bar{\alpha}$ decreases to the minimum, dead-zone value $\alpha_{\rm DZ} = 10^{-4}$. Beyond this radius it is assumed that $\bar{\alpha} = \alpha_{\rm DZ}$ (dashed horizontal line).

0.25 AU it reaches the minimum value $\bar{\alpha} = \alpha_{\rm DZ}$. This is the location at which the thin MRI-active layer completely disappears in Fig. 2.4. Outwards from this location it is simply assumed that $\bar{\alpha}$ remains at the constant dead-zone value.

Fig. 2.8 shows the resulting disc temperature, midplane density and pressure, and the surface density, as functions of radius. Most importantly, in line with expectations, the inner disc features a local gas pressure maximum. The pressure maximum is located at ~0.25 AU, where $\bar{\alpha}$ falls to the minimum value. Interestingly, this corresponds to the location where the MRI becomes suppressed at all heights, but not to the inner edge of the dead zone (which occurs where the MRI is suppressed at midplane but still survives higher up), which appears at about 0.09 AU (see Fig. 2.4). Therefore, this potential pebble trap lies in the region where the disc midplane is only weakly turbulent. The location of the pressure maximum is also the location of the overall maximum in the surface density. Inwards of this location, the MRI-driven accretion efficiency (i.e., $\bar{\alpha}$) increases; at a constant gas accretion rate, this leads to a decrease in the surface density. Finally, note that inwards of the pressure maximum, the temperature profile is nearly flat. The temperature at the location of the pressure maximum is ~1000 K, as expected from the ionization potential of potassium. Outwards from the pressure maximum, temperature decreases with radius much more steeply.

2.3.1.4. Viscous instability

The models considered here are steady-state models, meaning that all quantities are assumed to be constant in time, with a gas accretion rate that is also constant at all radii. A viscous disc that is in steady-state, however, can still be unstable to surface density perturbations (viscously unstable, Lightman and Eardley 1974; Pringle 1981). Consider the time evolution equation for

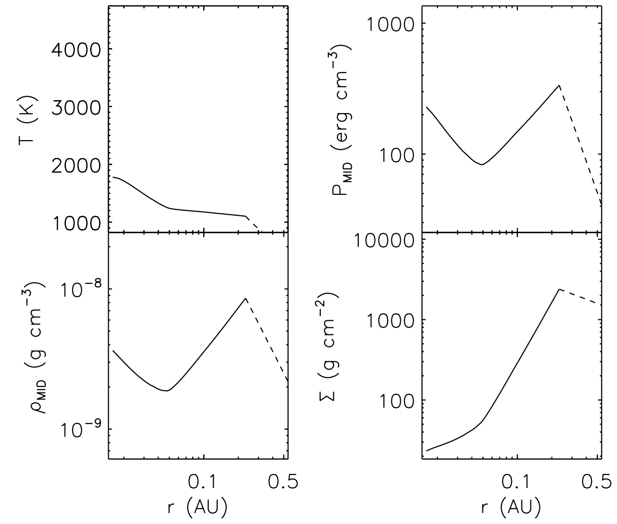


Figure 2.8: Various disc structure quantities as a function of radius, for the fiducial disc model. Top left: (Vertically constant) temperature. Top right and bottom left: Midplane pressure and midplane density. Bottom right: Surface density. Solid lines show the results of the self-consistent MRI-accreting model, ending at the radius where $\bar{\alpha}$ falls to $\alpha_{\rm DZ}$. Beyond this radius, dashed lines show the results based on an assumption that $\bar{\alpha} = \alpha_{\rm DZ}$.

the surface density, eq. (1.11). In steady-state, $\partial\Sigma/\partial t = 0$, corresponding to a solution $\Sigma = \Sigma_0$. Then, consider a small perturbation, $\Sigma_0 \rightarrow \Sigma_0 + \delta\Sigma$, leading to a perturbation in $x \equiv \bar{\nu}\Sigma$, $x_0 \rightarrow x_0 + \delta x$, where $\delta x = (\partial x/\partial\Sigma)\delta\Sigma$. Here, note that in a thin disc, both the hydrostatic and thermal equilibria are established on timescales much shorter than the viscous timescale, so on a viscous timescale $\bar{\nu}$ can be considered a function of Σ only (at a given orbital radius). Substituting $\delta\Sigma$, δx into eq. (1.11) yields

$$\frac{\partial(\delta x)}{\partial t} = \left(\frac{\partial x}{\partial\Sigma}\right)\frac{1}{r}\frac{\partial}{\partial r}\left[3\,r^{1/2}\frac{\partial}{\partial r}\left(r^{1/2}\,\delta x\right)\right]. \tag{2.29}$$

This is a diffusion equation in x, with a diffusion constant $\partial x/\partial\Sigma$. If $\partial x/\partial\Sigma > 0$, a small perturbation δx diminishes. However, if $\partial x/\partial\Sigma < 0$, a small perturbation grows, i.e., leads to an instability.

This instability criterion can be re-formulated in terms of the steady-state solutions that are calculated here, if it is assumed that, following the perturbation in the surface density, the disc viscosity, temperature and the accretion rate quickly evolve to the steady-state solution in which the surface density equals the perturbed value $\Sigma_0 + \delta\Sigma$. Under that assumption, $\dot{M} \propto x$, and the instability develops if $\partial\dot{M}/\partial\Sigma < 0$, or, equivalently,

$$\partial\Sigma/\partial\dot{M} < 0. \tag{2.30}$$

The above assumption is likely to hold since the disc viscosity and accretion come from the MRI, which develops on the dynamical timescale (the thermal timescale also being shorter than

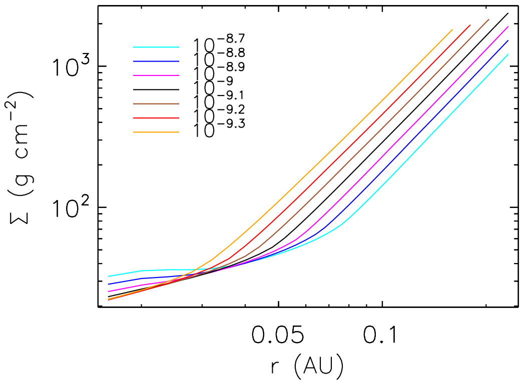


Figure 2.9: Surface density Σ as a function of radius, for a number of steady-state model solutions with $M_* = 1\,\mathrm{M}_\odot$, $\alpha_\mathrm{DZ} = 10^{-4}$ and varying accretion rate in the range $\dot{M} = 10^{-9.3} - 10^{-8.7}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. Outwards of ~0.03 AU, at a fixed orbital radius the surface density increases with decreasing accretion rate.

the viscous timescale). Therefore, the steady-state models with varying gas accretion rate can be used to check if the inner disc is viscously unstable.

This is done in Fig. 2.9 and Fig. 2.10. Fig. 2.9 shows the surface density as a function of radius for a number of steady-state solutions with varying gas accretion rate around the fiducial value of $\dot{M} = 10^{-9}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. The plot clearly shows that beyond ~0.03 AU the surface density increases with decreasing $\dot{M}$. Fig. 2.10 shows this more formally, i.e., it shows the derivative $\partial\Sigma/\partial\dot{M}$ evaluated using steady-state solutions at $\dot{M} = 10^{-9}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1} \pm 1\%$, with $\partial\Sigma/\partial\dot{M} < 0$ outwards of ~0.03 AU. Therefore, a large part of the inner disc is unstable to surface density perturbations, or viscously unstable.

What would be the outcome of this instability? If a positive perturbation in surface density (i.e., an over-density, $\delta\Sigma > 0$) leads to a lower gas accretion rate, the over-density would grow. Conversely, if $\delta\Sigma < 0$, producing a local under-density, the accretion rate increases, expelling even more material from the under-density. Therefore, the viscous instability would create axisymmetric over- and under-densities, i.e., rings and gaps in the disc.

Finally, consider the viscosity $\bar{\nu}$ in terms of the Shakura-Sunyaev $\bar{\alpha}$ parameter. The instability criterion can then be written as (within the steady-state formulation)

$$\frac{\partial(\ln\bar{\alpha})}{\partial(\ln\dot{M})} + 2\,\frac{\partial(\ln c_s)}{\partial(\ln\dot{M})} > 1. \tag{2.31}$$

The first and the second term on the left-hand side are plotted in Fig. 2.11 (from the same calculation as Fig. 2.10). This explicitly shows that it is the dependence of the MRI-driven $\bar{\alpha}$ on the gas accretion rate (and the surface density) that drives the instability, while variations in the sound speed c_s are small.

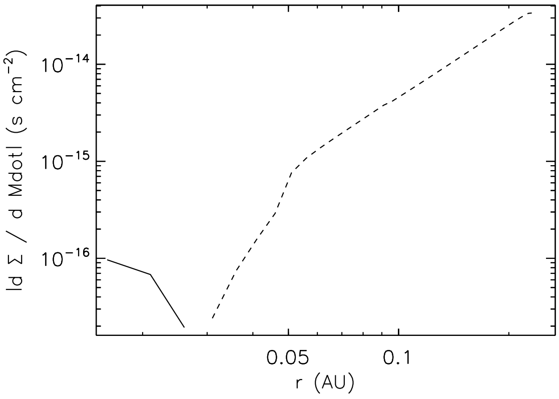


Figure 2.10: Derivative $|\partial\Sigma/\partial\dot{M}|$ as a function of radius for the fiducial disc model. The solid line denotes $\partial\Sigma/\partial\dot{M} > 0$ (viscously stable) while the dashed line denotes $\partial\Sigma/\partial\dot{M} < 0$ (viscously unstable). Outwards of ~0.03 AU, disc is viscously unstable.

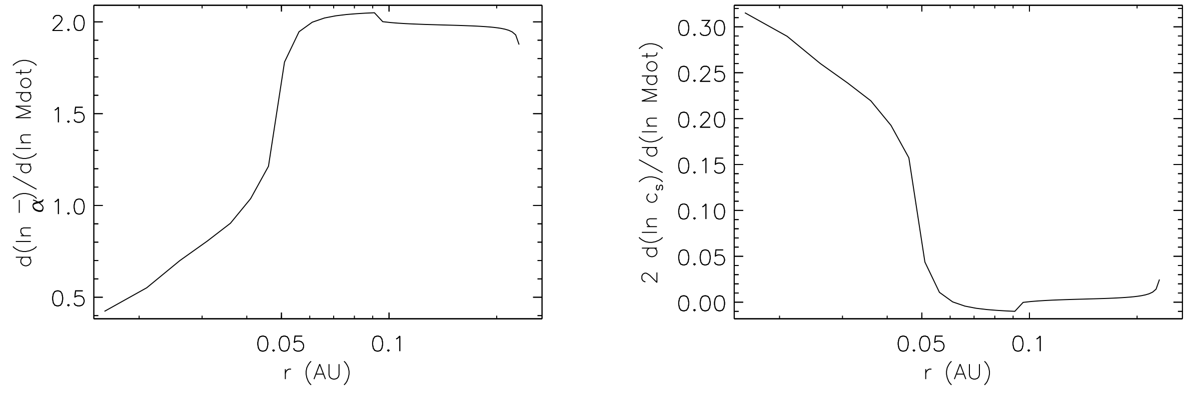


Figure 2.11: Derivatives $\partial(\ln\bar{\alpha})/\partial(\ln\dot{M})$ (top) and $2\partial(\ln c_s)/\partial(\ln\dot{M})$ (bottom) as a function of radius for the fiducial disk model.

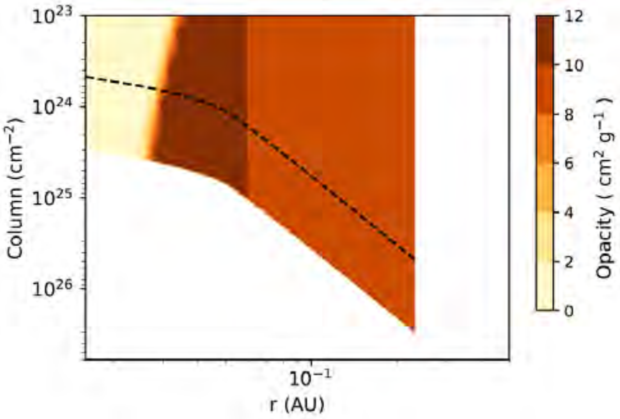


Figure 2.12: Rosseland mean opacity (in cm^2 g^{-1}) calculated *a posteriori* for the fiducial disc model, as a function of disc location (with height in units of vertical column density). The dashed curve denotes one pressure scale height. Over most of the region of interest in the disc, the derived opacity is within a factor of two of 10 cm^2 g^{-1}, consistent with the *a priori* assumed value of opacity.

2.3.1.5. Opacity

In all of the models presented in this chapter, disc opacity is assumed to be a constant, $\kappa = 10\,\mathrm{cm^2\,g^{-1}}$. In general, disc opacity is a function of pressure and temperature. The adopted constant value is representative of opacity due to small dust grains. Fig. 2.12 shows an *a posteriori* calculation of opacity due to dust grains, and also due to gas atomic and molecular lines (Zhu et al. 2009), using the disc temperatures and densities inferred from the steady-state model. In the bulk of the inner disc the calculated value is roughly in agreement with the adopted value. The disagreement is greatest in the innermost disc, where the temperature is higher than the sublimation temperature of dust grains, since the opacity adopted in this work is inaccurate when dust is no longer the contributing species.

2.3.1.6. Validity of the strong-coupling limit

The models presented here assume that ambipolar diffusion suppresses the MRI based on the criterion given by eq. (2.11). This criterion is based on numerical simulations of the MRI performed in the strong-coupling limit (referring to the coupling between ions and neutrals in the gas, and the way both jointly respond to the magnetic field; Bai and Stone 2011). The strong-coupling limit is valid if $\rho_\mathrm{n} \gg \rho_\mathrm{i}$, i.e. if the density of neutrals is much greater than the density of ions, and if the recombination timescale t_rcb for ions is much shorter than the dynamical timescale, i.e. $t_\mathrm{rcb} \ll t_\mathrm{dyn}$. The first condition is easily fulfilled here, since the only ionized species is potassium.

To check whether the second condition is fulfilled, we can calculate the recombination timescale t_rcb as described in Section 2.2.2.2 and compare it to t_dyn. The result is shown in Fig. 2.13. The region where $t_\mathrm{rcb} < t_\mathrm{dyn}$ is shown in green here, showing that this condition

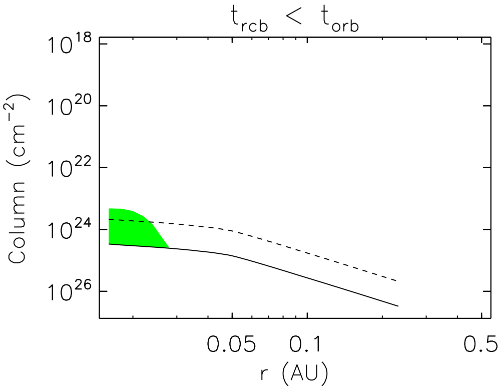


Figure 2.13: Comparison of the recombination timescale (t_{rcb}) and the dynamical timescale (t_{dyn}) as a function of location in the fiducial disk model. The solid black curve denotes the disc midplane, and the dashed curve denotes one pressure scale height. The green region shows where $t_{\rm rcb} < t_{\rm dyn}$ (and thus where the single-fluid approximation is valid).

is fulfilled only near disc midplane inwards of ∼0.03 AU. This result indicates that collisional recombinations are too slow to establish the ionization equilibrium on the orbital timescales (at which the MRI develops). It thus follows that more realistic (more complex) chemical networks ought to be considered in order to rely on eq. (2.11) (such as in the work by Bai 2011b).

2.3.2. Varying model parameters

2.3.2.1. Viscosity in the dead zone

In the results presented so far, the viscosity parameter in the dead zone (and the zombie zone) is assumed to be $\alpha_{\rm DZ} = 10^{-4}$. However, this is an uncertain parameter, with weak constraints from numerical simulations. To explore the sensitivity of the inner disc structure to this value, in Figs. 2.14–2.17 I show the results for two models in which the accretion rate $\dot{M}$ and the stellar mass M_* are the same as in the fiducial model, but viscosity parameter in the dead zone is taken to be $\alpha_{\rm DZ} = 10^{-5}$ and 10^{-3}.

Fig. 2.15 shows that, since $\alpha_{\rm DZ}$ is the minimum value of $\bar{\alpha}$, its precise value can drastically change the location of the pressure maximum $r_{P_{\rm max}}$ (where $\bar{\alpha}$ falls to $\alpha_{\rm DZ}$). Larger $\alpha_{\rm DZ}$ implies a smaller $r_{P_{\rm max}}$. In fact, since outwards of ∼0.05 AU $\bar{\alpha}$ decreases approximately as a power-law of radius, there is a simple (approximate) relationship between the two, $r_{P_{\rm max}} \propto \alpha_{\rm DZ}^{-1/4}$ (see also Fig. 2.17).

In terms of other disc features, the models with varying $\alpha_{\rm DZ}$ appear to be qualitatively the same to the fiducial model. The value of $\alpha_{\rm DZ}$ does affect the magnetic field strength outwards from ∼0.09 AU, where larger $\alpha_{\rm DZ}$ requires a smaller B (shown in Fig. 2.15). This corresponds to the region where a dead zone exists around disc midplane (see Fig. 2.14). In this region, the dead-zone viscosity contributes significantly to the total vertically-averaged viscosity parameter.

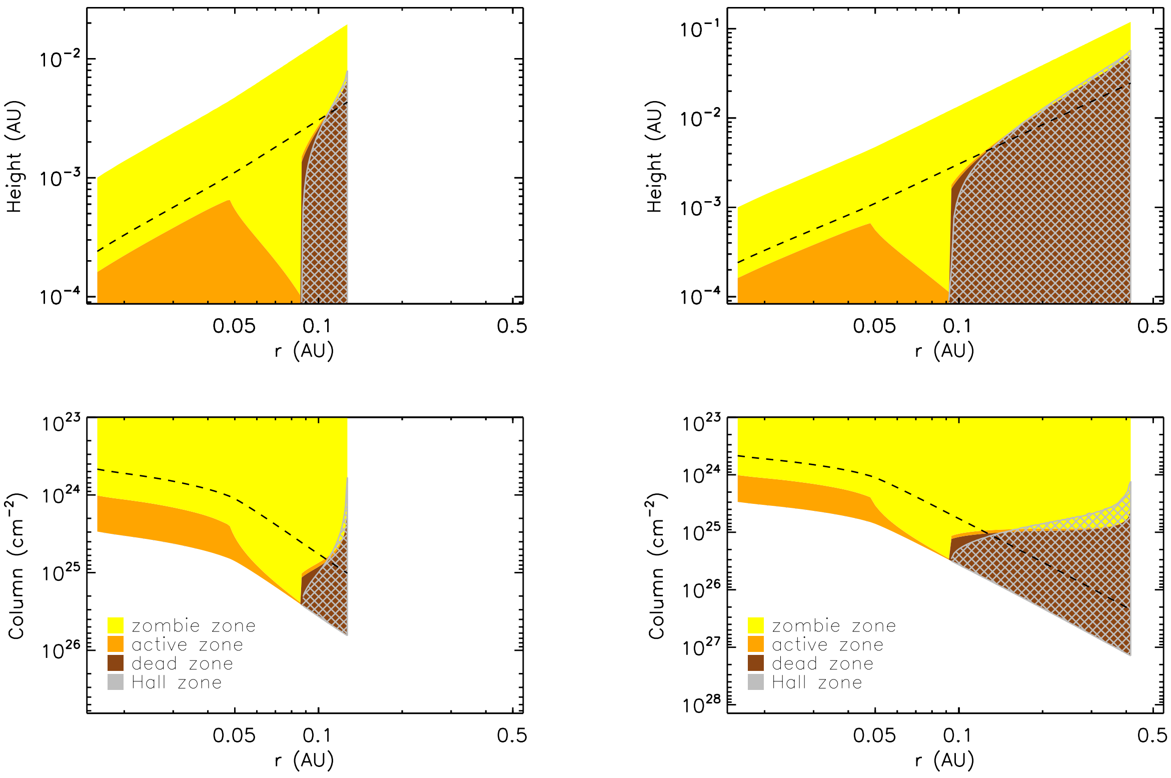


Figure 2.14: Same as Fig. 2.4, but now for $\alpha_{\rm DZ} = 10^{-3}$ (left) and 10^{-5} (right).

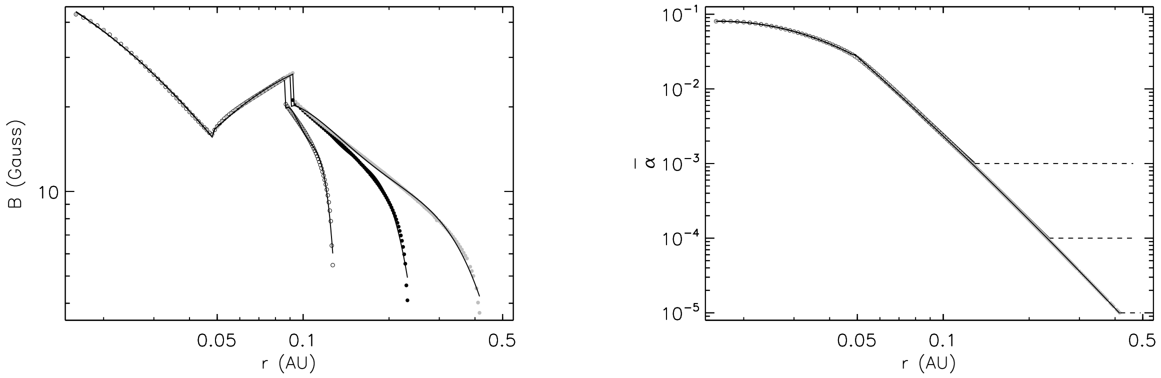


Figure 2.15: Left: Magnetic field strength B as a function of radius. Right: $\bar{\alpha}$ as a function of radius. Both figures show results for the fiducial model, $\alpha_{\rm DZ} = 10^{-4}$ (filled black circles), results for $\alpha_{\rm DZ} = 10^{-3}$ (empty circles) and $\alpha_{\rm DZ} = 10^{-5}$ (filled grey circles).

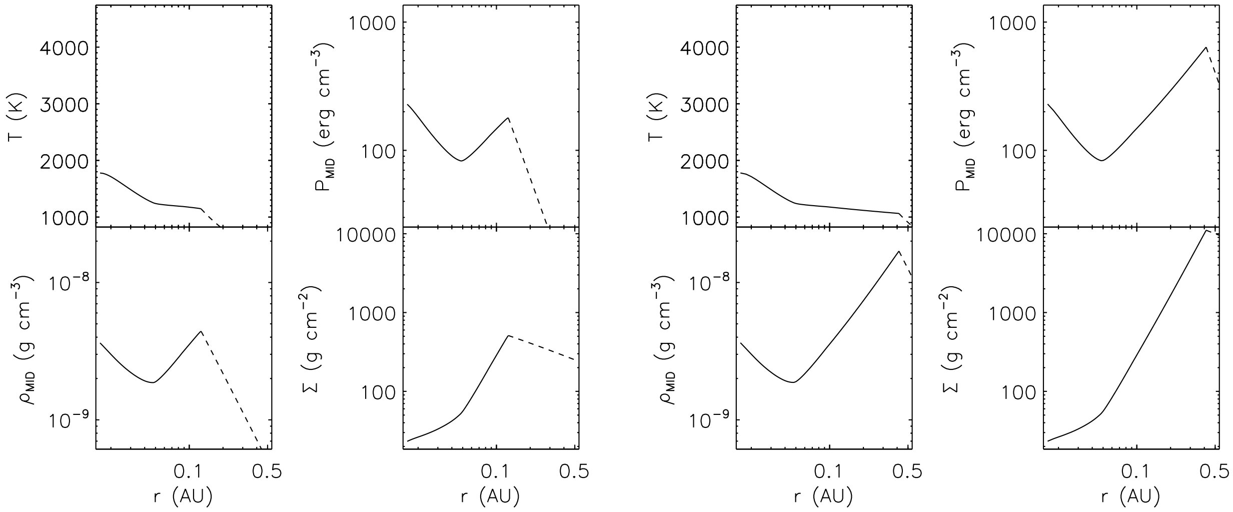


Figure 2.16: Various disc structure quantities as a function of radius: same as Fig. 2.8, but now for $\alpha_{\rm DZ} = 10^{-3}$ (left) and 10^{-5} (right).

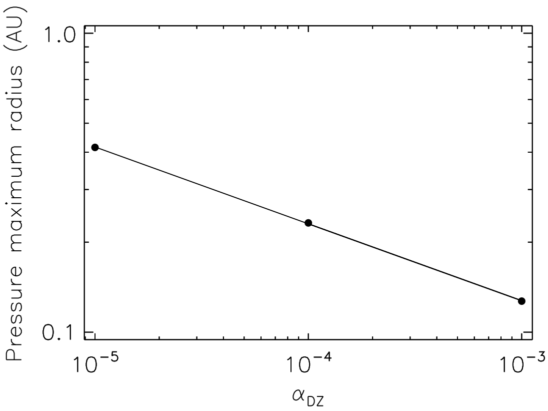


Figure 2.17: Radial location of the pressure maximum as a function of $\alpha_{\rm DZ}$, showing the approximate power-law dependence $r_{P_{\rm max}} \propto {\alpha_{\rm DZ}}^{-1/4}$.

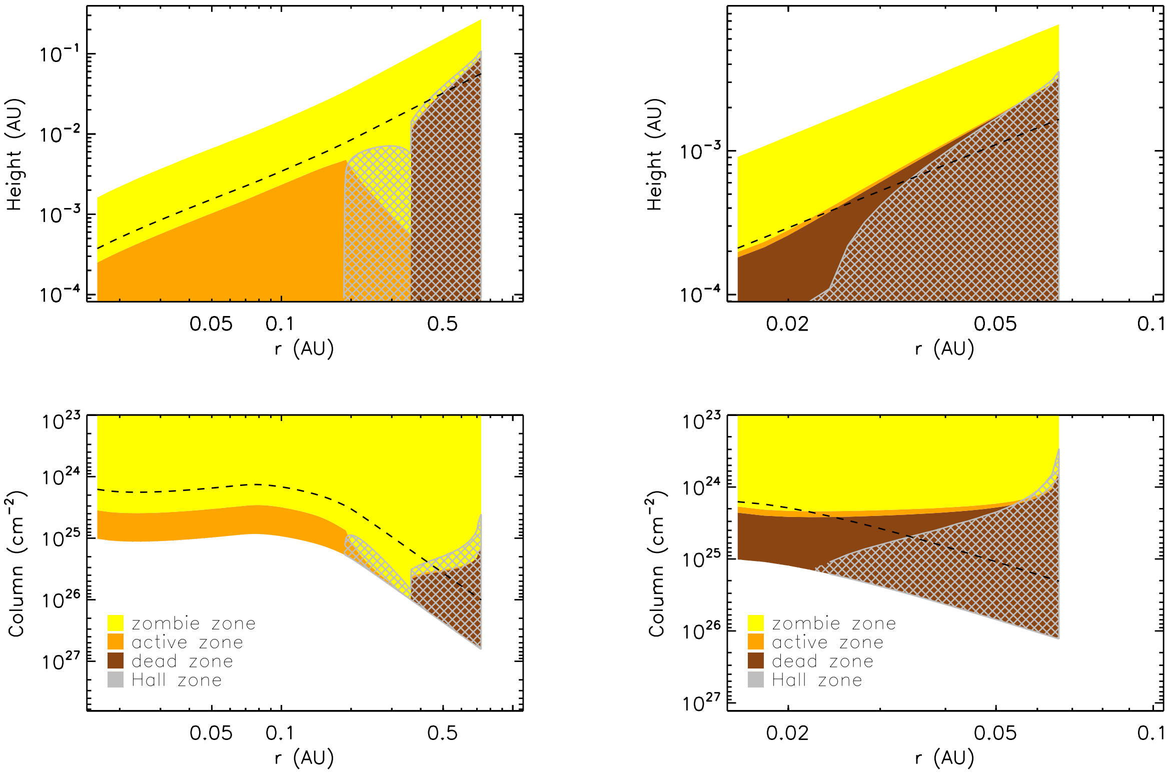


Figure 2.18: Same as Fig. 2.4, but now for an accretion rate $\dot{M} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ (left) and $10^{-10}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ (right).

2.3.2.2. Accretion rate

Next, I discuss the effects of varying the gas accretion rate $\dot{M}$ while keeping the stellar mass M_* and the viscosity parameter in the dead zone α_{DZ} constant. Results for a model with $\dot{M} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ and a model with $\dot{M} = 10^{-10}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ are shown in Figs. 2.18–2.22.

The main effect of varying $\dot{M}$ is that the viscous dissipation rate changes directly proportionally. Thus, higher $\dot{M}$ makes the disc hotter, and more ionized, increasing the MRI-driven $\bar{\alpha}$ at any given radius (where the MRI is active). For the outer regions, this means that for higher $\dot{M}$ the pressure maximum moves radially outwards, approximately as a power law $r_{P_{\mathrm{max}}} \propto \dot{M}^{1/2}$ (see Fig. 2.22). In the inner regions, the viscosity parameter saturates at around $\bar{\alpha} \sim 0.1$ (shown in Fig. 2.20). This happens because in the innermost disc (at high enough accretion rate) all of potassium becomes ionized.

Some qualitative differences also arise in where the disc is MRI-active. Fig. 2.18 shows that for $\dot{M} = 10^{-10}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, the midplane is not turbulent anywhere in the disc. Additionally, Fig. 2.18 shows for $\dot{M} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ the Hall zone (where the Hall effect might significantly affect the accretion and the disc structure) extends inwards of the dead zone inner edge, implying that at high $\dot{M}$ the pressure maximum might not move as outwards as implied by Fig. 2.22.

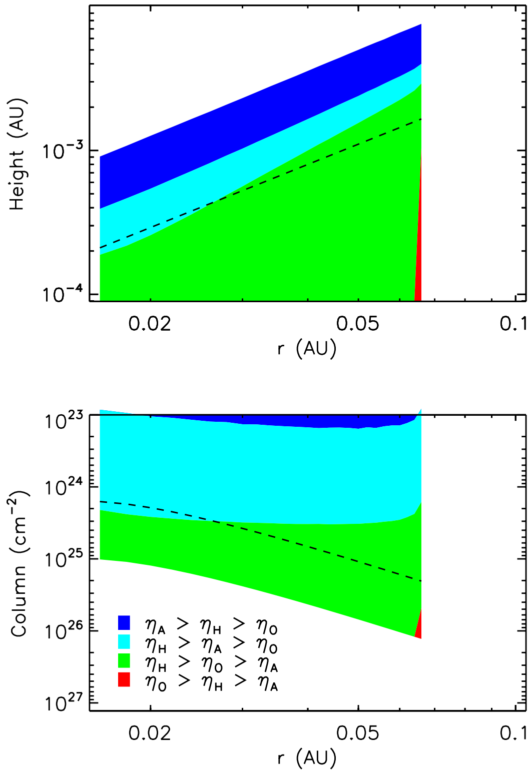


Figure 2.19: Same as Fig. 2.3, but for an accretion rate $\dot{M} = 10^{-10}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$. This is the only one of the disc models shown here in which an Ohmic-dominated region arises (red region in bottom right corner of both panels).

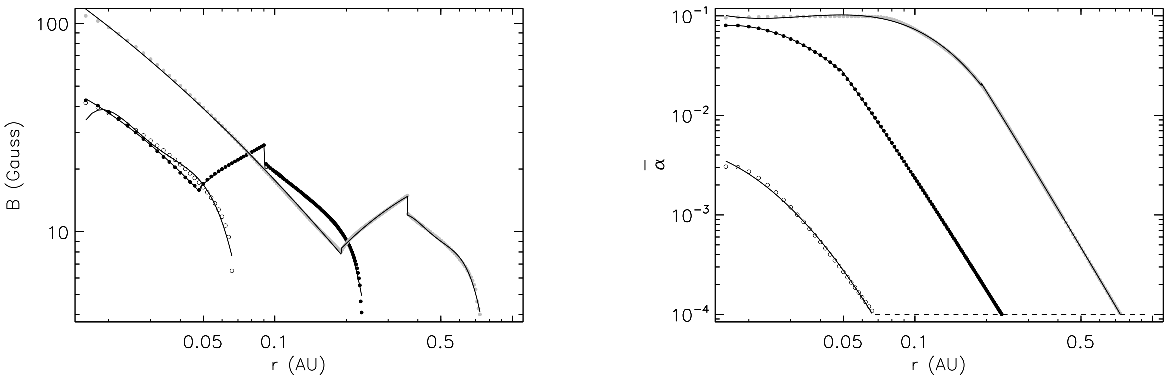


Figure 2.20: Left: Magnetic field strength B as a function of radius. Right: $\bar{\alpha}$ as a function of radius. Both figures show results for the fiducial model with $\dot{M} = 10^{-9}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ (filled black circles), a model with $\dot{M} = 10^{-10}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ (empty circles) and $\dot{M} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ (filled grey circles).

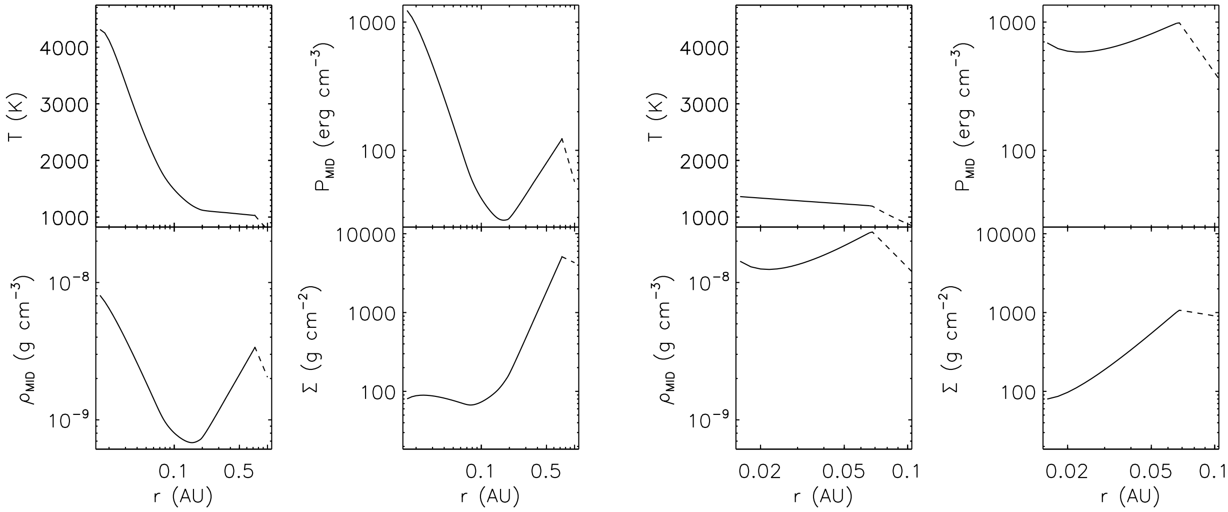


Figure 2.21: Same as Fig. 2.8, but now for an accretion rate $\dot{M} = 10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ (left) and $10^{-10}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ (right).

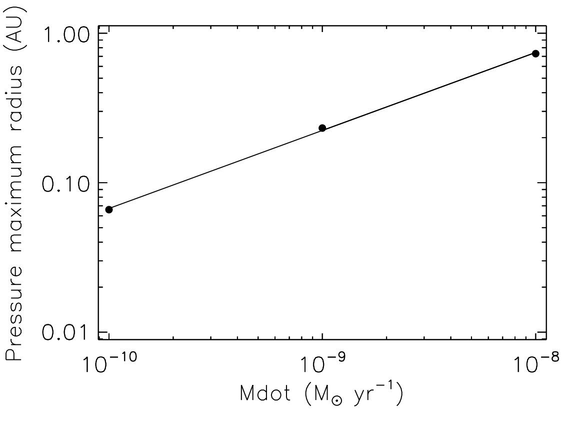


Figure 2.22: Radial location of the pressure maximum as a function of the accretion rate $\dot{M}$, showing the approximate power law dependence $r_{P_{\max}} \propto \dot{M}^{1/2}$.

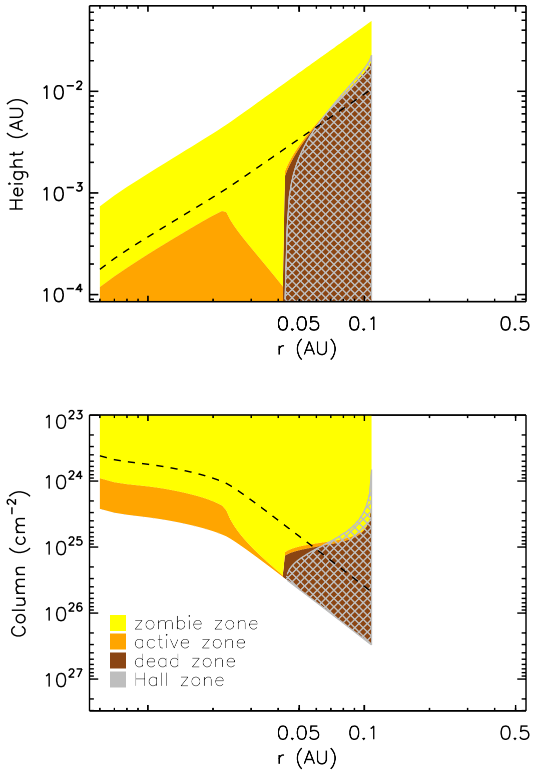


Figure 2.23: Same as Fig. 2.4, but for now for a stellar mass $M_* = 0.1\,\mathrm{M}_\odot$.

2.3.2.3. *Stellar mass*

Figs. 2.23–2.25 show the results for a model with the fiducial $\dot{M}$ and α_{DZ}, but with a stellar mass of $M_* = 0.1\,\mathrm{M}_\odot$ (and a stellar radius $R_* = 1\,\mathrm{R}_\odot$). These results show that the structure of the inner disc around stars of different mass is virtually the same, only shifted radially inwards. The location of the pressure maximum scales with the stellar mass as a power law, $r_{P_{\mathrm{max}}} \propto M_*^{-1/3}$. This follows from the Shakura-Sunyaev equations, in which the stellar mass and orbital radius can always be combined into M_*/r^3 (neglecting f_r, since for $r \gg R_*$, $f_r \approx 1$). Note that these models do not include heating of the disc by stellar irradiation, and so the stellar radius R_* only affects the disc structure through f_r, which can be neglected at r larger than a few R_*.

2.4. Discussion and conclusions

I used the Shakura-Sunyaev α-disc model coupled with a parametrization of the MRI-driven α to investigate the structure of the thermally-ionized inner regions of protoplanetary discs in steady-state. I produced models of discs with various gas accretion rates (10^{-10}–$10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$), stellar masses (0.1–1 $\mathrm{M}_\odot$) and minimum values of α (i.e., value of α in MRI-dead zones; 10^{-5}–

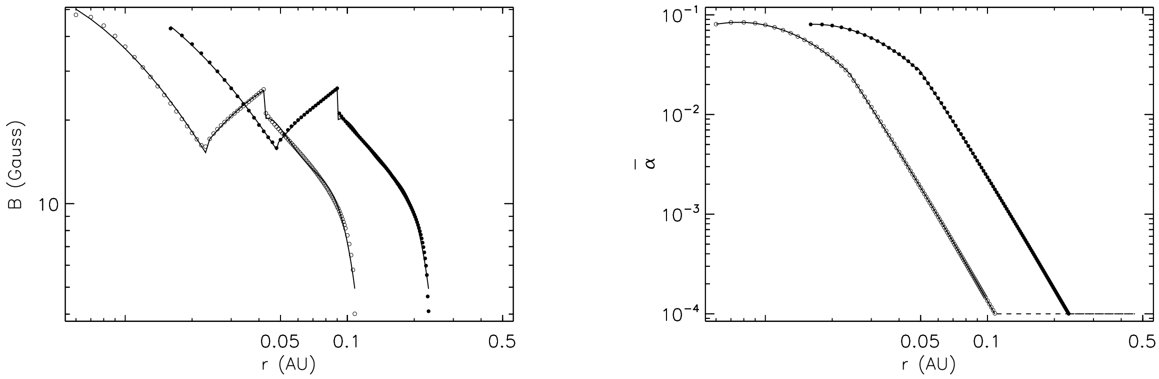


Figure 2.24: Left: Magnetic field strength B as a function of radius. Right: $\bar{\alpha}$ as a function of radius. Both figures show results for the fiducial model with $M_* = 1\,\mathrm{M}_\odot$ (filled black circles) and a model with $M_* = 0.1\,\mathrm{M}_\odot$ (empty circles).

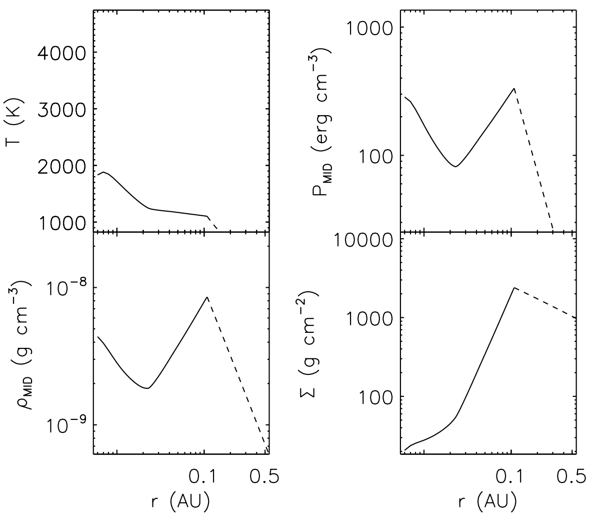


Figure 2.25: Same as Fig. 2.8, but now for a stellar mass $M_* = 0.1\,\mathrm{M}_\odot$.

10^{-3}).

All of the models feature a local gas pressure maximum at a few tenths of an AU from the star. Radially outwards from the pressure maximum the MRI is completely suppressed; just inwards of it the MRI is suppressed around disc midplane, but an MRI-active zone exists above (and below) the midplane. It is only closer to the star (and not in all models) that the disc midplane becomes MRI-active. In other words, at the location of the pressure maximum, the disc midplane is weakly-turbulent. This could have an important impact on the growth and evolution of pebbles that become trapped at this location. The exact orbital radius of the pressure maximum scales with model parameters via simple power-laws: $r_{P_{\max}} \propto \alpha_{\rm DZ}^{-1/4}$, $M_*^{1/3}$ and $\dot{M}^{1/2}$.

These models include suppression of the MRI by Ohmic and ambipolar diffusion, but not the Hall effect. To explore how the Hall effect would modify the obtained results, I calculated which regions of the disc have a Hall Elsasser number $\chi < 1$. This Hall zone largely coincides with the Ohmic dead zone, and partially with the MRI-active zone. Therefore, the Hall effect could move the pressure maximum radially inwards, if it simply shuts off the MRI within this Hall zone. The effects are less clear if the Hall effect leads to an increased efficiency of accretion, in which case the arising disc structure would depend on the physics at the boundary between the MRI-viscous, turbulent region and Hall-dominated inviscid, laminar region.

Note also that the suppression of the MRI by ambipolar diffusion is considered in the strong-coupling limit, which is found to be violated over most of the disc due to long recombination timescales. However, the models presented here consider only a single ionized species, and a more complex chemical network would have more recombination pathways. In particular, small dust grains can adsorb free charges, which then recombine on grain surfaces faster than in the gas phase. Including dust grains in the chemistry would have far-reaching consequences, since this might effectively reduce the ionization levels compared to the case where grains are absent, and might lead to a stronger suppression of the MRI (Sano et al. 2000; Ilgner and Nelson 2006; Wardle 2007; Salmeron and Wardle 2008; Bai and Goodman 2009), although dust can also positively contribute to the disc ionization state at high temperatures (Desch and Turner 2015).

Additionally, the above results show that the location of the pressure maximum does not correspond to the dead zone inner edge, but rather to the active zone outer edge. However, there could be no such outer edge in a disc that is also ionized by stellar X-rays and cosmic rays. As discussed in Section 1.2.3, at short orbital radii the X-rays and the cosmic rays may activate the MRI in the disc upper layers. The thermally-ionized active zone in the inner disc would then smoothly connect to the outer disc, see Fig. 1.5. In that case, the exact location of the pressure maximum, coinciding with the minimum in disc viscosity, would additionally depend on the relative importance of thermal and non-thermal sources of ionization. In Chapters 5 and 6 I will revisit the inner disc model, and consider both the effects of X-rays and cosmic rays, and the dust grains.

Finally, at a given orbital radius, the vertically-averaged viscosity parameter increases with gas accretion rate so much that the gas surface density decreases, $\partial\Sigma/\partial\dot{M} < 0$. This implies that the steady-state inner disc is unstable to perturbations in the gas surface density. This so-called viscous instability could lead to gas concentrating into rings. However, further study is required to determine if the disc is stabilized when a more detailed disc chemical and thermal structure is considered.

3 EARLY STAGES OF PLANET FORMATION IN THE INNER DISC

3.1. Introduction

In Chapter 2 I discussed a steady-state model of the gas in the inner regions of protoplanetary discs. In this chapter, I examine the evolution of dust in the obtained gas structure.

The main goal of this chapter is to examine if the evolution of dust leads to enhancement of solids at short orbital periods, as necessary for the formation of the super-Earths. To create a dust-rich inner disc, the radial drift of dust particles from the outer disc needs to be stopped or slowed down. The radial drift of small particles (that are in the Epstein drag regime) slows down closer to the star in conventional disc models, i.e., in which the gas density and pressure increase radially inwards (Youdin and Shu 2002; Youdin and Chiang 2004; Birnstiel et al. 2010; 2012; Drazkowska et al. 2016). This can concentrate dust in the inner disc to some extent. On the other hand, if dust grains are large and thereby only marginally coupled to the gas, the radial drift can be fully stopped inside an axisymmetric local gas pressure maximum that is expected to form if the accretion in the inner disc is driven by the MRI (Kretke et al. 2009; Dzyurkevich et al. 2010; Drazkowska et al. 2013; Chatterjee and Tan 2014). How coupled the dust grains are to the gas depends on their size (Weidenschilling 1977). Therefore, I consider both the spatial evolution of the dust, and the evolution of the dust grain size, including growth, fragmentation and radial drift.

Furthermore, to form the solid core of a planet, accumulated dust grains must grow larger. As discussed in Chapter 1, collisions and sticking between small dust grains produce millimetre to centimetre-sized pebbles. These pebbles do not stick as efficiently as small grains, raising the question of how solids proceed to grow beyond these sizes. One suggestion is that larger solid bodies, tens of kilometres in size (so-called planetesimals), may grow directly from small pebbles, by the concentration of pebbles into dense clumps that become gravitationally unstable (Youdin and Goodman 2005). Following dust evolution calculations in sections 3.2 and 3.3, I investigate the possibility of planetesimal formation in the inner disc in Section 3.4.

An important caveat to the calculations presented in this chapter is that the dust is evolved in a gas disc that is fixed in time. Dust grains can act to suppress the MRI by lowering the coupling between the gas and the magnetic field (e.g. Sano et al. 2000; Ilgner and Nelson

2006), and could thus significantly alter the gas disc structure. Moreover, at high dust-to-gas ratios, which I show can be achieved in the inner disc, dust becomes dynamically important and the dynamical back-reaction on the gas should be taken into account. In Section 3.5, these effects, and how they might influence the obtained results, are discussed qualitatively. The self-consistent feedback of dust enhancement on the gas disc structure should be addressed in future studies.

3.2. Methods

3.2.1. Gas disc model

The structure of the gas disc (gas surface density, temperature, pressure and viscosity) is obtained from the steady-state models of the inner protoplanetary gas disc presented in Chapter 2. In these models the viscosity (α parameter[1]) is determined self-consistently with the disc structure (Shakura and Sunyaev 1973), thermal ionization and MRI criteria (Bai and Stone 2011; Bai 2011a). The parameters of the model are the stellar mass and radius (M_*, R_*), gas accretion rate ($\dot{M}_\mathrm{g}$), and assumed viscosity parameter inside the MRI-dead zones (α_DZ). Recall that α_DZ is also the minimum viscosity parameter, in the sense that $\alpha = \alpha_\mathrm{DZ}$ even if the MRI is active, if it implies a lower value. In this work I primarily use $M_* = 1\,\mathrm{M}_\odot$, $R_* = 2.33\,\mathrm{R}_\odot$, $\dot{M}_\mathrm{g} = 10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ and $\alpha_\mathrm{DZ} = 10^{-4}$. The disc structure for these parameters is shown in Fig. 3.1. The local gas pressure maximum is at an orbital distance of $\sim 0.7\,\mathrm{AU}$, and temperature and surface density at that location are $\sim 1000\,\mathrm{K}$ and $\sim 5000\,\mathrm{g\,cm^{-2}}$, respectively. Outwards from the pressure (and the surface density) maximum the MRI is suppressed and $\alpha = \alpha_\mathrm{DZ}$.

Consider, briefly, how the disc structure depends on these parameters. For a higher gas accretion rate $\dot{M}_\mathrm{g}$ the radial α profile is roughly an outward-translated version of the one shown in the top panel of Fig. 3.1, and inward-translated for a smaller $\dot{M}_\mathrm{g}$. The radial location of the local gas pressure maximum scales with the accretion rate approximately as $\dot{M}_\mathrm{g}^{1/2}$. For a higher α_DZ the α falls to this value closer to the star, and vice versa, but the value of α as a function of orbital distance remains almost the same otherwise; as a result, the radial location of the local pressure maximum scales with the minimum dead-zone viscosity as $\alpha_\mathrm{DZ}^{-1/4}$. Furthermore, inwards of the pressure maximum the temperature has to be sufficiently high for thermal ionization of potassium to support the MRI and so it is always larger than 1000 K, regardless of the exact choice of $\dot{M}_\mathrm{g}$ and α_DZ. For a steady-state vertically-isothermal α-disc the accretion rate is $\dot{M}_\mathrm{g} = 3\pi c_\mathrm{s}^2 \alpha \Sigma_\mathrm{g}/\Omega$ (ignoring an additional factor which depends on the boundary condition at the inner disc edge, and which becomes unimportant far away from the edge). Further assuming (as is approximately the case) that the temperature at the location of the pressure maximum is constant regardless of the disc parameters, it follows that the maximum gas surface density approximately depends on the disc parameters[2] as $\dot{M}_{\rm g}^{1/4}$ and $\alpha_{\rm DZ}^{-5/8}$.

---

[1] For simplicity, in this chapter I denote the vertically-averaged viscosity parameter $\bar{\alpha}$ simply as α.

The dynamical back-reaction of the dust on the gas and the effect of the dust on the MRI are neglected in this chapter, and the timescale on which the dust evolves is much shorter than the Myr timescale on which the accretion rate evolves. Thus, the structure of the gas disc is held fixed, i.e. not evolved in time.

3.2.2. Dust evolution model

The dust particle size distribution is evolved using the two-population model of Birnstiel et al. (2012). The dust surface density $\Sigma_{\rm d}$ is evolved using the advection-diffusion equation

$$\frac{\partial \Sigma_{\rm d}}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(\Sigma_{\rm d}\bar{u} - D_{\rm gas}\Sigma_{\rm g}\frac{\partial}{\partial r}\left(\frac{\Sigma_{\rm d}}{\Sigma_{\rm g}}\right)\right)\right] = 0\,, \tag{3.1}$$

where r is the cylindrical radius, $\bar{u}$ is the dust advection velocity, $D_{\rm gas}$ is the gas diffusivity (assumed to be equal to the MRI-driven gas viscosity) and $\Sigma_{\rm g}$ is the gas surface density.

The dust advection velocity is a sum of the velocities due to advection with the accreting gas and radial drift. For particles with Stokes number $\mathrm{St}_i = \pi\rho_{\rm s}a_i/2\Sigma_{\rm g}$ (with $\rho_{\rm s}$ the internal density of the dust and a_i the particle size), and adopting the terminal velocity approximation (e.g. Takeuchi and Lin 2002), the dust velocity is given by

$$u_i = \frac{1}{1+\mathrm{St}_i^2}u_{\rm gas} + \frac{2}{\mathrm{St}_i + \mathrm{St}_i^{-1}}u_{\rm drift}\,, \tag{3.2}$$

where

$$u_{\rm drift} = \frac{c_{\rm s}^2}{2v_{\rm K}}\frac{{\rm d}\ln P}{{\rm d}\ln r}\,, \tag{3.3}$$

with $c_{\rm s}$ the speed of sound, $v_{\rm K}$ the Keplerian velocity and P the midplane gas pressure. Small particles ($\mathrm{St} \ll 1$) move with the gas, and larger particles can move faster or slower than the gas, depending on the sign of the pressure gradient.

In the Birnstiel et al. (2012) model the dust surface density $\Sigma_{\rm d}$ and dust advection velocity $\bar{u}$ are the sum and the mass-weighted average, respectively, of the surface density and velocity of two populations of particles: small monomer-sized particles ($a_0 = 1\,\mu$m) and large particles (a_1). The size of the large particles evolves in time and space. At first, dust grains grow, starting at the monomer size. Then, at each radius the size of the large particles is set by whichever process yields the smallest size limit: radial drift ($a_{\rm drift}$), where grains larger than $a_{\rm drift}$ radially drift more quickly than they can grow; drift-fragmentation ($a_{\rm df}$), where grains larger than $a_{\rm df}$ fragment due to relative radial drift velocities; or turbulent fragmentation ($a_{\rm frag}$), where grains

---

[2] Derivation of the scalings from the Shakura-Sunyaev equations takes into account the small correction due to the dependence of the temperature on the disc parameters and yields that the maximum surface density depends on the disc parameters as $\dot{M}_{\rm g}^{3/10}$ and $\alpha_{\rm DZ}^{-13/20}$. These small corrections are omitted here for simplicity.

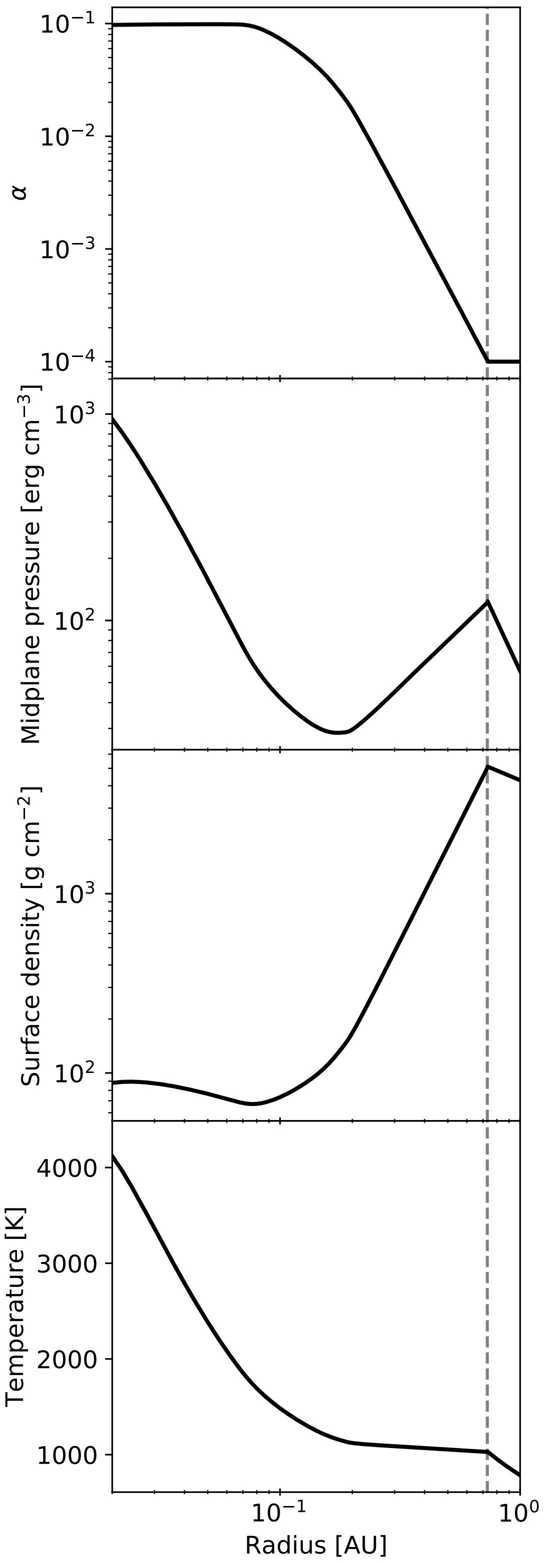


Figure 3.1: Gas disc structure from the steady-state model for $M_* = 1\,\mathrm{M}_\odot$, $R_* = 2.34\,\mathrm{R}_\odot$, $\dot{M}_\mathrm{g} = 10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ and $\alpha_\mathrm{DZ} = 10^{-4}$. From top to bottom: α parameter, midplane pressure, surface density and temperature, as functions of radius. Location of the local gas pressure maximum due to the MRI is indicated by the vertical dashed line.

larger than $a_{\rm frag}$ fragment due to relative velocities induced by turbulence.

In the innermost protoplanetary disc, inside the water ice line, the main dust species are silicate grains. Therefore, the bulk density of particles is set to $\rho_{\rm s} = 3\,{\rm g\,cm^{-3}}$, and the critical fragmentation velocity to $u_{\rm f} = 1\,{\rm m\,s^{-1}}$, based on experiments on collisions of silicate grains (Blum and Münch 1993; Beitz et al. 2011; Schräpler et al. 2012; Bukhari Syed et al. 2017) of similar size (the regime applicable in the Birnstiel et al. 2012 model used here).[3]

The viscosity parameter α due to the MRI-turbulence from the gas disc model is determined as a vertical average at each radius. Here, it is assumed that this vertically-averaged α equals the strength of turbulence that particles feel, which in turn determines the particle size due to turbulent fragmentation ($a_{\rm frag}$) and the radial turbulent mixing (diffusivity $D_{\rm gas}$). However, the viscosity (and the level of turbulence) can be different at the disc midplane compared to the upper layers of the disc, depending on where the non-ideal magnetohydrodynamic effects suppress the MRI. Dust tends to settle towards midplane, its scale height being determined by the balance between gravitational settling and vertical turbulent mixing (e.g. Youdin and Lithwick 2007). The use of a vertically-averaged α could thus be invalid in regions where the disc midplane is weakly turbulent. Nevertheless, I proceed with this assumption and check the robustness of the obtained results by swapping the vertically-averaged α parameter for the midplane value in one run, and recover qualitatively the same results.

3.2.3. Numerical methods

The advection-diffusion equation (3.1) is integrated using an explicit first order in time and second order in space finite element method. The advection term is integrated with an upwind scheme using a van Leer flux limiter. The numerical scheme has been developed and described in detail by Owen (2014). The modification made for this work, implemented by J. E. Owen, is the inclusion of the Birnstiel et al. (2012) dust evolution algorithm. The simulations in this work use 262 cells in the radial direction with a 0.002 AU spacing inwards of 0.4 AU and a 0.01 AU spacing outwards. The inner boundary is set to 0.016 AU, and the outer boundary is set to 1 AU (i.e. outside the pressure maximum, but inside the water ice line). The time-step is set with respect to the spatial resolution, the advection speed and the diffusion coefficient, so that it obeys the Courant-Friedrichs-Lewy condition. Following Birnstiel et al. (2012), at each time step the size of the large particles is updated to the smallest of the four size limits ($a_{\rm growth}$, $a_{\rm drift}$, $a_{\rm df}$ or $a_{\rm frag}$). The outer boundary condition is a constant dust accretion rate. This is not,

---

[3] Note that simulations of grain collisions (Meru et al. 2013) indicate that the critical fragmentation velocity could be significantly higher for porous grains than for compact ones, for a range of porosities that is not robustly covered by the above experiments. In the fragmentation-limited regime particle size depends quadratically on $u_{\rm f}$. If porosity is important, and $u_{\rm f}$ is higher by, e.g., a factor of 10, particle sizes (and their Stokes number) would be larger by a factor of 100, strongly affecting how coupled particles are to the gas flow and how susceptible to radial drift. However, there is a variety of processes that lead to particles becoming less porous (compactified) – e.g., collisions that result in coagulation (simulations by Meru et al. 2013, experiments by Kothe et al. 2010), collisions that result in bouncing (Weidling et al. 2009) and collisions of larger grains with monomers (Schräpler and Blum 2011). Therefore, I use the compact grain value of $u_{\rm f} = 1\,{\rm m\,s^{-1}}$.

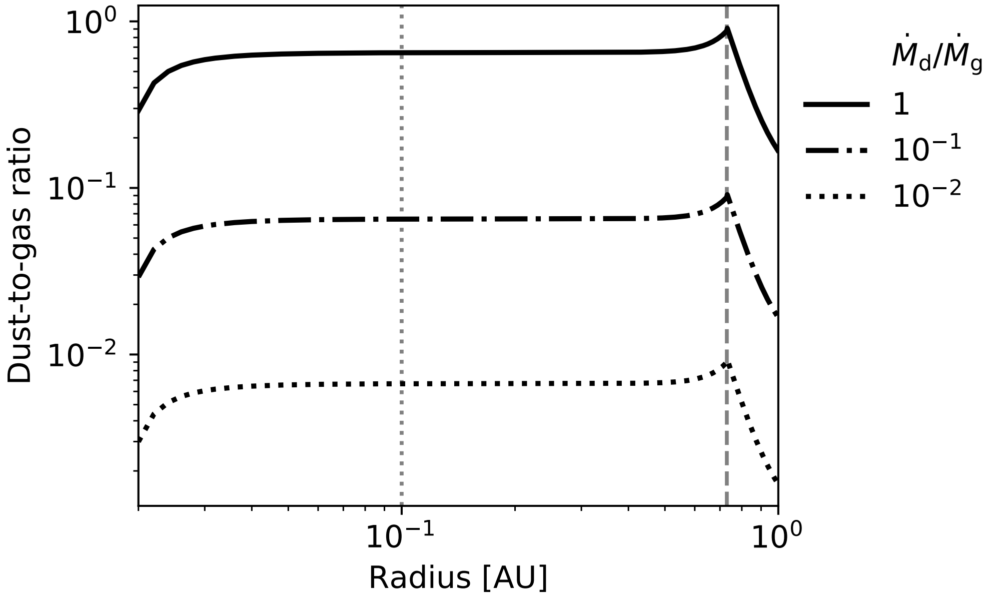


Figure 3.2: Dust-to-gas ratio Σ_d/Σ_g as a function of radius after 0.2 Myr, for various dust accretion rates at the outer boundary, as indicated in plot legend. Location of the local gas pressure maximum due to the MRI is indicated by the vertical dashed line, and location of the dust sublimation line by the vertical dotted line.

in general, a fixed quantity and it can vary with time, determined by the radial drift of dust from the outer disc (e.g. Birnstiel et al. 2012). Therefore, in this work it is varied in different simulations. Dust particle size at the outer boundary is calculated self-consistently.

Furthermore, the effects of dust sublimation are neglected. Note that the temperature in the adopted gas disc model exceeds the dust sublimation temperature (~1500 K) only inwards of 0.1 AU.

3.3. Results

Initially, the dust-to-gas ratio is 0.01 at all radii and all dust grains are monomers ($a = 1\,\mu$m). The dust is evolved for 0.2 Myr, by which time it has reached steady state. The steady-state dust-to-gas ratio is shown in Fig. 3.2 as a function of radius for three different values of dust accretion rates $\dot{M}_d$ at the outer boundary condition. For any given dust accretion rate, the steady-state dust-to-gas ratio is roughly constant inwards of the pressure maximum (indicated by the vertical dashed line), and it decreases outwards from the pressure maximum. There is only a moderate accumulation of dust at the pressure maximum, compared to the rest of the inner disc. This implies that the pressure maximum does not efficiently trap dust particles.

Essentially, the particles do not feel significant gas drag, and thus do not significantly feel the effect of the change in the sign of the pressure gradient inwards of the pressure maximum. This happens because, as a result of the MRI-induced turbulence and particle fragmentation, the particle size is very small. Fig. 3.3 shows the three dust size limits (due to radial drift, drift-fragmentation and turbulent fragmentation) as functions of radius, calculated in steady state. The smallest of the three (a_{frag}), due to turbulent fragmentation, sets the size of the population of large particles in these simulations, which dictates the evolution of dust overall. Particle size is thus limited to only a few millimetres near the pressure maximum (indicated by the vertical dashed line), and the particles are monomer-sized in the innermost disc.

The particle size determines, through the Stokes number, how coupled the dust is to the gas. Thus, the particle size determines to what extent the particles move with the accreting gas towards the star and also by how much they are slowed down or sped up by the gas drag. In this case, inwards of the pressure maximum (and outwards from the pressure minimum at $\sim 0.2\,\mathrm{AU}$) the gas drag acts outwards ($u_{\mathrm{drift}} > 0$). However, as dust particles are small inside the pressure maximum ($\mathrm{St} \sim 4 \times 10^{-4}$ even for the large particles) and their size further decreases inwards, the dust advection velocity is outwards only in a very narrow region. Consequently, after accounting for the diffusivity (i.e., the radial turbulent mixing of dust), the accumulation of solids inside the pressure maximum is moderate compared to the rest of the inner disc.

The dust advection velocity used here is a mass-weighted average of the velocity of the monomer-sized particles and the large particles (of size a_{frag}, Fig. 3.3). The monomer-sized particles are advected by the gas through the pressure maximum, and it can be shown that an individual large particle will also not be trapped. This is because the dust particles are in the fragmentation limit, in which the particles are fragmented faster than they drift. The drift timescale for the large particles inside the pressure trap (the region inwards of the pressure maximum where their advection velocity is outwards) can be estimated by

$$t_{\mathrm{drift}} \approx \frac{1}{2}\frac{d_{\mathrm{trap}}}{u_1(r_{P_{\mathrm{max}}})}\,, \tag{3.4}$$

where $d_{\mathrm{trap}} \approx 0.06\,\mathrm{AU}$ is the radial width of the trap, and the particle velocity is $u_1 \approx 2\,\mathrm{cm\,s^{-1}}$ (see eq. (3.2)) is evaluated just inwards of the pressure maximum. The velocity u_1 decreases inwards, and so this estimate, $t_{\mathrm{drift}} \approx 7400\,\mathrm{yr}$, is a lower limit. The collisional (i.e., fragmentation) timescale for the large particles, $t_{\mathrm{col}} = (n\sigma\Delta v)^{-1}$, is much shorter in comparison. Here $n = f_m \rho_{\mathrm{d}}/m_{\mathrm{d}}$ is the number density of large particles, f_m is the mass fraction of the large particles ($f_m = 0.75$ in the fragmentation limit, Birnstiel et al. 2012), $\rho_{\mathrm{d}} \approx \Sigma_{\mathrm{d}}\Omega/(\sqrt{2\pi}c_{\mathrm{s}})$ is the midplane mass density of particles, m_{d} is mass of a single particle, σ is the collisional cross section and $\Delta v \approx \sqrt{3\alpha\mathrm{St}}c_{\mathrm{s}}$ is the typical relative velocity between the particles due to turbulence

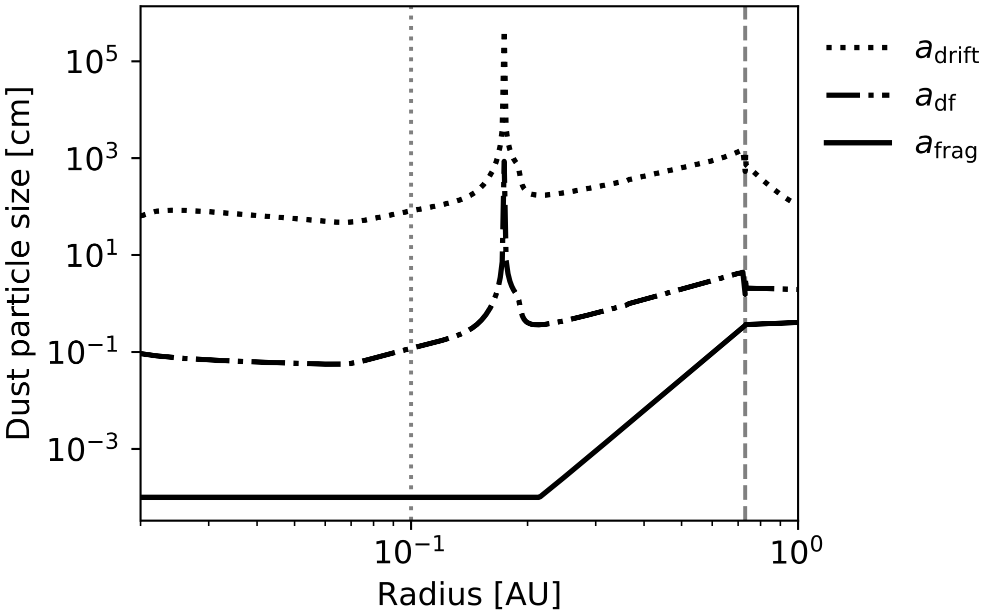


Figure 3.3: Dust particle size limits due to radial drift ($a_{\rm drift}$), drift-fragmentation ($a_{\rm df}$), and turbulent fragmentation ($a_{\rm frag}$) as functions of radius after 0.2 Myr. Location of the local gas pressure maximum due to the MRI is indicated by the vertical dashed line, and location of the dust sublimation line by the vertical dotted line. Spikes in $a_{\rm drift}$ and $a_{\rm df}$ correspond to gas pressure extrema.

(Ormel and Cuzzi 2007a). This yields the collisional timescale of:

$$\begin{aligned} t_{\rm col} &= \sqrt{\frac{8}{27\pi}\frac{\Sigma_{\rm g}}{f_m\Sigma_{\rm d}}}\sqrt{\frac{\rm St}{\alpha}}\frac{1}{\Omega} \\ &\approx 4\left(\frac{f_m}{0.75}\right)^{-1}\left(\frac{\Sigma_{\rm d}}{0.01\Sigma_{\rm g}}\right)^{-1}\left(\frac{\rm St}{10^{-4}}\right)^{1/2}\left(\frac{\alpha}{10^{-4}}\right)^{-1/2} \\ &\times\left(\frac{\Omega}{10\,{\rm yr}^{-1}}\right)^{-1}{\rm yr} \end{aligned} \tag{3.5}$$

where I have expressed the particle size $a_{\rm frag}$ in terms of the Stokes number St. At the pressure maximum $\alpha = 10^{-4}$ and so $t_{\rm col} \approx 0.08\Sigma_{\rm g}/\Sigma_{\rm d}\,{\rm yr} \ll t_{\rm drift}$ for dust-to-gas ratios $\Sigma_{\rm d}/\Sigma_{\rm g} \gtrsim 0.01$. Thus, instead of becoming trapped in the pressure maximum, dust particles fragment and flow inwards.

In the innermost disc turbulent fragmentation yields monomer-sized particles that are entrained with the gas ($\bar{u} \sim u_{\rm gas}$). And so the radial drift inwards is slowed by particles becoming well coupled to the gas. This is also why, in steady state, the dust-to-gas ratio is roughly constant inwards of the pressure maximum.

Finally, despite the pressure maximum not trapping the inflowing particles, the dust-to-gas ratio is enhanced. Because the dust moves with the gas in the innermost disc, the steady state dust-to-gas ratio there is directly proportional to the ratio of dust-to-gas accretion rates at the outer boundary, $\dot{M}_{\rm d}/\dot{M}_{\rm g}$. The initial dust-to-gas ratio everywhere is $\Sigma_{\rm d}/\Sigma_{\rm g} = 10^{-2}$. For this ratio to be constant over time, the ratio of accretion rates would have to be $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 10^{-2}$ at the outer boundary. However, the growth of dust grains in the outer disc, and their radial drift inwards, means that $\dot{M}_{\rm d}/\dot{M}_{\rm g} > 10^{-2}$ at the outer boundary of the inner disc (i.e., dust accretes inwards preferentially compared to gas). Hence, the dust-to-gas ratio in the inner disc in steady-state is $\Sigma_{\rm d}/\Sigma_{\rm g} > 10^{-2}$. In other words, as Fig. 3.2 shows, radial drift of grains from the outer disc leads to an enrichment of solids in the inner disc.

What level of the enrichment is attainable depends on the ratio of dust and gas accretion rates $\dot{M}_{\rm d}/\dot{M}_{\rm g}$, i.e. how quickly the grains drift from the outer disc relative to gas accretion. Since the grain growth in the outer disc is limited by radial drift rather than fragmentation (Birnstiel et al. 2012), high grain drift rates are possible. For example, assuming that $\Sigma_{\rm d}/\Sigma_{\rm g} = 10^{-2}$ in the outer disc, achieving $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 1$ requires the radial drift velocity of grains ($\approx 2\ {\rm St}\ u_{\rm drift}$) to reach $10^2 u_{\rm gas}$. For the standard α-disc model and $\alpha = 10^{-4}$ (extrapolation of the disc model shown in Fig. 3.1) this is satisfied if grains grow to ${\rm St} \sim 10^{-2}$. This corresponds to a particle size an order of magnitude below the radial drift limit ($a_{\rm drift}$) throughout the outer disc. Therefore, the grains easily grow large enough to achieve accretion rates of $\dot{M}_{\rm d}/\dot{M}_{\rm g} \gtrsim 1$. These grains in the outer disc will contain ices which will evaporate as grains drift across the ice lines, towards the inner disc. However, even in the outermost disc silicates account for a considerable portion of the total solid mass (e.g., adopting the abundances of oxygen and carbon in their main molecular carriers from Öberg et al. 2011, 23% of the total oxygen and carbon mass in the outer disc is in silicates and other refractories). Hence, Fig. 3.2 features dust accretion rates up to $\dot{M}_{\rm d}/\dot{M}_{\rm g} = 1$, in which case the dust-to-gas ratio in the inner disc also approaches unity.

3.4. Implications for planetesimal formation

The above results show that the MRI yields a dust-enhanced inner disc, although at the expense of the dust particle size. Furthermore, as there is no trap for the dust particles, the accumulation of dust is limited by the dust inflow rate from the outer disc, and does not increase indefinitely. In this section I explore if further concentration of particles via the streaming instability (SI; Youdin and Goodman 2005) and subsequent gravitational collapse into planetesimals could be the next step towards forming the close-in super-Earths and mini-Neptunes.

The SI can greatly concentrate dust particles if the ratio of dust-to-gas bulk densities is $\rho_{\rm d}/\rho_{\rm g} \gtrsim 1$ (Youdin and Goodman 2005; Johansen and Youdin 2007). This is most likely to be attained in the disc midplane, as dust particles gravitationally settle. The settling is balanced by turbulent stirring. One source of turbulence is the MRI. To reach $\rho_{\rm d}/\rho_{\rm g} \geqslant 1$ in the midplane in the presence of such turbulence, the dust-to-gas surface density ratio $\Sigma_{\rm d}/\Sigma_{\rm g}$ needs to be greater than or equal to $Z_{\rm cr1} = \sqrt{\alpha/({\rm St} + \alpha)}$ (Carrera et al. 2017).

Even in discs that are weakly turbulent or completely laminar, as dust settles the dust-gas interaction leads to turbulence (self-stirring) which can prevent clumping by the SI. In this case, SI can only successfully concentrate dust particles if the dust-to-gas ratio $\Sigma_{\mathrm{d}}/\Sigma_{\mathrm{g}}$ is greater than a critical value Z_{cr2} that depends on the particle Stokes number (Johansen et al. 2009; Carrera et al. 2015). For $\mathrm{St} < 0.1$ (relevant to the simulations shown here) this critical value has been most recently revised by Yang et al. (2017), who find

$$\log Z_{\mathrm{cr2}} = 0.1 \log^2 \mathrm{St} + 0.2 \log \mathrm{St} - 1.76\,. \tag{3.6}$$

Small dust grains that are entrained with the gas do not participate in the SI. Hence, to compare the above results against the SI criteria, I use the dust-to-gas ratio $f_{\mathrm{m}}\Sigma_{\mathrm{d}}/\Sigma_{\mathrm{g}}$ of the large grain population only (where $f_{\mathrm{m}} = 0.75$ is the mass fraction of large particles; Birnstiel et al. 2012).

The top panel of Fig. 3.4 compares this dust-to-gas ratio, for the outer boundary condition $\dot{M}_{\mathrm{d}}/\dot{M}_{\mathrm{g}} = 1$, with the above two criteria for the onset of the SI. This plot shows that in the inner disc, turbulence due to the MRI is generally more prohibitive to dust settling than the turbulence due to dust-gas interactions. Both conditions are fulfilled only near the pressure (and density) maximum.

Provided that the SI successfully concentrates dust particles in the disc midplane, the dust bulk density there may reach up to $100-1000$ times the local gas density (Johansen and Youdin 2007). Gravitational collapse of such particle concentrations will occur if the dust density exceeds the local Roche density (below which the star can tidally disrupt the fragment), $\rho_{\mathrm{Roche}} = 9M_*/4\pi r^3$. Comparison of ρ_{Roche} with the midplane gas densities in the steady-state MRI disc (Fig. 3.4, bottom panel) shows that the condition for gravitational collapse is only satisfied near the pressure (and density) maximum, or at larger orbital distances. Importantly, the bottom panel of Fig. 3.4 shows that the possibility of gravitational collapse of solids at short orbital distances is severely limited by the Roche density.

Overall, given the stellar and disc parameters used here, the SI and the gravitational collapse pathway to planetesimals is viable in the inner disc only in a very narrow region near the pressure maximum. For gas accretion rates larger than the one used here ($\dot{M}_{\mathrm{g}} > 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, that could be expected in the early phase of disc evolution, e.g. Manara et al. 2012) this conclusion will hold, while for sufficiently smaller accretion rates planetesimals would not form in this way anywhere near or inwards of the pressure maximum. Firstly, in both cases dust evolution is expected to yield roughly the same steady-state dust-to-gas ratios given the same ratio of dust and gas accretion rates at the outer boundary. To confirm this conclusion, I repeat the dust evolution calculations for the gas accretion rate of $\dot{M}_{\mathrm{g}} = 10^{-9}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, obtaining results very similar to those above. In addition to similar steady-state dust-to-gas ratios, the particle Stokes number sharply drops inwards of the pressure maximum, and the SI is similarly triggered only around the pressure maximum. This is because the slope of the increase in α inwards of the pressure maximum is roughly the same for different $\dot{M}_{\mathrm{g}}$. Hence the SI criteria is expected to be fulfilled only near the pressure maximum for higher gas accretion rates as well.

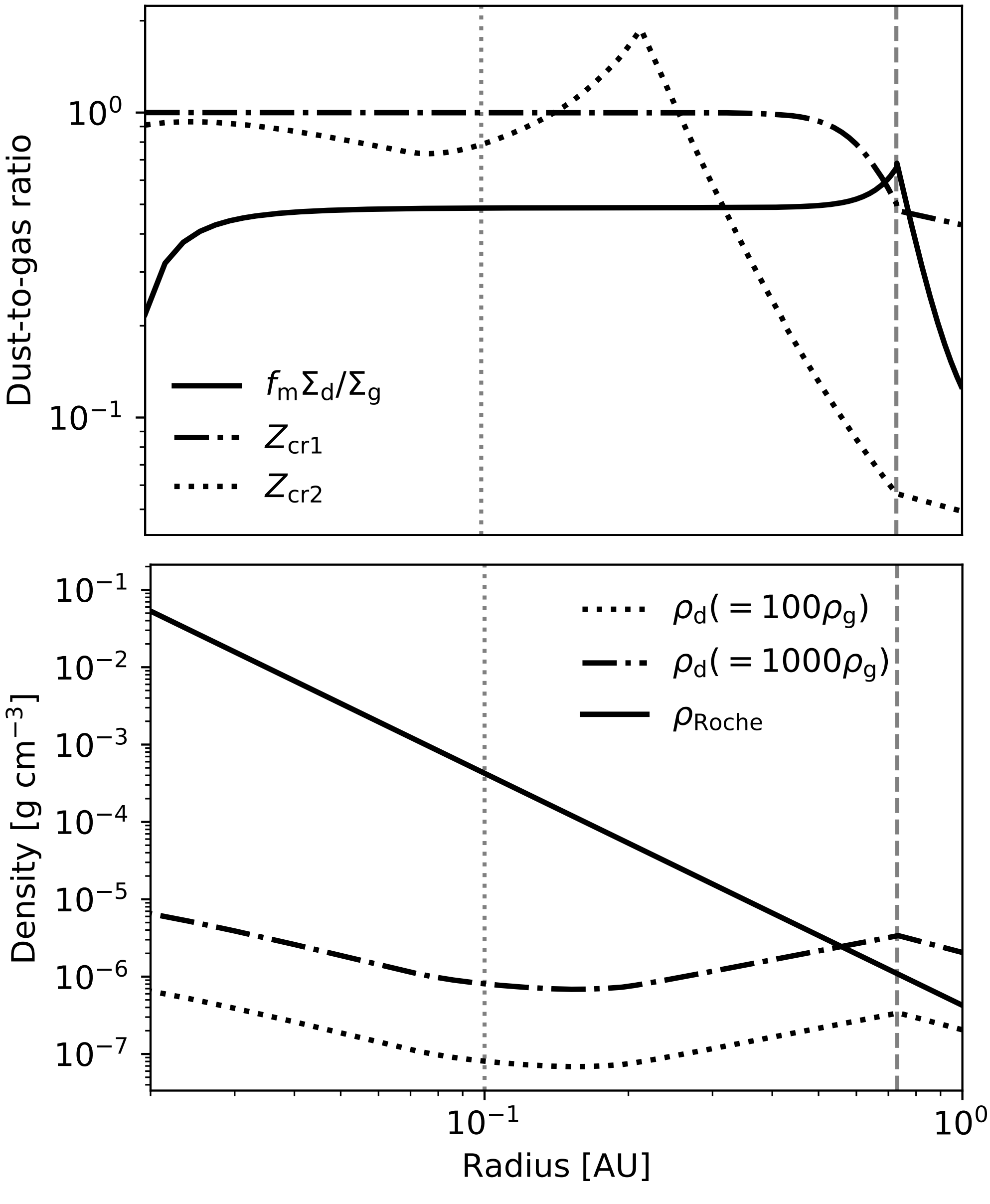


Figure 3.4: Top panel: Dust-to-gas ratio of large dust grains $f_m\Sigma_d/\Sigma_g$ as a function of radius (when $\dot{M}_d/\dot{M}_g = 1$ at the outer boundary; $f_m = 0.75$), compared to the two criteria for the onset of the streaming instabilities.
Bottom panel: Expected peak local dust densities ρ_d if streaming instabilities successfully concentrate particles in the disc midplane as functions of radius, compared to the Roche density ρ_{Roche}. In both panels the vertical dashed line indicates location of the local pressure maximum, and the dotted line indicates the location of the dust sublimation line.

Secondly, even if SI is successfully triggered, to form planetesimals the peak dust density needs to be above the Roche density. The peak dust density scales with the midplane gas density at the pressure maximum, so it scales with the gas accretion rate approximately as $\dot{M}_{\mathrm{g}}^{-1/2}$, and with the radial location of the pressure (and density) maximum as $r_{P_{\mathrm{max}}}^{-1}$. Since the peak dust density is larger than the Roche density at the pressure maximum for $\dot{M}_{\mathrm{g}} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, and the Roche density $\rho_{\mathrm{Roche}} \propto r^{-3}$, for $\dot{M}_{\mathrm{g}} > 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ the gravitational collapse criterion will also be fulfilled near the pressure maximum, while for a sufficiently smaller accretion rate (including $\dot{M}_{\mathrm{g}} = 10^{-9}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$) peak dust density will be too low.

Moreover, it is important to note that, in order for the SI to operate, there must be a relative azimuthal velocity between the dust and gas (e.g. Squire and Hopkins 2018), in addition to the criteria discussed above. At the pressure maximum itself, however, this relative velocity vanishes.

However, the calculations presented here do not include the effects of dust on the gas dynamics and on the MRI. The latter effect in particular may relax some of the constraints on the SI, and this possibility is discussed in Section 3.5.

3.5. Discussion and conclusions

In this chapter I have investigated the earliest phase of planet formation that is the evolution of dust grains in the inner disc that is viscously accreting due to the MRI. The MRI-accreting inner disc features a local gas pressure maximum at the orbital distance of a few tenths of AU. Taking into account the effect of the MRI-induced turbulence on the dust grain size, I find that fragmentation of particles due to turbulent relative velocities limits the particle size to below few millimetres. As a result, the particles are not efficiently accumulated inside the pressure maximum as hypothesised by Chatterjee and Tan (2014); Hu et al. (2018). Regardless of that, as the particles become well coupled to the gas, the radial drift is stopped in the inner disc, and the dust-to-gas ratio is enhanced throughout the inner disc. Thus, the local gas pressure maximum might play a lesser role in the planet formation in the inner disc than previously thought.

The pressure maximum is, however, still the location of a local density maximum in both gas and dust. I explored if the resulting inner disc structure that is enriched in dust could be susceptible to the onset of the streaming instabilities. This pathway to planetesimals seems to be viable only in a narrow region near the pressure (and density) maximum, for the chosen disc parameters.

The gas is not, however, evolved in this work and effects of the growing amounts of dust onto the MRI have not been taken into account. Dust grains lower the gas ionization levels by absorbing free charges and enhancing recombination rates, as ions recombine on the grains (Draine and Sutin 1987; Ilgner and Nelson 2006). Charged grains are not themselves well coupled to the magnetic field as they are too massive, and so their presence promotes the non-ideal MHD effects which can suppress the MRI (Sano et al. 2000; Ilgner and Nelson 2006; Wardle 2007; Salmeron and Wardle 2008; Bai and Goodman 2009). The likely result of taking dust effects into account is thus weakened turbulence, and the change of the disc structure in the longer term. The consequences can only be investigated by modelling both the gas and the dust self-consistently. However, one can sketch out a potential scenario by considering the relevant timescales.

Assuming that the steady-state solution of the gas structure (Fig. 3.1) is reached before dust starts affecting the MRI, we can expect the dust enhancement of the inner disc to ensue. At a certain dust-to-gas ratio the dust will suppress the MRI, and we can expect the levels of turbulence to adapt almost instantly, as the timescale of the magnetic field regeneration is the orbital timescale t_{orb} (e.g. Balbus and Hawley 1991). With the decreasing levels of turbulence, the dust particle size will rapidly grow due to particle coagulation. The growth due to coagulation happens on the timescales of $\Sigma_{\mathrm{g}}/\Sigma_{\mathrm{d}} t_{\mathrm{orb}}$ (e.g. Brauer et al. 2008), so faster than $10^2 t_{\mathrm{orb}}$ if the inner disc is enriched in dust.

Concurrently, due to larger particle size and lower turbulent stirring the particles would vertically settle towards the midplane and radially towards the pressure maximum. Such formation of a ring of solids could potentially trigger formation of larger bodies (as hypothesized by e.g. Chatterjee and Tan 2014). The larger particle size and the settling towards the midplane would likely trigger the streaming instability (inwards of the pressure maximum; see Fig. 3.4). However, it is unclear if this could lead to the formation of planetesimals, as a gravitational collapse is unlikely due to the low bulk dust densities and high Roche density in the inner disc.

The gas surface density will change slowly in comparison to the above processes, on the long viscous timescale, $\sim 10^3 - 10^5 t_{\mathrm{orb}}$ from the pressure minimum to the pressure maximum in the model considered here. The decrease in viscosity would lead to an increase in the gas surface density on this timescale. The gas surface density maximum and the pressure maximum would move inwards. However, if thresholds to form planetesimals are crossed, the small dust grains suppressing the MRI would be removed from the disc. This would increase the viscosity, decrease the gas surface density and move the pressure maximum outwards. Whether these processes are balanced in another kind of a steady state, or the behaviour of the inner disc is dynamic and quasi-periodic, must be investigated through self-consistent modelling of dust, gas and the MRI.

At high dust-to-gas ratios the dust also becomes dynamically important, and affects the gas disc structure through the drag backreaction (Nakagawa et al. 1986). The gas rotation profile is then driven towards Keplerian, and as a result the radial gas pressure profile flattens. This, in turn, slows down the radial drift of dust particles. If dust already piles up in the inner disc due to radial drift being slower than in the outer disc, the dust backreaction amplifies the effect (Drazkowska et al. 2016). In this work, dust enhancement is driven by the dust grains already being completely coupled to the gas in the innermost disc, and thus the effect of dust backreaction would be limited. However, the backreaction would become important if the dust grains grow (e.g. due to the suppression of the MRI-induced turbulence discussed above), especially near the pressure maximum. If the dust grains grow in the innermost disc where the pressure gradient is negative, the backreaction would slow down the loss of dust to the star. However, the backreaction would also limit the concentration of dust that can be achieved at the pressure maximum, since it acts to flatten the overall gas pressure profile (Taki et al. 2016).

In summary, the results presented here suggest that the inner disc might support formation of the super-Earths. However, the self-consistent evolution of the dust and the gas needs to be studied.

4 ATMOSPHERES OF PLANETS FORMED IN THE INNER DISC

4.1. Introduction

It was shown in Chapter 3 that planetesimal formation through the streaming instability and gravitational collapse is challenging despite the dust enhancement in the inner disc. Although these conclusions could change when the feedback of dust onto gas is taken into account, it is not presently clear how exactly planetesimals or cores would arise in the inner disc, and it is not possible to predict properties of solid cores formed in the inner disc. Nevertheless, orbital distances, radii and masses of many close-in planets have been well determined observationally. Thus, in this chapter I use observational results to perform a separate test of super-Earth formation in the inner disc by considering accretion of planetary atmospheres and their subsequent evolution.

Recall, from Chapter 1, that the atmospheres of many of the close-in super-Earths and mini-Neptunes are H/He dominated (e.g. Jontof-Hutter et al. 2016) and they typically make up $0.1-10\%$ of their total mass (Lopez and Fortney 2014; Wolfgang and Lopez 2015). Thus, they are considerably more massive than the atmospheres of the planets in the inner Solar system. Outgassing of hydrogen from a rocky core is not sufficient to explain the majority of these atmospheres (Rogers et al. 2011). Thus these atmospheres are most likely composed of gas accreted from the protoplanetary disc after the formation of a solid core. If so, these atmospheres are presumably formed steadily through core accretion. In core accretion, the gaseous envelope is connected to the disc. It remains in a quasi-hydrostatic equilibrium and contracts as it cools, allowing more gas to be accreted (e.g. Rafikov 2006; Lee et al. 2014; Lee and Chiang 2015).

Lee et al. (2014) (see also Lee and Chiang 2015; 2016; Lee et al. 2017) argue that accretion onto a super-Earth-sized solid core is an efficient process, and that there is a theoretical problem of how to stop the super-Earth cores from undergoing runaway accretion and becoming gas giants (Mizuno 1980). A solution was proposed in Lee and Chiang (2016), where they suggest that the final assembly of super-Earth cores only occurs towards the end of the protoplanetary disc lifetime, so that their atmospheres are accreted from the gas-poor "transition discs", i.e., during disc dispersal.

On the other hand, the inner disc models discussed in Chapter 2 predict in a steady state gas surface densities that are considerably lower than those of the minimum mass solar nebula at short orbital periods. If the super-Earths had to have formed in a gas-poor inner disc, perhaps they have formed in the gas-poor inner disc arising due to steady-state MRI accretion, a possible alternative to the Lee and Chiang (2016) proposal. Therefore, I use the atmospheric accretion model of Lee and Chiang (2015) to calculate the atmospheres accreted onto super-Earth-sized cores in the MRI-accreting inner disc.

Additionally, after disc dispersal, super-Earths are subject to atmospheric mass loss due to photoevaporation by high-energy stellar irradiation. It has been shown theoretically (e.g. Owen and Wu 2013; Lopez and Fortney 2013) and observationally (e.g. Lundkvist et al. 2016; Fulton et al. 2017; Fulton and Petigura 2018) that this mass loss can significantly change the planet properties. Therefore, in this study I take this mass loss into account. Finally, I evolve the planets in time, to their present day ages, and compare the resulting planet properties to the observed ones.

4.2. Methods

4.2.1. Accretion of planetary atmospheres

In this work, it is assumed that solid super-Earth-sized planet cores accrete their gaseous envelopes in a gas-poor inner disc that is viscously accreting due to the MRI. The model of the disc structure is the same one as in Chapter 3, and it is assumed that the disc structure does not evolve in time. The gaseous envelope is assumed to be in quasi-hydrostatic equilibrium, contracting as it cools, allowing more gas to be accreted. Then, the accreted envelope mass fraction (ratio of envelope mass and core mass) after time t can be estimated by the scaling relations from Lee and Chiang (2015) (with an additional factor that accounts for varying gas surface density from Lee et al. 2017; Fung and Lee 2018):

$$X(t) = 0.07 \left(\frac{t}{1\,\mathrm{Myr}}\right)^{0.4} \left(\frac{0.02}{Z}\right)^{0.4} \left(\frac{\mu}{2.37}\right)^{3.4} \left(\frac{M_{\mathrm{core}}}{5\mathrm{M}_{\oplus}}\right)^{1.7} \left(\frac{f_{\Sigma}}{0.1}\right)^{0.12} \tag{4.1}$$

for dusty atmospheres, and

$$X(t) = 0.18 \left(\frac{t}{1\,\mathrm{Myr}}\right)^{0.4} \left(\frac{0.02}{Z}\right)^{0.4} \left(\frac{\mu}{2.37}\right)^{3.3} \left(\frac{M_{\mathrm{core}}}{5\mathrm{M}_{\oplus}}\right)^{1.6} \left(\frac{1600\,\mathrm{K}}{T_{\mathrm{rcb}}}\right)^{1.9} \left(\frac{f_{\Sigma}}{0.1}\right)^{0.12} \tag{4.2}$$

for dust-free atmospheres. Here Z is the metallicity of the atmosphere; $\mu = 1/(0.5W + 0.25Y + 0.06Z)$ is the mean molecular weight, with $W = (1 - Z)/1.4$, $Y = 0.4(1 - Z)/1.4$; T_{rcb} is the temperature at the radiative-convective boundary inside the atmosphere; $f_{\Sigma} = \Sigma_{\mathrm{g}}/\Sigma_{\mathrm{MMSN}}$ is the ratio of the gas surface density (Σ_{g} from the adopted inner disc model, Fig. 3.1) and the gas surface density profile of the minimum mass solar nebula ($\Sigma_{\rm MMSN} = 1700(d/1\,{\rm AU})^{-3/2}\,{\rm g\,cm^{-2}}$, where d is the orbital radius, Hayashi 1981). Furthermore, it is assumed that the gas adiabatic index is $\gamma = 1.2$, and that in dusty atmospheres $T_{\rm rcb} = 2500\,{\rm K}$ which arises from the disassociation of Hydrogen (see Lee and Chiang (2015), their section 2.1).

I use the expressions (4.1,4.2) to calculate how much gas a planet accretes in 1 Myr as a function of core mass, for various metallicities Z and gas surface density factors f_Σ in the case of dusty atmospheres (with $T_{\rm rcb} = 2500\,{\rm K}$), and various Z, f_Σ and $T_{\rm rcb}$ in the case of dust-free atmospheres.

4.2.2. Photoevaporation of planetary atmospheres

These accreted atmospheres are subject to photoevaporation following disc dispersal. I use a simplified estimate of how the photoevaporation changes the accreted atmospheres. First, for a given planet core mass $M_{\rm core}$ and (accreted) envelope mass fraction X I find the photospheric radius of the planet $R_{\rm p}$. For this I use a simple model of an atmosphere at hydrostatic equilibrium (Owen and Wu 2017), in which the solid core is surrounded by an adiabatic convective envelope, on top of which is an isothermal radiative photosphere.

Next, the mass-loss timescale due to high-energy stellar irradiation is (Owen and Wu 2017)

$$t_{\dot{X}} = \frac{4\pi d^2 G M_{\rm core}^2 X(1+X)}{\eta\pi L_{\rm HE}}\frac{1}{R_{\rm p}^3}\,, \tag{4.3}$$

where d is the orbital radius of the planet and $L_{\rm HE}$ is the stellar high-energy flux. I consider a Sun-like star, as in the adopted disc model.

To determine the final envelope mass fractions, I do not explicitly evolve the atmospheres in time. Instead, one can use the fact that most of the mass loss happens during the first $\sim 100\,{\rm Myr}$ after disc dispersal, since during this period the stellar high-energy flux $L_{\rm HE}$ is saturated ($L_{\rm HE} = L_{\rm sat} \sim 10^{-3.5} L_\odot$ for a Sun-like star) and after this time it quickly decays (Jackson et al. 2012; Tu et al. 2015).

Thus, one can estimate the final envelope mass fraction as follows. If a planet's mass-loss timescale $t_{\dot{X}}$ is longer than 100 Myr at the time of disc dispersal, the planet does not suffer significant mass loss. In this case it can be assumed that such a planet remains unchanged by the photoevaporation.

On the other hand, a planet with $t_{\dot{X}} < 100\,{\rm Myr}$ loses mass. For a given core mass $M_{\rm core}$ and orbital distance d, the mass-loss timescale as a function of the envelope mass fraction, $t_{\dot{X}}(X)$, peaks at $X \equiv X_{\rm peak}$ of a few percent, decreasing for both smaller and larger X (Owen and Wu 2017). Thus, for a planet with a small accreted envelope mass fraction ($X < X_{\rm peak}$), the mass loss further shortens the loss timescale. If such a planet's initial mass-loss timescale is less than 100 Myr it is subject to runaway mass loss, and it can be assumed that it is completely stripped of its atmosphere. For a planet with a large accreted atmosphere ($X > X_{\rm peak}$), the mass-loss timescale increases towards the peak value $t_{\dot{X}}(X_{\mathrm{peak}})$ as the planet loses mass. If M_{core} and d are such that the $t_{\dot{X}}(X_{\mathrm{peak}}) \geqslant 100\,\mathrm{Myr}$, one can assume that such a planet will end up with an envelope mass fraction X corresponding to a mass-loss timescale of $t_{\dot{X}} = 100\,\mathrm{Myr}$. However, if $t_{\dot{X}}(X_{\mathrm{peak}}) < 100\,\mathrm{Myr}$, the mass-loss timescale at first increases, and then decreases, entering the runaway regime. In this case a planet loses its entire atmosphere despite an initially large accreted X. This simple prescription adequately captures the basic physics of atmospheric photoevaporation (see far left panel of fig. 6 in Owen and Wu 2017).

4.3. Results

4.3.1. Accretion of planetary atmospheres

Using the scaling relations (4.1,4.2) I calculate the envelope mass fractions that planetary cores of various masses accrete from the gas disc in 1 Myr. Results are shown in Fig. 4.1 for both dusty and dust-free atmospheres of various metallicities, ranging from solar ($Z = 0.02$) to the metallicity of Neptune's atmosphere ($Z = 0.2$; Karkoschka and Tomasko 2011), gas surface density factors $f_{\Sigma} = 10^{-4} - 1.88$, and radiative-convective boundary temperatures $T_{\mathrm{rcb}} = 1600 - 2500\,\mathrm{K}$.

The dependence on metallicity Z is non-monotonous, as the accreted envelope mass fraction depends separately on the metallicity and on the mean molecular weight (which is set by the metallicity). The smallest accreted atmospheres, with the rest of the parameters fixed, have $Z \sim 0.1$.

The radiative-convective boundary temperature T_{rcb} is expected to be roughly constant in dusty atmospheres, so I only explore the effect of this parameter in dust-free atmospheres. In the latter, T_{rcb} is related to the temperature of the environment. Additionally, accretion of both dust-free and dusty atmospheres depends the density of the environment. Here we are interested in atmospheres that are accreted in the inner disc, near or inwards of the pressure maximum. The location of the pressure maximum is determined by the extent of thermal ionization of potassium in the disc model used here, and so this corresponds to disc temperatures of $T \gtrsim 1000\,\mathrm{K}$, regardless of the exact disc parameters (e.g. gas accretion rate). For a disc temperature of $T \sim 1000\,\mathrm{K}$, numerical models of the accreting atmospheres give $T_{\mathrm{rcb}} \sim 1600\,\mathrm{K}$ (Lee and Chiang 2015), which thus sets a lower bound on T_{rcb} for the calculations in this work. Moreover, the location of the pressure maximum is also where the gas surface density is highest (see Fig. 3.1). Colder atmospheres in a more dense environment accrete more. So, to show the maximum accreted atmospheres in an MRI-accreting disc, I plot a set of dusty and dust-free atmospheres (of various metallicities Z) for the maximum $f_{\Sigma} = 1.88$, and the minimum $T_{\mathrm{rcb}} = 1600\,\mathrm{K}$ (the latter refers only to the dust-free atmospheres). Conversely, the maximum temperature at which equation (4.2) is valid (due to the limitations of the opacity tables used

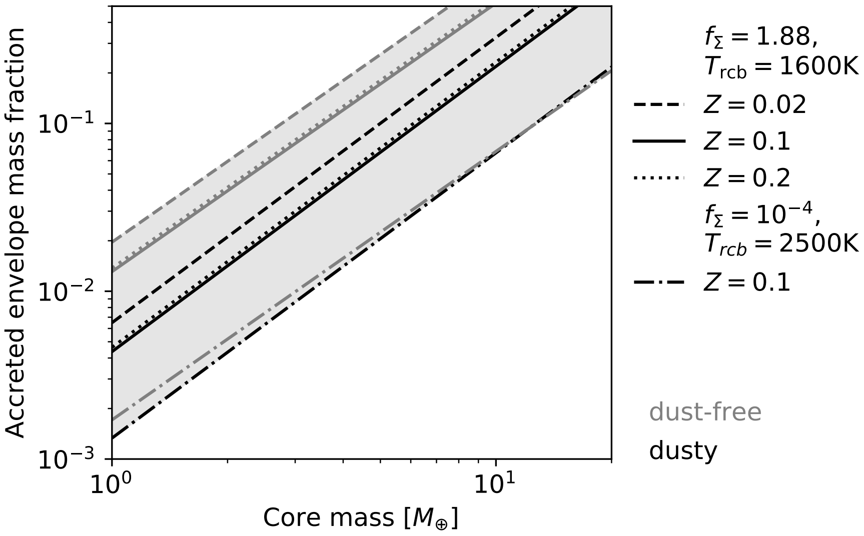


Figure 4.1: Envelope mass fraction of atmospheres accreted in 1 Myr as a function planet core mass for dusty (black lines) and dust-free (grey lines) atmospheres and a variety of metallicities Z, gas surface density factors f_Σ and (in the case of dust-free atmospheres) radiative-convective boundary temperatures T_rcb, as indicated in plot legend. Grey region indicates the total range of expected envelope mass fractions (except for those that would reach an envelope mass fraction of $X = 0.5$ within 1 Myr and thereby expected to undergo runaway accretion to form gas giants; these are not shown).

by Lee et al. 2014) is $T_\mathrm{rcb} = 2500\,\mathrm{K}$ [1] , and the minimum gas surface density in the inner disc model with respect to the minimum mass solar nebula is $f_\Sigma \approx 10^{-4}$ (corresponding to the inner disc edge in Fig. 3.1). Hotter atmospheres in lower-density environments accrete less, and so to show the smallest accreted dust-free atmospheres, I plot the $f_\Sigma = 10^{-4}$ atmospheres, with $T_\mathrm{rcb} = 2500\,\mathrm{K}$ for the dust-free atmospheres, and with metallicity $Z = 0.1$ (since, as noted above, $Z \sim 0.1$ yields the smallest atmosphere for any given set of other parameters).

Finally, atmospheres that grow above a threshold of $X = 0.5$ undergo runaway accretion and end up as gas giants (Rafikov 2006). The scaling relations (4.1,4.2) are not applicable in this case. Therefore, Fig. 4.1 is cut off at $X = 0.5$, and the grey region indicates how small or large super-Earth/mini-Neptune atmospheres may be at the time of disc dispersal. Overall one can see that, if the cores are formed 1 Myr before the dispersal, runaway accretion is avoided for the majority of relevant core masses, but they do accrete significant gaseous envelopes of up to a few × 10% of core mass.

The envelopes shown in Fig. 4.1 have been calculated assuming that the accretion lasts for 1 Myr. Since disc lifetimes can be longer (Mamajek 2009), these envelopes could be conservative estimates if planets form sooner than 1 Myr before disc dispersal. If, for example, the envelopes are accreted for 5 Myr, the envelope mass will double. It is not expected for the results to be very sensitive to the exact disc parameters, as long as the cores accrete their atmospheres in a thermally-ionized MRI-active inner disc. Note, however, that the disc model used here implies that the extent of such inner disc does not encompass all observed sub-Neptunes for all relevant accretion rates; e.g. the gas pressure maximum is at an orbital period longer than 100 days only for gas accretion rates of $\dot{M}_{\mathrm{g}} \gtrsim 3 \times 10^{-9}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$. Thus, a planet with a longer orbital period might spend at least some time in a colder MRI-dead zone, which is not taken into account in this work.

---

[1] Note that, assuming T_rcb is directly proportional to the disc temperature T, and scaling from the numerical models' result that $T_\mathrm{rcb} \sim 1600\,\mathrm{K}$ corresponds to $T \sim 1000\,\mathrm{K}$, yields $T \sim 1500\,\mathrm{K}$ for $T_\mathrm{rcb} \sim 2500\,\mathrm{K}$. The disc temperature in the model used here only exceeds 1500 K at radii <0.1 AU, so a maximum T_rcb of 2500 K is indeed roughly valid over most of the inner disc.

4.3.2. Photoevaporation of planetary atmospheres

To further check the consistency of core accretion in the MRI-accreting inner disc with observations, we need to consider whether these accreted atmospheres survive photoevaporation. I calculate the final (remaining) envelope mass fraction of the minimum and maximum possible accreted atmospheres (corresponding respectively to dusty atmospheres with $f_{\Sigma} = 10^{-4}$ and $Z = 0.1$, and dust-free atmospheres with $Z = 0.02$, $f_{\Sigma} = 1.88$ and $T_{\mathrm{rcb}} = 1600\,\mathrm{K}$) for each core mass and as a function of orbital period. Results are shown in Fig. 4.2. In the case of the maximum accreted atmospheres (top panel), the atmospheres would undergo runaway accretion for core masses $\gtrsim 8\,\mathrm{M}_{\oplus}$ (indicated by the hatched region), which are thus excluded here.

The figures show that the orbital period at which the atmosphere can be completely evaporated decreases with increasing core mass, and cores that retain their atmospheres generally evolve towards a 1% envelope mass fraction as expected from theory. At 100 days the atmospheres are unaffected by photoevaporation, and at periods shorter than 1 day all planets are predicted to end up as bare cores. Massive cores are predicted to keep their $1-50\%$ atmospheres at the majority of orbital periods, and planets with Earth-mass cores are safe from complete mass loss at periods larger than 50 days.

Note that here the orbital period determines the level of high-energy flux that planet experiences and planet equilibrium temperature (and thus planet radius), but does not directly reflect variations in temperature and density of the protoplanetary disc inside which the atmospheres were accreted. As discussed above, the effect of the disc temperature on the accreted envelope mass fraction is negligible for dusty atmospheres. For dust-free atmospheres the dependence is monotonous and the extent of the effect is explored by considering the minimum T_{rcb} expected in the inner disc, and the maximum T_{rcb} for which the scaling relations (4.1,4.2) are valid. Similarly, the dependence on the disc density is explored by considering the smallest and largest values of the ratio of the MRI-disc model and the minimum mass solar nebula surface densities. Thus, by calculating the effect of photoevaporation on both the minimum and maximum accreted atmospheres shown in Fig. 4.1 for each core mass, I encompass the possible range of disc densities and temperatures.

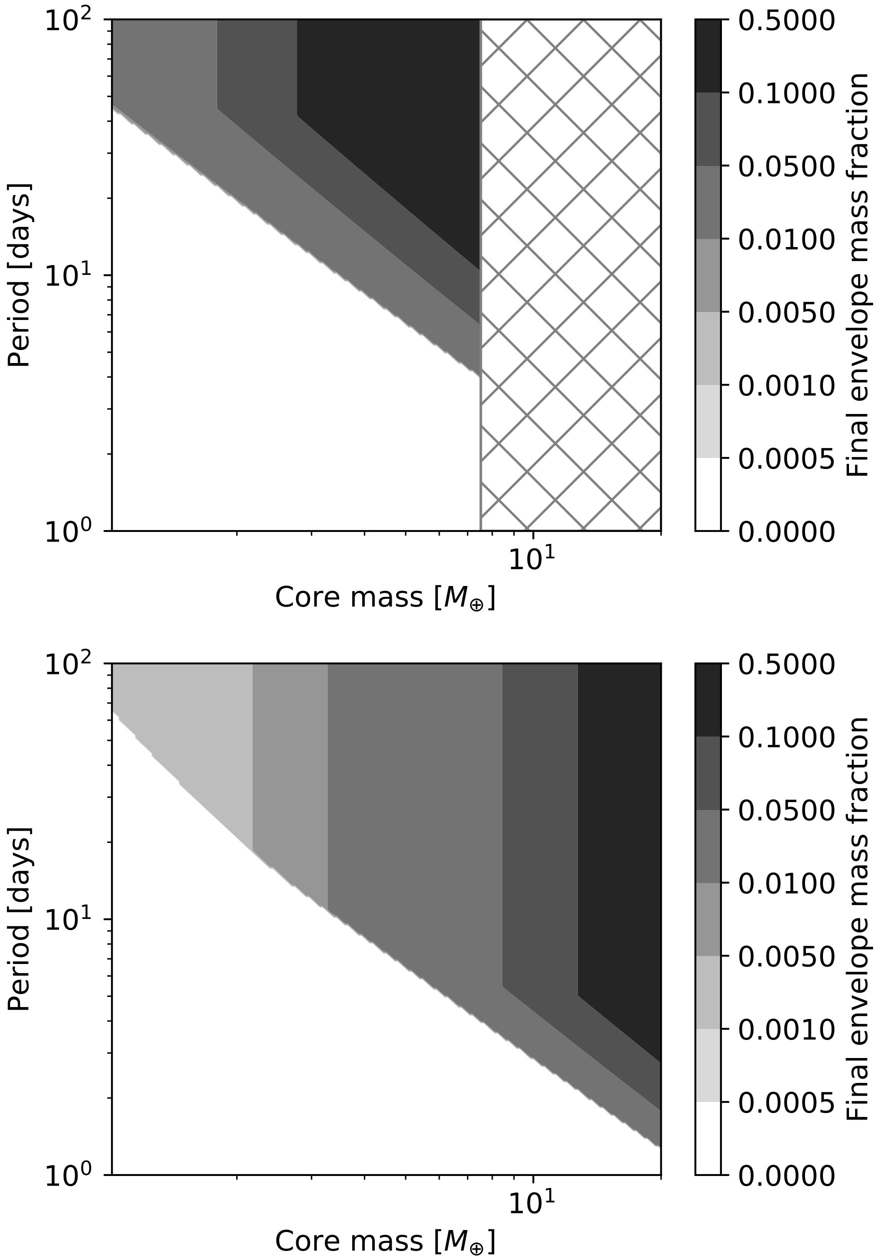


Figure 4.2: Maximum (top) and minimum (bottom) envelope mass fraction of the atmospheres after accounting for photoevaporation, as functions of planet core mass and orbital period. In the top figure, the hatched region indicates the core masses for which the planets would undergo runaway accretion and are thus excluded from here.

4.4. Comparison to observations

Overall, Fig. 4.2 shows that the envelopes formed in a gas-poor inner disc due to the MRI survive photoevaporation for a large range of orbital periods, and the low gas surface densities are not a hindrance to the formation of mini-Neptunes. On the contrary, the final envelope mass fractions of the planets that do keep their atmospheres are typically overestimated. The planets with core mass larger than $2\,\mathrm{M}_\oplus$ are predicted to either have a $> 1\%$ atmosphere or to be completely evaporated. On the other hand, from the observations, the typical envelope mass fraction of mini-Neptunes that hold onto their atmospheres is 1% (Wolfgang and Lopez 2015). To look into this further, I compare the predictions of these calculations against the observed mass-radius relationship for sub-Neptune planets in Fig. 4.3, and against measured masses and radii of individual sub-Neptune planets in Fig. 4.4.

For the observations in Fig. 4.3, I show the probabilistic best-fitting mass-radius relationship of Wolfgang et al. (2016): a power law $M/\mathrm{M}_\oplus = 2.7(R/\mathrm{R}_\oplus)^{1.3}$ (indicated by the dashed line) with a standard deviation of $\pm 1.9\,\mathrm{M}_\oplus$ due to an intrinsic scatter in planet mass (the dark grey region), and an upper limit constraint on the planet density corresponding to a mass-radius relationship for solid cores of Earth-like composition $M/\mathrm{M}_\oplus = (R/\mathrm{R}_\oplus)^4$ (dotted line; Valencia et al. 2010). Additionally, the above mass-radius relationship does not capture a significant feature of the observed radius distribution of sub-Neptunes, a decrease in occurrence rates of planets with radii of $1.5-2\,\mathrm{R}_\oplus$ (indicated here by the sheer grey region; Fulton et al. 2017).

To show the predictions of the above atmospheric calculations in the mass-radius plane, I take the calculated envelope mass fraction as a function of core mass and period and re-calculate the planet radius at the planet age of 5 Gyr as a function of core mass and period, using the same simple atmospheric evolution model of Owen and Wu (2017). The results are shown for the minimum and maximum accreted atmospheres (Fig. 4.2, and excluding the completely evaporated planets) in Fig. 4.3 (the medium grey and the light grey region respectively). Note that the light grey region has a cut-off at about $8\,\mathrm{M}_\oplus$ because the massive cores that, given the parameters of the maximum accreted atmospheres, would be subject to runaway accretion, are excluded. The solid-line contours show how the planet mass and radii change as a function of period for the minimum accreted atmospheres. At the orbital period of 100 days the planets are largely unaffected by the atmospheric loss, and closer to the star the photoevaporation removes atmospheres of the lower-mass planets entirely. For the planets that keep their atmospheres at large periods a decrease in period means little to no change in planet mass. Consequently, for these planets a decrease in period results in an increase in planet radius as atmospheres are hotter and more expanded closer to the star due to stronger stellar irradiation. At small periods the atmospheric loss is significant for all planets, and the trend is reversed.

It is clear from Fig. 4.3 that for planets with radii $R \lesssim 2.3\,\mathrm{R}_\oplus$ the core accretion of atmospheres in the inner disc predicts larger planet radii than those observed, due to the overestimated envelope mass fractions. The predicted atmospheres fill in the range of planet radii of

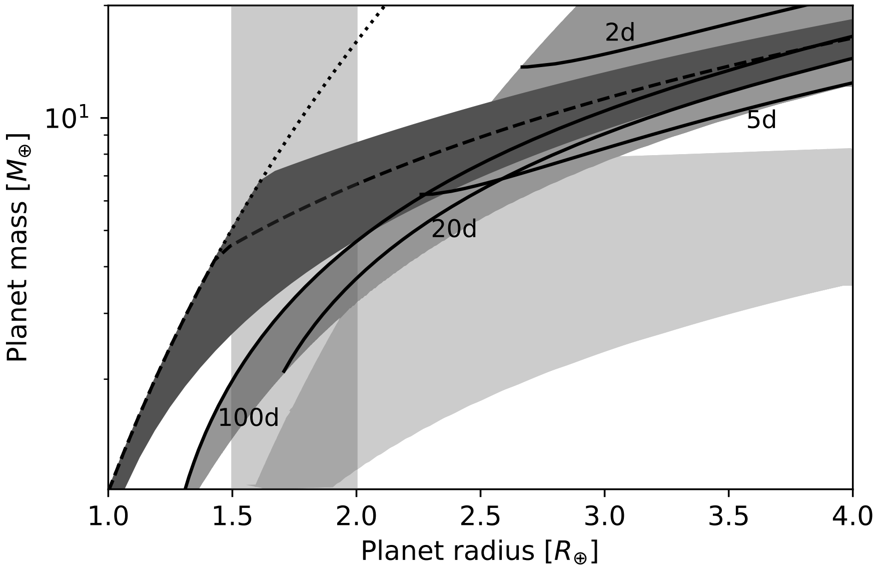


Figure 4.3: Mass-radius relationship for sub-Neptune planets: Earth-like composition solid cores (dotted line), probabilistic fit to observations mean value (dashed line) and scatter (solid dark grey region) (Wolfgang et al. 2016), region of low planet occurrence rates from the observed radius distribution of planets (sheer grey region) (Fulton et al. 2017), and predictions from the minimum and maximum accreted atmospheres and photoevaporation (medium grey and light grey region respectively) with orbital period contours for the minimum accreted atmospheres (solid-line contours).

$1.5-2\,\mathrm{R}_\oplus$, at which there is an observed decrease in planet occurrence rates (sheer grey region). For planets with $R \gtrsim 2.3\,\mathrm{R}_\oplus$ there is a region in which the observed (dark grey) and the predicted (medium and light grey) mass-radius relationships overlap. This overlap corresponds to the (minimum accreted) predicted atmospheres for orbital periods between 20 and 100 days, and a narrow range of short orbital periods ($2-5$ days). Notably, even for the minimum accreted atmospheres, the planet radii, at fixed planet mass, are smaller than those observed only for significant high-energy fluxes at orbital periods of less than about 2 days. Taking into account the full range of accreted atmospheres (up to the maximum accreted atmospheres shown in light grey) further suggests that the predicted atmospheres are typically larger than the atmospheres of the observed sub-Neptunes.

I further compare the predictions of the above calculations to sub-Neptune planets with measured masses and radii (taken from Wolfgang et al. 2016, excluding the planets where only the upper limit on the mass was known). The observed and the predicted radii and masses are shown in Fig. 4.4 in four panels corresponding to four orbital period bins. As in Fig. 4.3, the medium and light grey regions correspond to the predictions from the minimum and maximum atmosphere mass models respectively. To facilitate comparison against the planets that are bare solid cores in each period bin, Fig. 4.4 also shows the core masses that are predicted to lose their entire atmospheres in a given period bin (grey lines shown below the dotted lines that represent the Earth-like composition mass-radius relationship). Fig. 4.4 shows that the

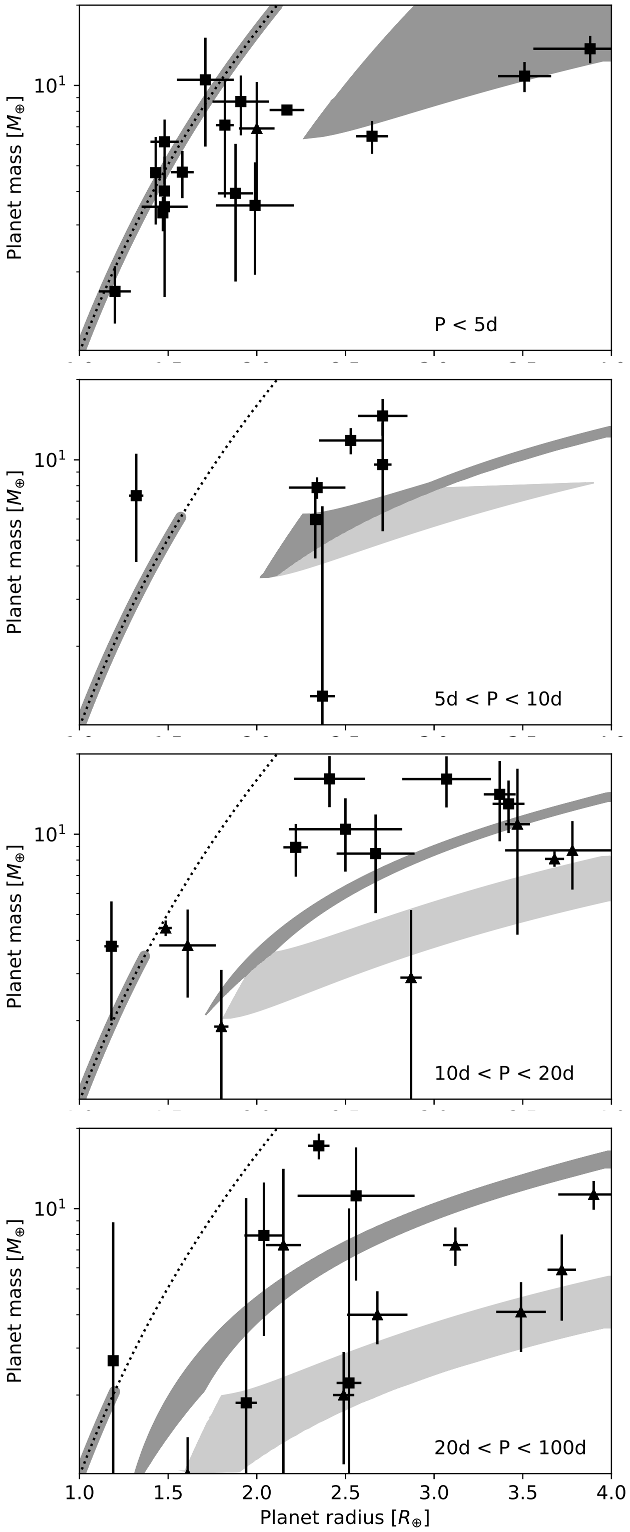


Figure 4.4: Sub-Neptune planets with measured masses and radii (square markers if the mass was determined using the radial velocity method, and triangles if the mass was determined using transit-timing variations), with uncertainties as listed in Wolfgang et al. (2016), Earth-like composition solid cores (dotted line), and predictions from the minimum and maximum atmosphere models (medium and light grey region respectively), in period bins as indicated in plot labels. This figure indicates that while the period range at which planets can be stripped by photoevaporation is consistent with the data, the planets typically have larger H/He envelopes that expected.

masses of the predicted bare cores and the period at which photoevaporation can strip them are largely consistent with those observed. That is, there are no observed planets consistent with the Earth-like composition that are (significantly) more massive than the largest core that the photoevaporation can strip (the upper limit of the grey line, the predicted bare cores) in each period bin. Fig. 4.4 also explicitly demonstrates that for the planets that maintain their atmospheres against the photoevaporation, the predicted planet radii are consistent with or larger than the those observed for the majority of the planets. At long orbital periods (20 – 100 days) all planets except one are consistent, within the observational uncertainties, with the predictions (the minimum accreted atmospheres in medium grey, the maximum accreted atmospheres in light grey, and the region in between). At intermediate periods (5 – 20 days), about a third of planets that are not bare cores have radii smaller than the predicted radii at the same mass. Finally, at short periods of less than 5 days, there are noticeably 5 planets with radii of $\sim 1.8 - 2\,\mathrm{R}_\oplus$ that are neither consistent with the mass-radius relationship of rocky cores, nor with the presence H/He envelopes. This suggests, potentially, that the cores of these planets could contain significant amounts of ice. Still, the majority of short-period planets are consistent with the predictions. Additionally, while there might be exceptions, the radius distribution of sub-Neptunes is consistent with cores being largely rocky (Owen and Wu 2017). Therefore, these results confirm the inference that typically the planets accrete too much gas in the calculations.

Overall, the atmospheres accreted in the inner disc are typically in agreement with or larger than those observed, with the exception of planets with significant high-energy fluxes within a very narrow range. This is because core accretion is so efficient that considerable atmospheres can be accreted in the hot and low-density MRI-accreting inner disc and also maintained against photoevaporation.

4.5. Discussion and conclusions

In this chapter I have examined accretion of planetary atmospheres formed in the inner disc that is viscously accreting due to the MRI. I have considered both the accretion and subsequent photoevaporation of close-in super-Earth planets, and compared the results to the observed properties of these planets.

If super-Earth and mini-Neptune cores indeed form in the inner disc, would the inferred low gas surface densities due to the MRI allow them to acquire the observed $0.1 - 10\%$ envelope mass fractions? It is found that they would. In fact, even after accounting for atmospheric evaporation, the calculated atmospheres tend to overestimate the observed ones.

Could the atmospheric accretion in the MRI-implied disc and the observations be brought into agreement, without invoking an assumption that cores form just before the beginning of disc dispersal (e.g. Ikoma and Hori 2012; Lee and Chiang 2016)? The calculations shown here do not include several effects which could contribute.

First of all, for core masses smaller than $10\,\mathrm{M}_\oplus$, the discrepancy could be explained by the "boil-off" or core powered mass-loss (Owen and Wu 2016; Ginzburg et al. 2018; see also Ikoma and Hori 2012), a process in which a planet atmosphere that had not cooled and contracted before the disc dispersal loses its mass. Upon the dispersal the stellar continuum radiation illuminates the planet and launches a Parker wind. The mass loss causes rapid contraction of the atmosphere, and the contraction in turn shuts off the mass loss. Planets that start out with few 10s of percent atmospheres, may be left with 1% after the boil-off. This process precedes the mass loss caused by the stellar high-energy flux considered above, and can operate at larger distances from the star.

Secondly, the scaling relations used to calculate the accreted atmospheres are derived assuming no sources of heating due to planetesimal accretion, or due to heat deposited in the hypothesized final stage of giant mergers of planetary embryos. The latter could be released for several kyr (e.g. Inamdar and Schlichting 2015), lowering the cooling rate of the atmosphere, and thus allowing less gas to be accreted. Furthermore, the scaling relations assume that the gas inside the planet's Hill sphere is bound and static. Three-dimensional numerical simulations suggest instead that disc material is recycled between the envelope and the disc (e.g. Ormel et al. 2015; Fung et al. 2015; Cimerman et al. 2017).

Finally, if the giant mergers happen between planets, after the disc has fully dispersed, they would likely result in significant atmospheric mass loss. Head-on collisions between Earth/super-Earth-sized planets with few-percent atmospheres can remove tens of percent of the total atmospheric mass (Liu et al. 2015; Inamdar and Schlichting 2016).

Nevertheless, to avoid the runaway accretion for more massive cores, the low gas surface densities the MRI provides are favourable compared to the MMSN environment. Furthermore, gas-poor conditions in this case are provided in a long-lived state, and not in a transient phase (e.g. a transition disc, as proposed by Lee and Chiang 2016). Therefore, the results presented here support the hypothesis that the super-Earths could have formed in the inner protoplanetary disc.

5 Improved model of the MRI-accreting inner disc

5.1. Introduction

In Chapter 2 I discussed models of the inner disc in which the disc structure and accretion due to the MRI were considered self-consistently. The obtained structure of the inner disc features a local gas pressure maximum, in line with theoretical expectations, and the location of the pressure maximum in these models was found to be consistent with the orbital distances of the close-in super-Earths. However, these models included a number of simplifications about the disc physical and chemical structure. The disc vertical structure was considered to be vertically-isothermal, and heating by stellar irradiation was neglected. However, disc temperature may vary vertically, and the temperature profile is particularly non-trivial when both accretion heating and stellar irradiation are considered (e.g. D'Alessio et al. 1998). The opacity of the disc due to dust grains was also taken to be a constant, whereas in reality it varies with temperature and the properties and the abundance of dust grains.

Furthermore, in Chapter 2 it was assumed that the only source of ionization in the disc is thermal (collisional) ionization of potassium, and that the potassium ions and free electrons may only recombine in the gas-phase. Potassium is a good representative of thermally-ionized species in the inner disc based on its low ionization potential and its abundance (Desch and Turner 2015). However, ions and free electrons also collide with dust grains present in the disc. When they are adsorbed onto grain surfaces, they quickly recombine. The presence of dust grains is thus known to reduce the ionization levels in discs and suppress the MRI (Sano et al. 2000; Ilgner and Nelson 2006; Wardle 2007; Salmeron and Wardle 2008; Bai and Goodman 2009; Mohanty et al. 2013). Importantly, dust grains heated to high temperatures ($\gtrsim 500\,\mathrm{K}$) can also emit electrons and ions into the gas phase. For electrons, the process is known as thermionic emission. Such temperatures are easily attainable in the inner disc, and thermionic and ion emission have been shown to be important sources of ionization there (Desch and Turner 2015). Additionally, in Chapter 2 we saw that an MRI-accreting inner disc features low gas surface densities, and so ionization of molecular hydrogen by the stellar X-rays may be comparable to thermal ionization as a source of free electrons.

In this chapter I present a new model of a steady-state disc that is accreting due to the MRI.

The vertical structure in this model is calculated self-consistently from viscous dissipation (due to the MRI-induced viscosity), stellar irradiation, and radiative and convective cooling (with realistic opacities due to dust grains). Ionization in the disc is driven by thermal ionization of potassium, thermionic and ion emission from dust grains, and ionization of molecular hydrogen by stellar X-rays, cosmic rays and radionuclides. Section 5.2 details all components of this model and the methods we can use to find a self-consistent steady-state solution for the disc structure. Section 5.3 presents the results, and in Section 5.4 I discuss the relative importance of the considered physical and chemical processes for the structure of the inner protoplanetary disc. In the next chapter I further investigate how the inner disc structure changes with dust grain size, dust-to-gas ratio, gas accretion rate and stellar mass, and discuss the implications for the formation of the super-Earths in the inner disc.

5.2. Methods

The disc is assumed to viscously accrete due to the MRI. The model of the disc structure is described in Section 5.2.1. The disc structure depends on the disc's radiative properties, i.e. opacities, and on the viscosity. The calculation of opacities is given in Section 5.2.2, and the prescription for the MRI-driven viscosity in Section 5.2.3. The MRI-driven viscosity is a function of the disc's ionization state, calculated using a chemical network described in Section 5.2.4. Disc structure, opacities, ionization and viscosity are calculated self-consistently at every point in the disc, using numerical methods given in Section 5.2.5.

The key parameters of the self-consistent disc model are the steady-state gas accretion rate $\dot{M}$, stellar mass M_*, stellar radius R_*, effective temperature T_*, and value of the viscosity in the absence of the MRI (the dead-zone viscosity). Additionally, disc opacities and ionization state, and thus the disc structure, depend on the properties of dust, namely, the dust-to-gas ratio f_{dg} and the maximum dust grain size $a_{\max}$.

5.2.1. The disc model

The disc model largely follows the work of D'Alessio et al. (1998; 1999). I consider a thin, axisymmetric, Keplerian, steady-state disc that is viscously accreting, and assume that the disc is in vertical hydrostatic equilibrium, heated by viscous dissipation and stellar irradiation, and that energy is transported by radiation and convection. The energy transport in the radial direction is neglected. Furthermore, at a given orbital radius, viscosity depends only on local conditions and the vertical mass column (see Section 5.2.3). Then, the structure of the disc at different radii is only coupled by the stellar irradiation, as it penetrates the disc along the line-of-sight to the central star.

5.2.1.1. Hydrostatic equilibrium

In a thin Keplerian disc in vertical hydrostatic equilibrium, at any given radius the gas pressure profile follows from

$$\frac{dP}{dz} = -\rho\Omega^2 z, \tag{5.1}$$

where P is the gas pressure, ρ is the gas volume density, Ω is the Keplerian angular velocity and z is the height above disc midplane. I adopt the ideal gas law.

5.2.1.2. Viscous heating and stellar irradiation

The disc is heated by the viscosity that drives the accretion. The viscosity is parametrized by $\nu = \alpha c_\mathrm{s}^2/\Omega$, where α is the Shakura and Sunyaev (1973) viscosity parameter and c_s is the isothermal sound speed. Local viscous dissipation rate at any location in the disc is given by

$$\Gamma_\mathrm{acc} = \frac{9}{4}\alpha P \Omega. \tag{5.2}$$

The flux generated by viscous dissipation that is radiated through one side of the disc is

$$F_\mathrm{acc} = \frac{3}{8\pi}\dot{M} f_r \Omega^2, \tag{5.3}$$

where $f_r = 1 - \sqrt{R_\mathrm{in}/r}$ comes from the inner disc edge thin-layer boundary condition, r is the orbital radius and $R_\mathrm{in} = R_*$ is the orbital radius of the inner disc edge (e.g. Frank et al. 2002), where I am assuming a zero-torque inner boundary condition.

I also consider heating due to stellar irradiation. Stellar flux propagates spherically from the star, and heating due to stellar irradiation at any location in the disc depends on the attenuation of this flux along the line-of-sight to the star. However, such treatment of the stellar flux in 2D coupled to a 1+1D disc model considered here, with no radial energy transport or scattering of stellar light, leads to multiple, or no solutions at all, for the equilibrium disc structure. Such behaviour does not appear in 2D disc models (e.g. Dullemond 2002). Therefore, it is necessary to use an assumption that the stellar flux hits the disc surface at an angle ϕ and then propagates vertically towards the midplane (Calvet et al. 1992; Chiang and Goldreich 1997). In this framework, the attenuation, i.e. the optical depth to the stellar flux, is approximated as $\tau_\mathrm{irr} \approx \tau_\mathrm{irr,z}/\mu$, where $\tau_\mathrm{irr,z}$ is the optical depth in the vertical direction, and $\mu = \sin\phi$. Local heating due to stellar irradiation is then given by

$$\Gamma_\mathrm{irr} = \kappa_\mathrm{P}^* \rho \frac{F_\mathrm{irr}}{\mu} e^{-\tau_\mathrm{irr,z}/\mu}, \tag{5.4}$$

where κ_P^* is the disc Planck opacity to stellar irradiation (see Section 5.2.2) and F_irr is the total incident stellar flux at the given orbital radius. In an optically-thick disc, the latter is given by

$$F_{\rm irr} = \sigma_{SB} T_*^4 \left(\frac{R_*}{s}\right)^2 \mu, \tag{5.5}$$

where s is the distance to the star (spherical radius) from the top of the disc at the given orbital radius.

The grazing angle ϕ is given by

$$\phi = \sin^{-1}\frac{4}{3\pi}\frac{R_*}{r} + \tan^{-1}\frac{d\log z_{\rm irr}}{d\log r}\frac{z_{\rm irr}}{r} - \tan^{-1}\frac{z_{\rm irr}}{r}. \tag{5.6}$$

Here the first term is the value of ϕ for a flat disc and comes from the finite size of the stellar disc, and the other two terms are due to disc flaring. Height $z_{\rm irr}(r)$ defines the surface in the disc at which the stellar flux is absorbed; here, $z_{\rm irr}(r)$ is the height above disc midplane at which the optical depth to stellar irradiation is $\tau_{\rm irr} = 2/3$. Specifically, in calculating the height $z_{\rm irr}$, the optical depth $\tau_{\rm irr}$ is obtained by integrating along spherical radius s to the star, as opposed to the approximation used in the local heating term. The height $z_{\rm irr}$ and the angle ϕ are determined self-consistently with the disc structure following the procedure by D'Alessio et al. (1999), as outlined in Section 5.2.5.

5.2.1.3. Radiative energy transport

We can consider the energy transport only in the vertical direction. In a thin disc the optical depth to disc radiation is much smaller in the vertical than in the radial direction. Thus, it can be expected that the radiation primarily transports heat vertically.

The frequency-integrated moments of the radiative transfer equation in the Eddington approximation (i.e. assuming that radiation is isotropic, as is valid in the optically thick regime) and the energy balance equation are

$$\frac{dF}{dz} = \Gamma_{\rm acc} + \Gamma_{\rm irr}, \tag{5.7}$$

$$\frac{dJ}{dz} = -\frac{3\rho\kappa_{\rm R}}{4\pi}F, \tag{5.8}$$

$$4\rho\kappa_{\rm P}(\sigma_{\rm SB}T^4 - \pi J) = \Gamma_{\rm acc} + \Gamma_{\rm irr}, \tag{5.9}$$

where F and J are radiative flux and mean intensity, respectively. I have also assumed here that the J and F weighted opacities can be approximated by the Planck mean opacity $\kappa_{\rm P}$ and the Rosseland mean opacity $\kappa_{\rm R}$, respectively (following e.g. Hubeny 1990, see Section 5.2.2).

Together with the ideal gas law, equations (5.1, 5.7-5.9) form a closed set of equations in P, F, J and T. Together with appropriate boundary conditions, they determine the disc vertical structure. One boundary condition is given at disc midplane, where due to symmetry the flux $F(0) = 0$. The rest of the boundary conditions are given at the top of the disc, at height $z_{\rm surf}$

above disc midplane. The boundary condition for the flux F is obtained by integrating eq. (5.7) from $z = 0$ to $z = z_{\mathrm{surf}}$, from which it follows that $F(z_{\mathrm{surf}}) = F_{\mathrm{acc}} + F_{\mathrm{irr}}$. The boundary condition for the mean intensity J is given by $J(z_{\mathrm{surf}}) = \frac{1}{2\pi} F(z_{\mathrm{surf}})$. Furthermore, we can assume that the gas pressure at the top of the disc has a small constant value, $P(z_{\mathrm{surf}}) = 10^{-10}\,\mathrm{dyn\,cm^{-2}}$, which is another boundary condition. This is arbitrary, but does not affect the results as long as the value is sufficiently small. Overall, thus, there are 4 boundary conditions for the three differential equations. Additionally, the temperature $T(z_{\mathrm{surf}})$ at the top of the disc follows from the algebraic eq. (5.9) and the boundary conditions for J and P.

Some applications of eq. (5.8) to accretion discs require an additional multiplicative factor on the right-hand side of the equation, the so-called flux-limiter, to preserve causality (Levermore and Pomraning 1981). This is not the case in the boundary value problem described above. The flux-limiter is introduced to ensure that $F < cE_{\mathrm{R}}$, where $E_{\mathrm{R}} = \frac{4\pi}{c} J$ is the radiation energy density. Here, this condition is already fulfilled everywhere in the disc, since at the top of the disc $F = cE_{\mathrm{R}}/2$ and vertically downwards E_{R} (or J) can only increase and F can only decrease (i.e., from eq. (5.7) and (5.8), $dF/dz > 0$ and $dJ/dz < 0$).

5.2.1.4. Energy transport by convection

If radiative energy transport yields a thermal structure such that gradient $\nabla = \frac{d\ln T}{d\ln P}$ is greater than the adiabatic gradient $\nabla_{\mathrm{ad}} = (\gamma - 1)/\gamma$, the gas is unstable to convection. In disc regions where this is the case we can assume that energy transport by convection is efficient and that gas is vertically isentropic, so $\nabla = \nabla_{\mathrm{ad}}$ at such locations (e.g. Shu 1992; Rafikov 2007). I adopt $\gamma = 1.4$. Then, the equations (5.8), (5.9) are replaced by

$$\frac{dT}{dz} = -\nabla_{\mathrm{ad}} \frac{T}{P} \rho(T, P) \Omega^2 z = -\nabla_{\mathrm{ad}} \frac{\mu m_{\mathrm{H}}}{k_{\mathrm{B}}} \Omega^2 z. \tag{5.10}$$

5.2.2. Opacities

Radiative transport is controlled by the Rosseland-mean opacity κ_{R} in optically-thick regions, eq. (5.8), and by the Planck-mean opacity κ_{P} in optically-thin regions, eq. (5.9). Additionally, absorption coefficient for the stellar flux is a Planck-mean opacity κ_{P}^{*} at the stellar effective temperature. In this work, it is assumed that the only source of these opacities are dust grains. Gas opacities are important in the very innermost regions of discs, where most dust species are sublimated. However, these regions are not of particular significance for the early stages of planet formation that we are interested in, as the dust is required to form solid planet cores. Beyond the silicate sublimation line, gas opacities may still be important in the hot, optically-thin regions (Malygin et al. 2014). However, the inner disc is significantly optically-thick and including gas opacities would only alter the structure of the hot disc atmosphere. Therefore, we can completely ignore the gas contribution to the opacities.

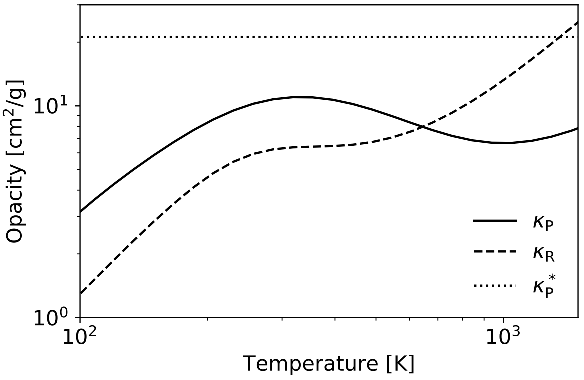


Figure 5.1: Planck-mean opacity $\kappa_{\rm P}$, Rosseland-mean opacity $\kappa_{\rm R}$, and Planck-mean opacity at the stellar effective temperature $\kappa_{\rm P}^*$, as functions of disc temperature, assuming dust-to-gas ratio $f_{\rm dg} = 10^{-2}$ and a maximum dust grain size $a_{\rm max} = 1\,\mu$m.

Furthermore, we can assume that the only dust species present are silicate grains. Other species that can be comparable in abundance to silicates are water ice and carbonaceous grains (e.g. organics, Pollack et al. 1994). However, due to their low sublimation temperatures, both water ice and carbonaceous grains are expected to be sublimated in the hot MRI-active regions.

To calculate the opacities, I use the optical constants of "astronomical silicates" from Draine (2003), and the MIESCAT module of the python wrapper RADMC3DPY for RADMC3D (Dullemond et al. 2012) to obtain dust absorption and scattering coefficients as functions of radiation wavelength and grain size. For each grain size a the coefficients are averaged within a size bin of width $\Delta \log a = 0.02$. Next, I adopt the grain bulk density of $\rho_{\rm g} = 3.3\,{\rm g\,cm^{-3}}$, and the grain size distribution given by $dn(a) \propto a^{-q} da$ (Mathis et al. 1977), with $q = 3.5$, minimum grain size $a_{\rm min} = 0.1\,\mu$m, and a maximum grain size $a_{\rm max}$. Then, the size-dependant absorption and scattering coefficients are weighted by grain mass and averaged over the grain size distribution. Finally, the absorption coefficient is integrated over frequency to obtain Planck-mean opacity ($\kappa_{\rm P}(T)$), and the total extinction coefficient yields Rosseland-mean opacity ($\kappa_{\rm R}(T)$). Following D'Alessio et al. (1998), I calculate the mean absorption coefficient for the stellar flux as a frequency-integrated absorption coefficient weighted by the Planck function at the stellar effective temperature ($\kappa_{\rm P}^* = \kappa_{\rm P}(T_*)$).

Fig. 5.1 shows the opacities per unit mass of gas, assuming a dust-to-gas ratio of $f_{\rm dg} = 10^{-2}$ and $T_* = 4400\,$K, for a maximum grain size $a_{\rm max} = 1\,\mu$m. Planck-mean opacity at the stellar effective temperature κ_P^* is a constant, since the wavelength-dependent dust absorption coefficient does not depend on the local temperature. The Planck-mean opacity $\kappa_{\rm P}$ is in general expected to increase with increasing temperature. This is because the wavelength at which the Planck function peaks is inversely proportional to the temperature, and for grains smaller than the wavelength of peak emission (and small grains contribute to the opacities most) absorption is expected to increase with decreasing wavelength. However, due to the silicate absorption feature at $10\,\mu$m, $\kappa_{\rm P}$ decreases with temperature in the range $\sim 500\,{\rm K} - 1000\,{\rm K}$.

5.2.3. Viscosity

The model of the MRI-driven viscosity closely follows that presented in Chapter 2 (see also Bai 2011a). Here only the main points are summarized. A well-ionized circumstellar disc which behaves according to laws of ideal magnetohydrodynamics is susceptible to the MRI (Balbus and Hawley 1991). Under such conditions, the MRI leads to turbulence, producing an accretion stress and acting as a source of viscosity. In terms of the Shakura-Sunyaev parameter α, the MRI yields

$$\alpha_{\mathrm{AZ}} = \frac{1}{3\beta} \tag{5.11}$$

where $\beta = P/P_B$ is the plasma parameter, and $P_B = B^2/8\pi$ is the magnetic field pressure (Sano et al. 2004).

However, even in the inner regions of protoplanetary discs, non-ideal magnetohydrodynamic effects quench the MRI, leading to the so-called dead zones. The non-ideal magnetohydrodynamic effects considered here are Ohmic and ambipolar diffusion. Ohmic diffusion will *not* supress the MRI if (Sano and Stone 2002)

$$\Lambda = \frac{v_{\mathcal{A}z}^2}{\eta_{\mathrm{O}}\Omega} > 1, \tag{5.12}$$

where Λ is the Ohmic Elsasser number, $v_{\mathcal{A}z} = B_z/\sqrt{4\pi\rho}$ is the vertical component of the local Alfven velocity and η_{O} is the Ohmic magnetic resistivity. Here, we can utilize the relationship between the vertical component of the magnetic field strength B_z and the r.m.s. field strength B, $B_z^2 \sim B^2/25$ (Sano et al. 2004). It is also assumed that B is vertically constant. The method of determining the value of B is described in Section 5.2.5.

Similarly, the ambipolar Elsasser number is defined by

$$Am = \frac{v_{\mathcal{A}}^2}{\eta_{\mathrm{A}}\Omega}, \tag{5.13}$$

where η_{A} is the ambipolar magnetic resistivity. However, in the strong-coupling limit valid in protoplanetary discs, the MRI can be active even if $Am < 1$, if the magnetic field is weak (Bai and Stone 2011). The criterion for the active MRI is then given by

$$\beta/\beta_{\mathrm{min}} > 1 \tag{5.14}$$

where the minimum value of β necessary to sustain the MRI is a function of the ambipolar Elsasser number,

$$\beta_{\mathrm{min}}(Am) = \left[\left(\frac{50}{Am^{1.2}}\right)^2 + \left(\frac{8}{Am^{0.3}} + 1\right)^2\right]^{1/2}. \tag{5.15}$$

Thus, whether the MRI will be active or not depends on the magnetic resistivities, η_{O} and η_{A}, as well as the magnetic field strength B. The magnetic resistivities, calculated following

Wardle (2007), express the coupling between the gas and the magnetic field. This is principally determined by the ionization of the gas.

If any of the two criteria given by eqns. (5.12), (5.14) is not fulfilled, the MRI is not active. In such MRI-dead zones, I assume there is a small residual viscosity α_{DZ}, driven either by propagation of turbulence from the MRI-active zone, or by hydrodynamic instabilities. In this chapter, I also impose a smooth transition between the active and the dead zone, necessary to ensure numerical stability in the integration of the equations of the disc structure. To cover both non-ideal effects that lead to dead zones, we can define $C = \min(\Lambda, \beta/\beta_{\mathrm{min}})$. Then, at any location in the disc, if $|C-1| < 0.5$,

$$\alpha = \alpha_{\mathrm{DZ}} + \frac{\alpha_{\mathrm{AZ}} - \alpha_{\mathrm{DZ}}}{1 + \exp\left(-\frac{C-1}{\Delta}\right)}, \tag{5.16}$$

where $\Delta = 10^{-2}$. The exact value of Δ makes very little difference to the results presented here.

5.2.4. Ionization

5.2.4.1. Chemical network

A simple chemical network is implemented following Desch and Turner (2015). The chemical network considers number densities of five species: atomic ions (n_{i}), free electrons (n_{e}), potassium ions ($n_{\mathrm{K^+}}$), neutral potassium atoms ($n_{\mathrm{K^0}}$) and potassium atoms adsorbed onto dust grains ($n_{\mathrm{K,cond}}$).

In the gas-phase potassium atoms are thermally ionized – collisions of neutral potassium atoms with hydrogen molecules produce potassium ions and free electrons. Potassium ions and free electrons recombine radiatively and also in three-body recombinations in collisions with hydrogen molecules (the latter process dominates at high densities present in the inner disc). Furthermore, non-thermal sources of ionization (e.g. X-rays) can ionize molecular hydrogen (Glassgold et al. 1997; Ercolano and Glassgold 2013). The charge is quickly transferred from the ionized hydrogen to other abundant gas species through collisions, producing molecular and atomic ions (e.g. $\mathrm{HCO^+}$, $\mathrm{Mg^+}$). Notably, in application to the MRI, the exact composition of the gas that this leads to is unimportant in the presence of dust, and simple chemical networks reproduce the gas ionization levels well (Ilgner and Nelson 2006). Thus, it is assumed that ionization of molecular hydrogen directly produces an atomic ion and a free electron at a rate $\zeta n_{\mathrm{H_2}}$. The atomic ion species in this chemical network may thus be understood as a representative of the various chemical species abundant in the gas-phase. Its mass is taken to be that of magnesium. It is assumed that the number density of molecular hydrogen is constant, which is valid for low ionization rates. Same as potassium, the atomic ions also recombine radiatively and in three-body recombinations.

Importantly, all gas-phase species collide with and are adsorbed onto dust grains at a rate

$$\mathcal{R}_{j,\,\mathrm{coll}} = n_j n_{\mathrm{gr}} \pi a_{\mathrm{gr}}^2 \left(\frac{8k_{\mathrm{B}}T}{\pi m_j}\right)^{1/2} \tilde{J}_j S_j, \tag{5.17}$$

where n_j is the number density of the gas-phase species, n_{gr} is the number density of the grains, a_{gr} is the grain size, m_j is the gas-phase species mass, $\tilde{J}_j$ is the modification of the collisional cross-sections for charged species due to dust grain charge (Draine and Sutin 1987), and S_j is the sticking coefficient. It is assumed that all grains have the same charge; this is valid since the dispersion of the distribution of charge states is generally found to be small(Draine and Sutin 1987). It is further assumed that potassium ions are quickly recombined on the grain surface to form condensed potassium atoms. The atomic ions are effectively destroyed upon adsorption.

At high temperatures electrons on the dust grains have a finite probability of leaving the grain, producing the so-called thermionic emission. The emission depends on the energy required for the electron to escape the grain. For a neutral grain this is the work function W, a property of the material out of which the grains are made. The rate at which free electrons are produced through thermionic emission is

$$\mathcal{R}_{\mathrm{therm}} = n_{\mathrm{gr}} 4\pi a_{\mathrm{gr}}^2 \lambda_{\mathrm{R}} \frac{4\pi m_{\mathrm{e}}(k_{\mathrm{B}}T)^2}{h^3} e^{-\frac{W_{\mathrm{eff}}}{k_{\mathrm{B}}T}}, \tag{5.18}$$

where

$$W_{\mathrm{eff}} = W + k_{\mathrm{e}} \frac{Ze^2}{a_{\mathrm{gr}}} \tag{5.19}$$

is the effective work function due to grain charge Ze, and k_{e} is the Coulomb constant.

Potassium atoms will also evaporate from the grains only at high temperatures. The vaporization rate of condensed potassium atoms is given by

$$\mathcal{R}_{\mathrm{K,evap}} = n_{\mathrm{K,cond}} \nu e^{-\frac{E_{\mathrm{a}}}{k_{\mathrm{B}}T}}, \tag{5.20}$$

where ν is the vibration frequency of potassium atoms on the dust grain surface lattice, and $E_{\mathrm{a}} = 3.26\,\mathrm{eV}$ is the binding energy, whose value is chosen to reproduce the condensation temperature of potassium (1006 K, Lodders 2003). These potassium atoms may be emitted into the gas phase as both neutral atoms and ions, contributing further to the gas' ionization state. The ratio of ions to neutrals among the emitted particles is given by

$$\frac{n_{\mathrm{K}}^{+}}{n_{\mathrm{K}^0}} = \frac{g_+}{g_0} e^{+\frac{W_{\mathrm{eff}}-\mathrm{IP}}{k_{\mathrm{B}}T}}, \tag{5.21}$$

where $\frac{g_+}{g_0}$ is the ratio of statistical weights of the ionized and neutral state of potassium, and IP the ionization potential of potassium.

Clearly, the contribution of the dust grains to gas ionization levels depends on the work function W of the grain material. I adopt the fiducial value of Desch and Turner (2015), $W = 5\,\mathrm{eV}$, and refer the reader to their work for a discussion of the experimental results supporting this choice. This value is close to the ionization potential of potassium, $\mathrm{IP} = 4.34\,\mathrm{eV}$, indicating that thermionic emission is important for the production of free electrons in the same temperature range as thermal ionization of potassium. Importantly, at this given value of the work function, grains become negatively charged at high temperatures as a large fraction of potassium evaporating from the grains is in ionized state. This results in a reduction of the effective work function W_{eff}, i.e., for negatively charged grains thermionic emission is higher.

In this work, I assume that the abundances of hydrogen and potassium atoms are $x_{\mathrm{H}} = 9.21\times10^{-1}$ and $x_{\mathrm{K}} = 9.87\times10^{-8}$, respectively, relative to the total number density of all atomic particles (Keith and Wardle 2014), and that the mean molecular weight is $\mu = 2.34 m_{\mathrm{H}}$[1]. The grain material density is $\rho_{\mathrm{gr}} = 3.3\,\mathrm{g\,cm^{-3}}$, the same as in the calculation of the dust opacities. The input for the chemical network are temperature T, pressure P, hydrogen ionization rate ζ, dust-to-gas ratio f_{dg} and dust grain size a_{gr}. All other kinetic rates, parameters and coefficients are the same as in Desch and Turner (2015). Note that the chosen value of the sticking coefficient for the electrons ($S_e = 0.6$) is compatible with the detailed calculation by Bai (2011a). The calculation therein is done for the values of work function of $1\,\mathrm{eV}$ and $3\,\mathrm{eV}$. The results suggest that for the work function of $5\,\mathrm{eV}$, at $1000\,\mathrm{K}$, S_e is indeed few times 0.1 for neutral grains, and increases further for negatively-charged grains.

For a given dust-to-gas ratio f_{dg} and grain size a_{gr}, the equilibrium number densities of electrons and ions, and the average grain charge, are pre-calculated and tabulated as functions of temperature T, pressure P and hydrogen ionization rate ζ. The equilibrium solution is found following the same method as Desch and Turner (2015). Time derivatives of all number densities are set to zero, and so rate equations yield an algebraic system of equations. For a given average grain charge Z, this system of equations is solved iteratively to find number densities of all five species. Then, the grain charge is found by solving the equation of charge neutrality.

5.2.4.2. *Effective dust-to-gas ratio*

The described chemical network incorporates only one grain size population. Ideally, we would consider a number of grain size populations, with the same size distribution used in the calculation of dust opacities. However, this would greatly enhance the computational complexity of the problem. At the same time, it is clear that dust grains of different size contribute differently to the equilibrium ionization levels. To the lowest order of approximation, all dust-related reaction rates are regulated by the total grain surface area. Thus, it can be expected that the ionization levels are most sensitive to the smallest grains. Bai and Goodman (2009) considered the effects of dust on the ionization levels in the cold regions of protoplanetary discs, in application to the

---

[1] The total number densities of molecular hydrogen and potassium are then related to the gas density as $n_{\mathrm{H_2}} = x_{\mathrm{H}}/(2 - x_{\mathrm{H}})\rho/\mu$ and $n_{\mathrm{K}} = 2x_{\mathrm{K}}/(2 - x_{\mathrm{H}})\rho/\mu$, respectively.

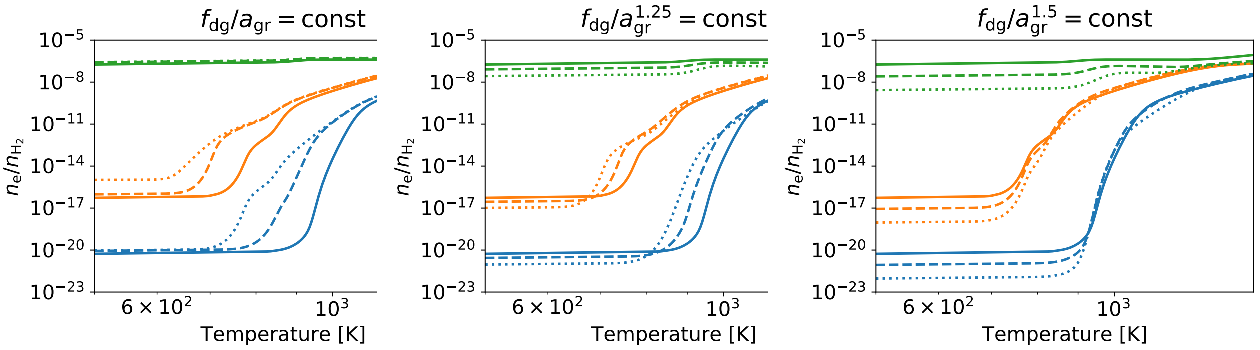


Figure 5.2: Ionization fraction (n_e/n_{H_2}) as a function of temperature. Different colours correspond to different combinations of pressure and hydrogen ionization rates, $P = 10^2\,\mathrm{dyn\,cm^{-2}}$ and $\zeta = 10^{-19}\,\mathrm{s^{-1}}$ (blue), $P = 10^{-2}\,\mathrm{dyn\,cm^{-2}}$ and $\zeta = 10^{-19}\,\mathrm{s^{-1}}$ (orange), $P = 10^{-2}\,\mathrm{dyn\,cm^{-2}}$ and $\zeta = 10^{-11}\,\mathrm{s^{-1}}$ (green). Different linestyles correspond to different grain sizes, $a_{gr} = 10^{-5}\,\mathrm{cm}$ (solid), $a_{gr} = 10^{-3}\,\mathrm{cm}$ (dashed), $a_{gr} = 10^{-1}\,\mathrm{cm}$ (dotted), each with a different dust-to-gas ratio such that the ratio f_{dg}/a_{gr}^p remains constant and as evaluated for $a_{gr} = 10^{-5}\,\mathrm{cm}$, $f_{dg} = 0.01$. Different panels show the calculations for different values of the exponent p, as indicated in each panel title. Exponent $p = 1$ is equivalent to keeping the total grain surface area constant. Adopting exponent $p = 1.5$ yields approximately the same threshold temperature at which ionization fraction sharply increases irrespective of grain size a_{gr}. See Section 5.2.4.2.

onset of the MRI due to non-thermal sources of ionization. They considered chemical networks with two grain size populations and found that the grain populations behave independently, as the charge transfer between the grains is negligible. They further found that the ionization levels are largely controlled by a quantity f_{dg}/a_{gr}^p where the exponent p varies between $p = 1$ (i.e. the total grain surface area) and $p = 2$.

We can repeat a similar exercise for the above chemical network, suited to the hot inner regions of protoplanetary discs. For a set of values of the exponent $p = 1, 1.25, 1.5$ grain size a_{gr} is varied while keeping f_{dg}/a_{gr}^p constant. The ionization levels are calculated as a function of temperature, and for different sets of pressure and hydrogen ionization rates, so as to probe different conditions in different regions of the inner disc. The results are shown in Fig. 5.2. The grain surface area ($p = 1$) controls the equilibrium ionization levels when the hydrogen ionization dominates, but does not determine the temperature at which the ionization levels rise due to thermionic and ion emission. On the other hand, for $p = 1.5$, this temperature depends very weakly on the grain size. That is, regardless of the actual dust grain size, a quantity $f_{dg}/a_{gr}^{1.5}$ regulates the temperature at which dense interior of the disc becomes ionized.

Therefore, we can proceed with a single grain size, $a_{gr} = 10^{-5}\,\mathrm{cm}$, in the chemical network, but, use an effective dust-to-gas ratio such that $f_{\mathrm{eff}} a_{gr}^{-p} = \int_{a_{\min}}^{a_{\max}} dn(a) m(a)/\rho_g a^{-p}$, where $n(a)$ is the same grain size distribution used to calculate dust opacities. Since it is found that thermionic and ion emission are more important than non-thermal sources of ionization in the inner disc, $p = 1.5$ is adopted. From Fig. 5.2 it appears that this choice will produce a large error in the ionization levels in the low-density non-thermally ionized disc regions. However, the discrepancy will be much lower for a size distribution than in Fig. 5.2, since most of the grains remain small.

5.2.4.3. Hydrogen ionization rate

In the calculation of the MRI-driven viscosity, I consider molecular hydrogen ionization rate due to radionuclides, cosmic rays and stellar X-rays. The ionization rate of molecular hydrogen due to short-lived and long-lived radionuclides is

$$\zeta_{\mathrm{R}} = 7.6 \times 10^{-19}\,\mathrm{s}^{-1}, \tag{5.22}$$

which predominantly comes from decay of ^{26}Al (Umebayashi and Nakano 2009). The ionization rate of molecular hydrogen due to interstellar cosmic rays is (Umebayashi and Nakano 2009)

$$\zeta_{\mathrm{CR}}(z) = \frac{\zeta_{\mathrm{CR,ISM}}}{2} e^{-\frac{\Sigma(z)}{\lambda_{\mathrm{CR}}}} \left(1 + \left(\frac{\Sigma(z)}{\lambda_{\mathrm{CR}}}\right)^{\frac{3}{4}}\right)^{-\frac{4}{3}}, \tag{5.23}$$

where $\zeta_{\mathrm{CR,ISM}} = 10^{-17}\,\mathrm{s}^{-1}$ is the interstellar cosmic-ray ionization rate, $\Sigma(z)$ is the integrated density column from the top of the disc to the height z above disc midplane, and $\lambda_{\mathrm{CR}} = 96\,\mathrm{g\,cm^{-2}}$ is the attenuation length for cosmic rays (Umebayashi and Nakano 1981).

For the ionization rate of molecular hydrogen due to stellar X-rays I use Bai and Goodman (2009) fits to Igea and Glassgold (1999) Monte Carlo simulations,

$$\zeta_{\mathrm{X}}(z) = \frac{L_{\mathrm{X}}}{10^{29}\,\mathrm{erg\,s^{-1}}} \left(\frac{r}{1\,\mathrm{AU}}\right)^{-2.2} (\zeta_1 e^{-(\Sigma(z)/\lambda_1)^{c_1}} + \zeta_2 e^{-(\Sigma(z)/\lambda_2)^{c_2}}), \tag{5.24}$$

where L_{X} is stellar X-ray luminosity, $\zeta_1 = 6 \times 10^{-12}\,\mathrm{s}^{-1}$, $\lambda_1 = 3.4 \times 10^{-3}\,\mathrm{g\,cm^{-2}}$, and $c_1 = 0.4$ characterize absorption of X-rays, and $\zeta_2 = 10^{-15}\,\mathrm{s}^{-1}$, $\lambda_2 = 1.6\,\mathrm{g\,cm^{-2}}$, and $c_2 = 0.65$ characterize the contribution from scattered X-rays. Here the attenuation lengths given by Bai & Goodman in terms of column densities of hydrogen nucleus are re-calculated into the surface density lengths using the hydrogen abundance given above. I adopt the relationship $L_{\mathrm{X}} = 10^{-3.5} L_{\mathrm{bol}}$ (Wright et al. 2011). For both cosmic rays and X-rays, I ignore the contribution coming through the other side of the disc. This is valid since in the inner disc the gas surface densities are mostly larger than the attenuation lengths of the ionizing particles and, even at low gas surface densities, this can only increase the ionization rates by at most a factor of 2.

5.2.5. Numerical methods

5.2.5.1. Equilibrium vertical disc structure

At a given orbital radius, magnetic field strength and grazing angle ϕ, the vertical disc structure is determined as a solution to the boundary value problem given by eqns. (5.1), (5.7)-(5.9) and the ideal gas law, and, where convectively unstable, by eqns. (5.1), (5.7) and (5.10). This boundary value problem is solved using the shooting method (Press et al. 2002). The equations are integrated from the top of the disc ($z = z_{\rm surf}$) to the disc midplane ($z = 0$). The height of the disc $z_{\rm surf}$ is then found, such that $F(0) = 0$. This root-finding problem is solved using the Ridders' method, with an exit criterion that $|F(0)| < 10^{-5}F(z_{\rm surf})$.

Equations of the vertical disc structure are integrated on a fixed grid with points uniform in the polar angle. Since eq. (5.9) is an algebraic equation, and for numerical stability, I use a fully implicit integration method. I use Runge-Kutta method of the second order, i.e. the trapezoidal method. In the trapezoidal method the equations (5.1), (5.7)-(5.9) are discretized as a system of nonlinear equations to be solved in every integration step (a system of nonlinear equations in P_{n+1}, F_{n+1}, J_{n+1}, T_{n+1} to be solved by a root-finding algorithm),

$$P_{n+1} = P_n + \frac{h}{2}\Omega^2(-\rho_n z_n - \rho_{n+1}z_{n+1}), \tag{5.25}$$

$$F_{n+1} = F_n + \frac{h}{2}(\Gamma_n + \Gamma_{n+1}), \tag{5.26}$$

$$J_{n+1} = J_n + \frac{h}{2}(-\frac{3}{4\pi})(\rho_n\kappa_{\rm R}(T_n)F_n + \rho_{n+1}\kappa_{\rm R}(T_{n+1})F_{n+1}), \tag{5.27}$$

$$0 = 4\rho_{n+1}\kappa_P(T_{n+1})(\sigma_{\rm SB}T_{n+1}^4 - \pi J_{n+1}) - \Gamma_{n+1}, \tag{5.28}$$

where h is the integration step, $\rho = \rho(T, P)$ is given by the ideal gas law and $\Gamma = \Gamma_{\rm acc} + \Gamma_{\rm irr}$. Here, $\Gamma_{{\rm irr},n} = \Gamma_{{\rm irr},n}(T_n, P_n, \tau_{{\rm irr},z,n})$, with the optical depth to stellar irradiation obtained using $\tau_{{\rm irr},z,n+1} = \tau_{{\rm irr},z,n} + \frac{h}{2}\kappa_{\rm P}^*(\rho_n + \rho_{n+1})$. Furthermore, viscous dissipation is a function of the MRI-driven viscosity α and thus depends on the local ionization levels. The latter are a function of the local temperature, pressure and the hydrogen ionization rate which depends on the column density from the top of the disc. Thus, $\Gamma_{{\rm vis},n} = \Gamma_{{\rm vis},n}(T_n, P_n, N_n)$, with the column density given by $N_{n+1} = N_n + \frac{h}{2}k_{\rm B}^{-1}(P_n/T_n + P_{n+1}/T_{n+1})$.

The equation for pressure P_{n+1} can be rearranged into an explicit form

$$P_{n+1} = \frac{P_n - \frac{h}{2}\Omega^2\rho_n z_n}{1 + \frac{h}{2}\Omega^2\frac{\mu m_{\rm H}}{k_{\rm B}}\frac{z_{n+1}}{T_{n+1}}}. \tag{5.29}$$

Then, the above system of equations is equivalent to a single nonlinear equation in T_{n+1}, greatly simplifying the problem. In every integration step the Ridders' method is used to solve this equation for the temperature T_{n+1} (down to a relative precision of 10^{-7}) and consequently for all other quantities. This includes the MRI-driven viscosity α, which is thus calculated self-consistently at each step of integration.[2] At every step and in every iteration of the root-solver opacities are interpolated from pre-calculated tables using cubic splines, and the ionization levels (e.g., free electron number density) using tri-linear interpolation.

---

[2] This is indeed necessary. An iterative method in which disc thermal structure is decoupled from the density structure and the heating terms (e.g. Dullemond 2002) does not converge to a solution in the case of MRI-driven viscosity.

Additionally, at each integration step it is checked if the resulting temperature gradient is unstable to convection, and if so, the temperature T_{n+1} is obtained analytically using

$$T_{n+1} = T_n + \frac{h}{2}\nabla_{\mathrm{ad}}\frac{\mu m_{\mathrm{H}}}{k_{\mathrm{B}}}\Omega^2(-z_n - z_{n+1}). \tag{5.30}$$

With T_{n+1} known, all other quantities follow same as above. A disc column can, in principle, become convectively stable again at some height above disc midplane. To calculate the mean intensity J at a boundary between a convective and a radiative zone, the energy balance equation (5.9) is used.

For some model parameters there can be a range of orbital radii and values of the magnetic field strength for which there are multiple solutions for the disc height z_{surf} (i.e. multiple solutions for the equilibrium vertical disc structure). This happens when a complex ionization structure leads to strong variations in the viscosity α as a function of height above disc midplane, making the total produced viscous dissipation a non-monotonous function of z_{surf}. When there are multiple solutions, the solution with smallest z_{surf} is chosen. It is likely that at least some of the additional solutions are unphysical, as the strong variations in both the levels of turbulence and the levels of ionization should be removed by turbulent mixing (see Section 5.3.2.3).

5.2.5.2. Magnetic field strength

At a given orbital radius, grazing angle ϕ and magnetic field strength B, the above procedure yields an equilibrium vertical disc structure characterized by a vertically-averaged viscosity

$$\bar{\alpha} = \frac{\int_0^{z_{\mathrm{surf}}} \alpha P dz}{\int_0^{z_{\mathrm{surf}}} P dz}. \tag{5.31}$$

As in the vertically-isothermal model (discussed in Chapter 2), for a sufficiently small and a sufficiently large magnetic field strength B the MRI is suppressed in the entire disc column and $\bar{\alpha} = \alpha_{\mathrm{DZ}}$. There can be an intermediate range of magnetic fields strength for which the MRI is active, and the vertically-averaged viscosity $\bar{\alpha}$ peaks at some value of B. At every orbital radius, B is chosen such that $\bar{\alpha}$ is maximized. The underlying assumption here is that the magnetic fields are strengthened by the MRI-driven turbulence, so that the magnetic field strength is a monotonously increasing function of $\bar{\alpha}$. To maximize $\bar{\alpha}(B)$, the Brent method is used, with a target absolute precision of 10^{-3} in $\log B$.

For some model parameters there can be a range of orbital radii where there are multiple local maxima in $\bar{\alpha}$ as a function of B. This is essentially for the same reasons that cause multiple solutions in disc height z_{surf} at a fixed value of B. In general B is chosen which corresponds to the global maximum in $\bar{\alpha}$. However, in some cases, this is a function of the grazing angle ϕ at a fixed orbital radius, and the procedure to determine the grazing angle (described below) does not converge. There a local maximum with a largest magnetic field strength is chosen.

5.2.5.3. Grazing angle

At any given orbital radius, the angle ϕ between the incident stellar radiation and the disc surface is determined self-consistently with the disc structure following D'Alessio et al. (1999). A self-consistent disc structure is found by iteratively updating the grazing angle ϕ and re-calculating the entire disc structure. I use a logarithmic grid for orbital radius. First, I calculate ϕ using eq. (5.6) by assuming that $z_{\rm irr} = 0$ and solve for the vertical disc structure and the magnetic field strength at all radii. I integrate through the obtained disc structure along lines-of-sight to the star to calculate $\tau_{\rm irr}(r, z)$, which yields an updated $z_{\rm irr}$ at each radius.

Critically, to calculate the updated value of the grazing angle ϕ at each radius, the derivative $d\log z_{\rm irr}/d\log r$ is approximated by assuming that $z_{\rm irr}$ is a power-law, $z_{\rm irr} \propto r^b$, within a radius bin centered at the given radius. So, at each radius I fit for the slope b using $z_{\rm irr}$ at that radius and at a number of radial grid points interior and exterior to it. Then, the value of the grazing angle is updated and the vertical disc structure re-calculated at all radii.

This procedure is repeated until the grazing angle has converged at every radius, i.e., until the relative difference in ϕ between two consecutive iterations is less than 0.5% at all radii. For the first radial point I always assume the flat-disc approximation ($z_{\rm irr} = 0$) and do not include it in the fitting routine. In this work I use a total of 100 radial points between 0.1 AU and 1 AU, and a total of 10 radial points in fits for $d\log z_{\rm irr}/d\log r$.

5.3. Results

I investigate structure of inner regions of protoplanetary discs using a model of a viscous steady-state accretion disc in which viscosity comes from the MRI. In Chapter 2, I discussed a vertically-isothermal, constant-opacity disc with a single thermally-ionized chemical species. In Section 5.3.1 I consider a model in which the vertical structure of the disc is calculated self-consistently from viscous heating, heating by stellar irradiation, radiative and convective energy transport, and self-consistent radiative properties of dust, and the disc's ionization state is calculated by only considering the thermal (collisional) ionization of potassium, using the Saha equation. Then, in Section 5.3.2 I present results of the full model that also includes other relevant chemical species, including dust grains, and non-thermal sources of ionization.

Throughout this chapter I assume the gas accretion rate $\dot{M} = 10^{-8}\,{\rm M_\odot\,yr^{-1}}$, stellar mass $M_* = 1\,{\rm M_\odot}$, stellar radius $R_* = 3\,{\rm R_\odot}$, effective stellar temperature $T_* = 4400\,{\rm K}$, and viscosity in the MRI-dead zone $\alpha_{\rm DZ} = 10^{-4}$. The chosen gas accretion rate is the median from observations (Hartmann et al. 1998; Manara et al. 2016; 2017). The chosen stellar parameters for a solar-mass star are from the stellar evolution models of Baraffe et al. (2015), for the stellar age of 5×10^5 yr. At later ages, stellar luminosity decreases, as the star contracts. Maximizing the stellar luminosity allows us to examine the maximum effect that the stellar irradiation has on the inner disc. Furthermore, I assume a dust-to-gas ratio $f_{\rm dg} = 10^{-2}$ and a maximum dust grain size $a_{\rm max} = 1\,\mu$m. In this work the focus is on examining the main physical and chemical processes that shape the structure of the inner disc, given the above model parameters. How the disc structure changes as a function of the model parameters is investigated in the next chapter.

5.3.1. Disc thermal structure and the MRI

In this section I explore how the thermal processes shape the inner disc structure. The chemistry considered in this section is simple, the disc's ionization state is set by the thermal ionization of potassium only, all of which is assumed to be in the gas phase. The effects of dust on the ionization levels, as well as non-thermal sources of ionization, are explored in Section 5.3.2.

5.3.1.1. Thermal structure of the inner disc

Fig. 5.3 shows the temperature as a function of orbital radius and height above disc midplane for three models of varying complexity. The left-hand panel shows a disc which has the same constant opacity ($\kappa_{\rm R} = \kappa_{\rm P} = 10\,{\rm cm}^2\,{\rm g}^{-1}$) as in Chapter 2, but where the disc vertical structure is calculated self-consistently from viscous heating and cooling by radiation and convection. The middle panel is a model in which the opacities are also determined self-consistently. Finally, the right-hand panel is a model which further accounts for the heating by stellar irradiation.

Noticeably, the temperature profiles deviate from vertically-isothermal. Here, the dashed lines show the disc pressure scale height, and the solid lines show the disc photosphere (where the Rosseland-mean optical depth is $\tau_{\rm R} = 2/3$). In all three models the temperature increases towards the midplane below the photosphere, as the disc becomes more and more optically-thick to its own radiation.

The resulting temperature gradient becomes sufficiently high, that the disc becomes convectively unstable, as shown in Fig. 5.4. From the disc midplane to a height of a couple of pressure scale heights, energy is thus transported by convection, and there we can assume that the temperature gradient is isentropic. Importantly, the strong temperature gradient that gives the rise to the convective instability is not specific to the MRI-driven accretion. It is a feature of any active optically-thick disc where the accretion heat is released near midplane (see also Garaud and Lin 2007). This is shown analytically in Section 5.4.

The model that also includes heating due to stellar irradiation, shown in the right panels of Fig. 5.3 and Fig. 5.4, features a temperature inversion in the disc upper layers. This temperature inversion has been discussed in detail by D'Alessio et al. (1998). The dotted line here indicates the irradiation surface $z_{\rm irr}$ at which $\tau_{\rm irr} = 2/3$. Above this line the disc upper layers are heated by stellar irradiation. Below, disc becomes optically-thick to incident irradiation and temperature drops. Then, below the disc photosphere (the solid line), disc is optically-thick to its own radiation and temperature increases again.

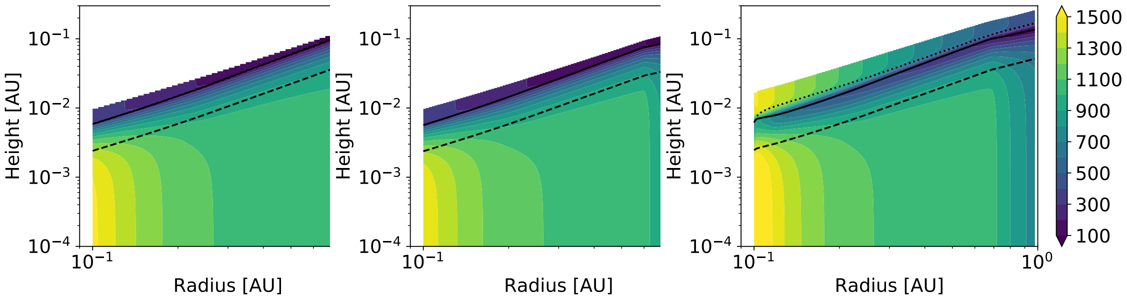


Figure 5.3: Temperature as a function of location in the disc for a viscously-heated constant-opacity model (left), viscously-heated model with realistic opacities (middle), and a viscously and irradiation-heated model with realistic opacities (right). In each panel the solid line shows the disc photosphere ($\tau_{\rm R} = 2/3$) and the dashed line shows the pressure scale height ($P = e^{-1/2} P_{\rm mid}$). The dotted line in the right-hand panel shows the surface at which $\tau_{\rm irr} = 2/3$. Note that the inclusion of heating by stellar irradiation does not strongly affect the disc midplane temperature. See Section 5.3.1.1.

In the irradiated disc, close to the star the disc midplane is as hot as the disc upper layers, and further away the midplane is significantly hotter than the upper layers. It would appear that the accretion heating dominates in the inner disc. However, the total flux of stellar radiation absorbed by a vertical column in the disc at a given radius, shown in Fig. 5.5, is in fact of the order of few tens of the total flux generated by viscous dissipation ($F_{\rm irr} \sim 10 F_{\rm acc}$) throughout the inner disc. Nevertheless, the disc temperature near midplane is weakly affected by irradiation. I discuss this result in Section 5.4.2.

The ratio $F_{\rm irr}/F_{\rm acc}$ varies non-monotonically, following the grazing angle, shown in the right panel of Fig. 5.5. At the inner edge of the calculation domain, the first ~10 points (shown in grey) are affected by the boundary effects. This is a well known problem in disc models that account for stellar irradiation using the grazing angle prescription (Chiang et al. 2001). Importantly, far away from the inner edge, the value of the grazing angle and the disc structure do not depend on the disc structure at the inner edge.

5.3.1.2. *Ionization levels and non-ideal MHD effects*

The ionization structure of the models with a self-consistent vertical structure is qualitatively different than in the vertically-isothermal models discussed in Chapter 2. In the vertically-isothermal models the ionization fraction increases with height above disc midplane, as the temperature is constant and density decreases (which follows from the Saha equation). Here, as the temperature decreases with height, so does the ionization fraction, as shown in the middle three panels of Fig. 5.6. Here, the top three panels show the viscosity α as a function of height above disc midplane, middle panels show the free electron number density relative to the number density of molecular hydrogen $n_{\rm e}/n_{\rm H_2}$, and bottom panels show the Ohmic and ambipolar resistivities. Solid lines are for the viscously-heated model with self-consistent dust

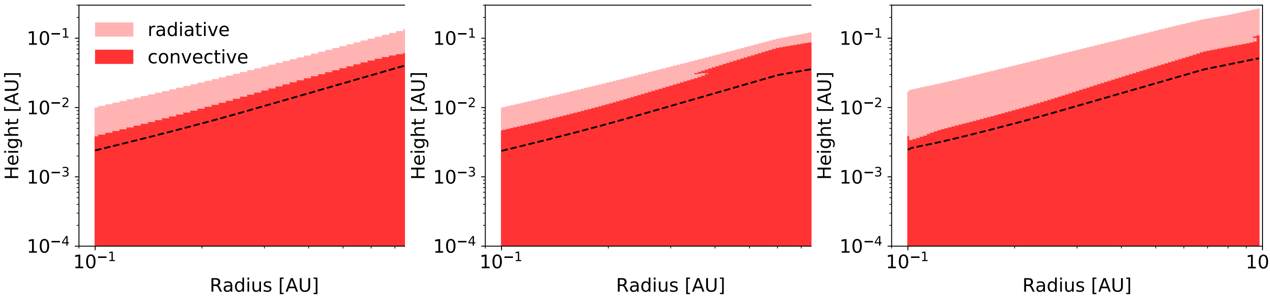


Figure 5.4: Radiative and convective zones (light and dark red, respectively) for a viscously-heated constant-opacity model (left), viscously-heated model with realistic opacities (middle), and a viscously and irradiation-heated model with realistic opacities (right). In each panel the dashed line shows the pressure scale height ($P = e^{-1/2}P_{\rm mid}$). In all three models the disc is convectively unstable within few scale heights. See Section 5.3.1.1.

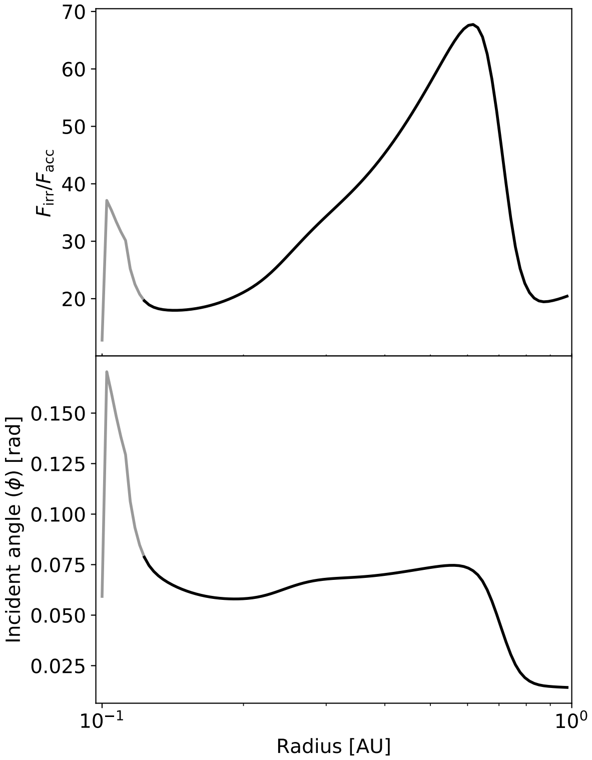


Figure 5.5: Ratio of the total irradiation heating to the total viscous dissipation ($F_{\rm irr}/F_{\rm acc}$) as a function of radius (top) and the incident angle to stellar irradiation (ϕ) as a function of radius (bottom). The total (vertically-integrated) absorbed stellar flux $F_{\rm irr}$ is at least an order of magnitude larger than the total viscous dissipation $F_{\rm acc}$ at any given radius. Gray lines indicate the region affected by the inner boundary condition for the incident angle ϕ. See Section 5.3.1.1.

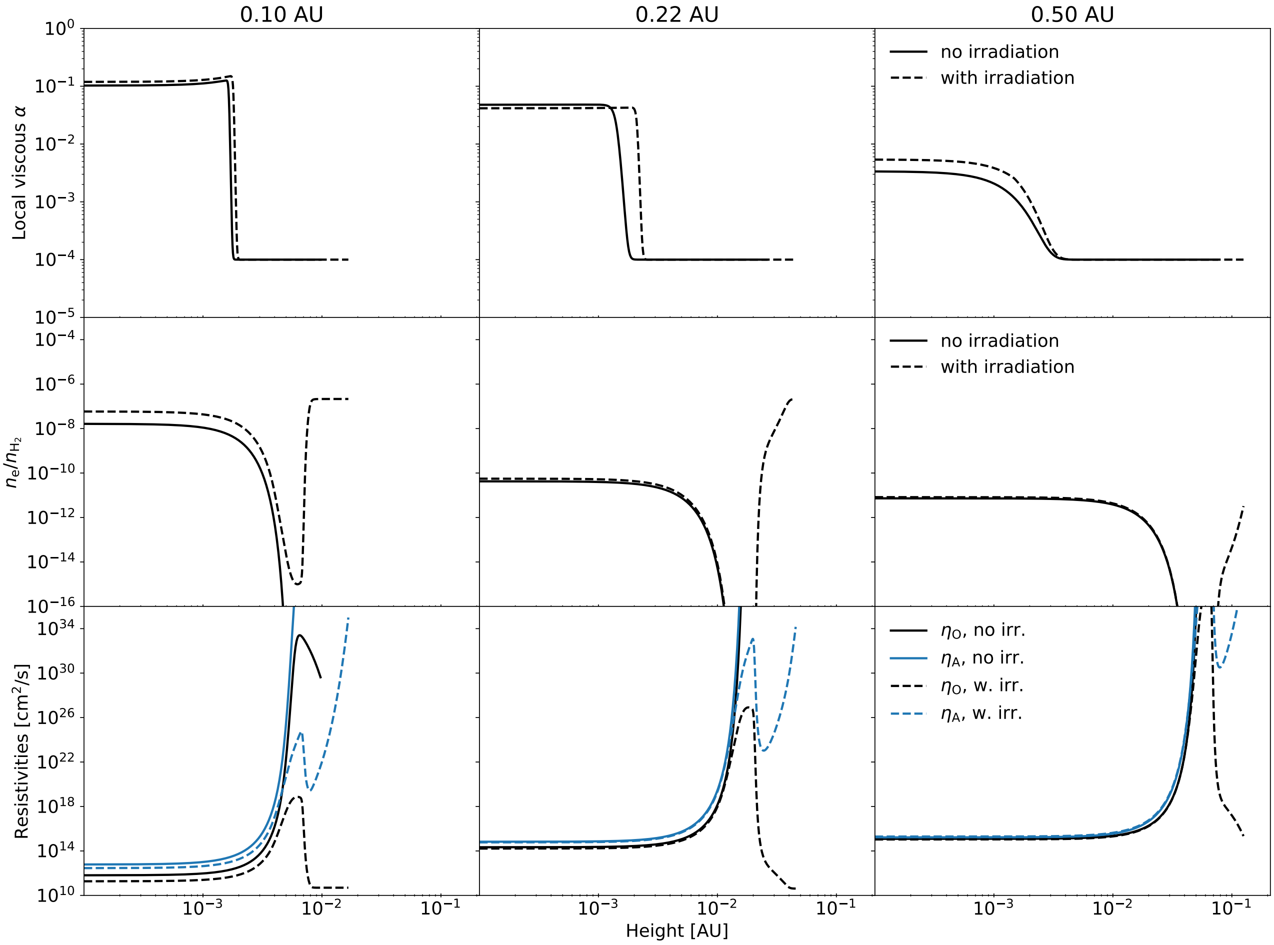


Figure 5.6: Local viscous parameter α (top), ionization fraction $n_{\mathrm{e}}/n_{\mathrm{H}_2}$ (middle) and magnetic resistivities (bottom) at three different orbital radii (as indicated in panel titles) in a model with realistic opacities, but no irradiation, and a model that also includes irradiation, as shown in plot legend. Despite the high ionization fraction in the irradiation-heated disc upper layers, ambipolar diffusion quenches the MRI there. See Section 5.3.1.2.

opacities, and dashed lines are for the model that also includes heating by stellar irradiation. In the latter model, the ionization does increase in the uppermost irradiated layers.

As a consequence of the above, in non-irradiated discs both ambipolar and Ohmic resistivity increase with height. In other words, in these self-consistent models, both ambipolar and Ohmic diffusion quench the MRI from above. The Ohmic resistivity decreases in the uppermost layers of the irradiated disc, however, ambipolar diffusion does not.

5.3.1.3. MRI-active and dead zones

The differences in the ionization structure of the vertically self-consistent models and the vertically-isothermal models lead to differences in where in the disc the MRI is active. Plot of the viscosity α as a function of orbital radius and height above disc midplane, Fig. 5.7, shows that the MRI-active zone (where $\alpha > \alpha_{\mathrm{DZ}}$) remains around midplane. This is qualitatively different from the vertically-isothermal case, where the MRI-active zone is around midplane

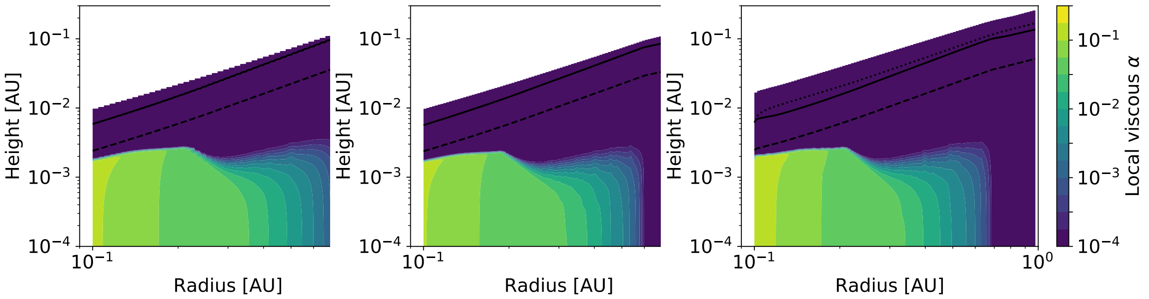


Figure 5.7: Local viscous parameter α as a function of location in the disc for a viscously-heated constant-opacity model (left), viscously-heated model with realistic opacities (middle), and a viscously and irradiation-heated model with realistic opacities (right). In each panel the solid line shows the disc photosphere ($\tau_{\rm R} = 2/3$) and the dashed line shows the pressure scale height ($P = e^{-1/2}P_{\rm mid}$). The dotted line in the right-hand panel shows the surface at which $\tau_{\rm irr} = 2/3$. Note that the heating by stellar irradiation has a very weak effect on the extent of the MRI-active region. See Section 5.3.1.3.

in the innermost region, but at a certain orbital distance rises into the upper layers. Such a configuration in which an MRI-active zone is nested between a dead zone at midplane, and ambipolar-dead zone above, arises in the vertically-isothermal case because there the ionization levels increase with height above disc midplane. Furthermore, the three panels correspond to the same three models as above: viscously-heated constant-opacity model, viscously-heated model with self-consistent opacities, and a fully self-consistent model which also includes heating by stellar irradiation. Noticeably, stellar irradiation makes little difference to the extent of the MRI-active zone, since irradiation affects the midplane temperature weakly, and the hot uppermost layers are dead due to the ambipolar diffusion.

Finally, Fig. 5.8 compares radial profiles of the vertically-averaged viscosity $\bar{\alpha}$ and the magnetic field strength B from four models of varying complexity. The grey solid lines show the results of the viscously-heated vertically-isothermal constant-opacity model from Chapter 2. The black dotted line is for the disc which has the same constant opacities, but where the disc vertical structure is calculated self-consistently from viscous heating and cooling by radiation and convection. The black dashed line is for the model in which the opacities are also determined self-consistently and the black solid line is for the model which further accounts for the heating by stellar irradiation.

In line with theoretical expectations and the previous study, the vertically-averaged viscosity $\bar{\alpha}$ decreases as a function of orbital radius. At some distance from the star the MRI is completely quenched and the viscosity reaches the minimum value, $\bar{\alpha} = \alpha_{\rm DZ}$. Remarkably, the $\bar{\alpha}$ radial profile is both qualitatively and quantitatively similar in these models. We can expect this to be the case in general, as long as the disc is optically thick, so that the heating by stellar irradiation may be neglected. However, the radial profile of the magnetic field strength B reveals qualitative differences. In the vertically-isothermal model, $B(r)$ features a sharp drop

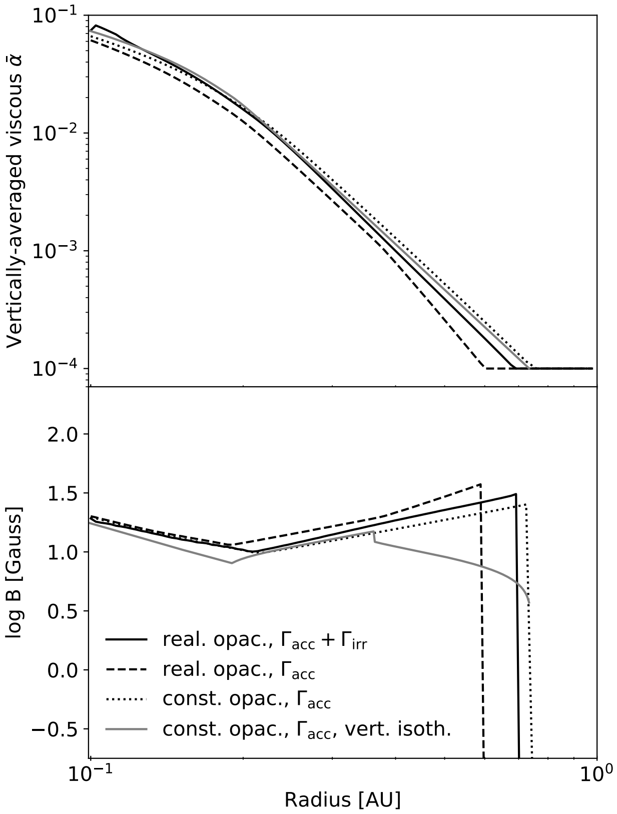


Figure 5.8: Vertically-averaged viscosity parameter ($\bar{\alpha}$) and the magnetic field strength (B) as functions of radius in a vertically-isothermal and self-consistent models of varying complexity (a viscously-heated constant-opacity model, viscously-heated model with realistic opacities, and a viscously and irradiation-heated model with realistic opacities), as indicated in plot legend. The vertically-averaged viscosity parameter profile is similar in all four models, but the magnetic field strength profile is qualitatively different in the vertically-isothermal model. See Section 5.3.1.3.

at $\sim 0.35\,\mathrm{AU}$, which corresponds to the rise of a dead zone at disc midplane described above (see also Chapter 2). From that point, until the MRI is completely quenched at $\sim 0.7\,\mathrm{AU}$, the MRI is active in a thin layer high above the midplane, between a dead zone at disc midplane and a zombie zone above. In the self-consistent models, however, this configuration does not appear and the magnetic field strength remains high until the MRI is completely quenched.

5.3.2. Disc chemical structure and the MRI

In this section I build upon the model from the previous section. First, I explore the effects of dust on the disc's ionization state and on the MRI. Then, I produce the full model of the inner disc by also considering non-thermal sources of ionization.

5.3.2.1. Effects of dust

Fig. 5.9 shows the radial profiles of the vertically-averaged viscosity $\bar{\alpha}$, magnetic field strength, midplane temperature and the midplane free electron number density for a model with a self-consistent vertical structure and thermal ionization only (solid lines), and a model in which the dust effects on the ionization state are included (dashed lines). The latter model includes gas-phase collisional (thermal) ionization and recombination, adsorption of free charges on the dust grains, and thermionic and ion emission from the dust grains. Non-thermal ionization is also included, but at a negligible rate of $\zeta = 10^{-25}\,\mathrm{s}^{-1}$ to isolate the high-temperature effects. As Fig. 5.9 shows, the radial profile of $\bar{\alpha}$ and the radial extent of the MRI-active zone are overall similar for both models. Interestingly, when the dust effects are included, $\bar{\alpha}$ is higher at a given orbital radius, as is the ionization fraction, i.e. $n_\mathrm{e}/n_{\mathrm{H}_2}$. Concurrently, the midplane temperature is lower in the model that includes dust.

Fig. 5.10 shows the vertical structure of the model with dust at three different orbital radii. Top panels show the viscosity α as a function of height, middle panels show the number densities of free electrons and ions, and bottom panels show the various contributions to the free electron production rate (in $\mathrm{s}^{-1}\,\mathrm{cm}^{-3}$). As in the thermally-ionized disc (Fig. 5.7), the MRI is active at the hot disc midplane. The ionization levels decrease with height above the midplane as temperature decreases, and increase again in the disc atmosphere heated by stellar irradiation; here, however, ambipolar diffusion quenches the MRI.

The plots of free electron production rates show that thermionic emission is the dominant ionization source. Thus, Fig. 5.9 is misleading in the sense that, while the differences in the global structure are very small when dust effects are added, it is not because dust effects are small. Clearly, dust dominates the chemistry in the inner disc. Rather, for the parameters assumed here, the ionization levels as a function of temperature and density are similar to the levels obtained from gas-phase thermal ionization only.

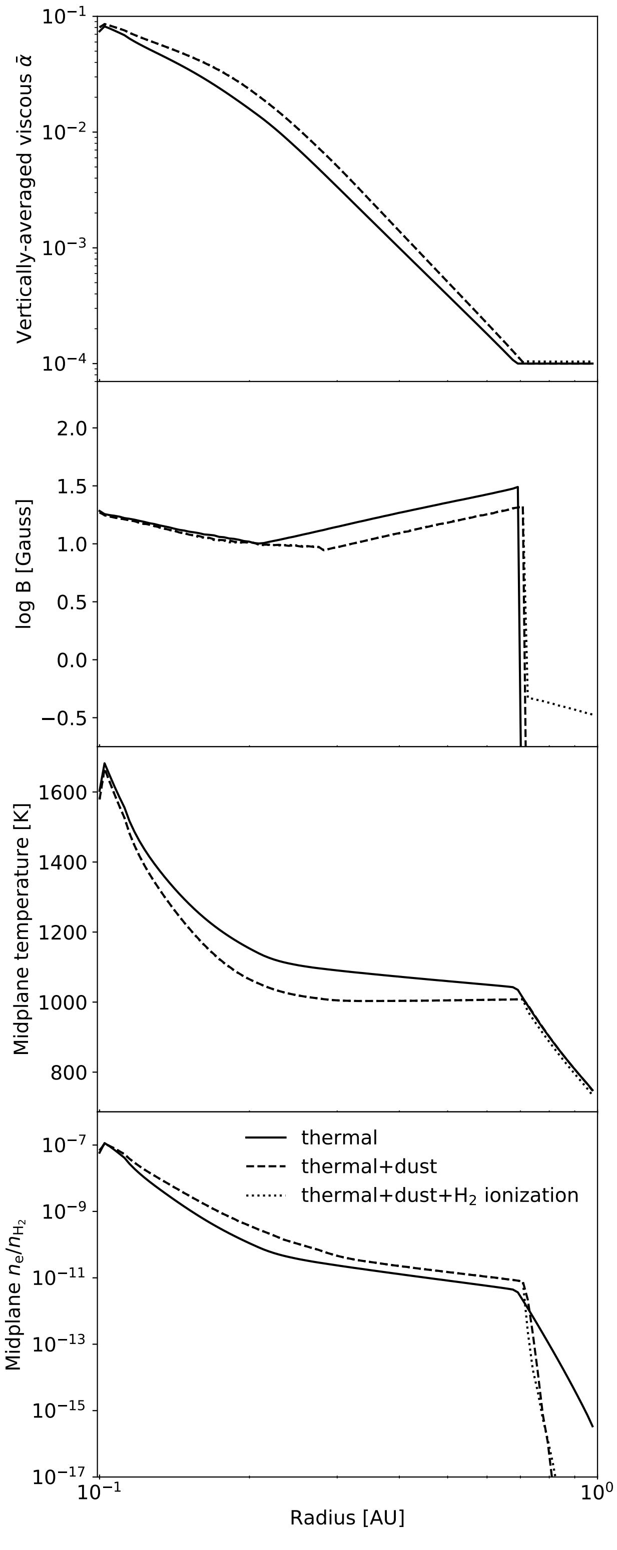


Figure 5.9: Vertically-averaged viscous parameter ($\bar{\alpha}$), magnetic field strength (B), midplane temperature and the midplane free electron fraction (n_e/n_{H_2}) as functions of radius for a model with thermal ionization only, a model with thermal ionization and dust effects (with $\zeta = 10^{-25}\,\mathrm{s}^{-1}$), and a model which also includes non-thermal sources of ionization ($\zeta = \zeta_R + \zeta_{CR} + \zeta_X$). The disc structure is quantitatively similar in all three models; however, the primary source of ionization in the models that include dust are thermionic and ion emission. See Section 5.3.2.1.

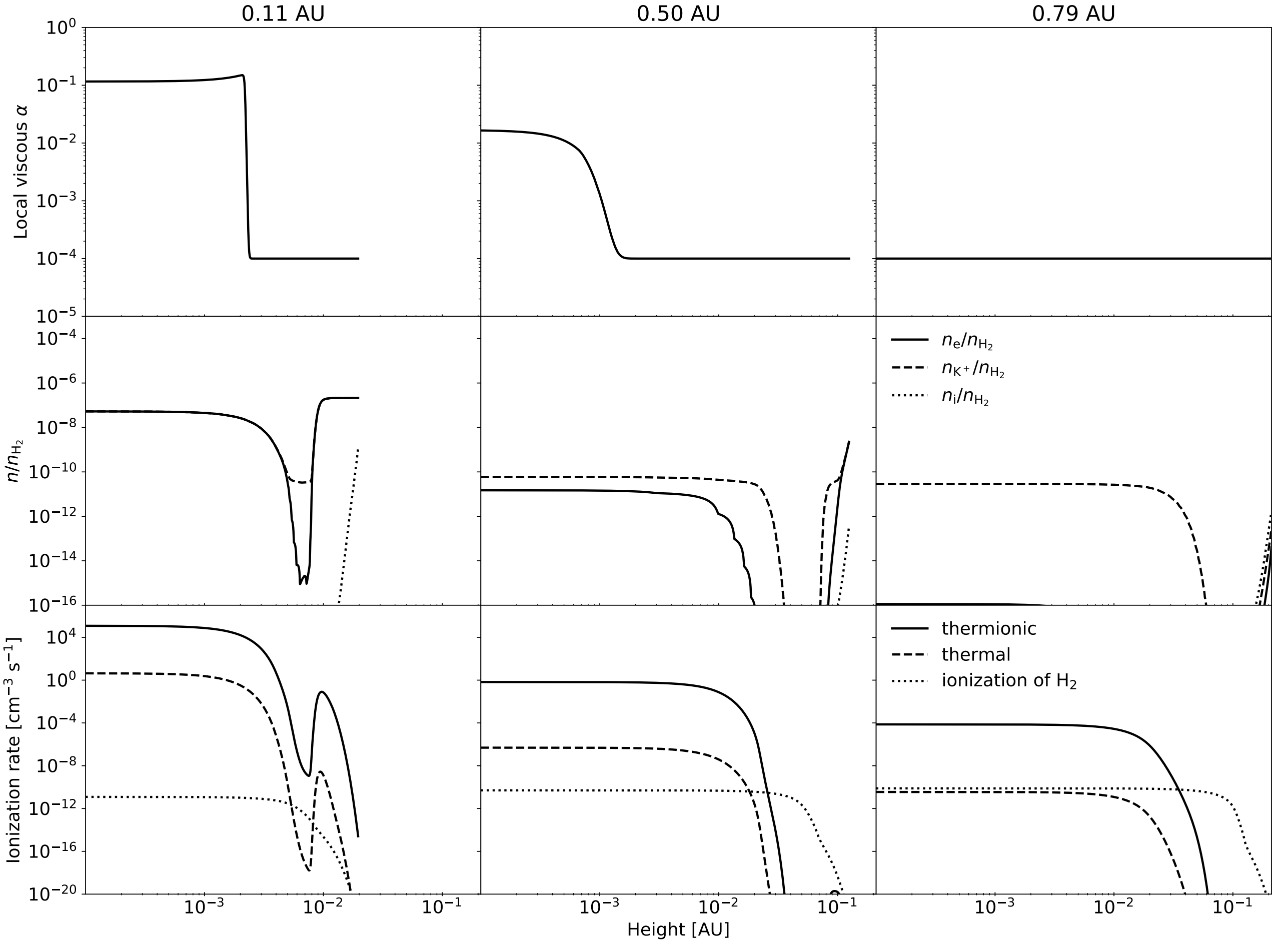


Figure 5.10: Local viscous parameter α (top), ionization fraction $n_\mathrm{e}/n_{\mathrm{H}_2}$ (middle) and ionization rates (bottom; thermionic $\mathcal{R}_\mathrm{therm}$, thermal $k_2 n_{\mathrm{H}_2} n_{\mathrm{K}^0}$ and non-thermal ζn_{H_2}) at three different radii (as indicated in panel titles) for the model with thermal ionization and dust effects (with $\zeta = 10^{-25}\,\mathrm{s}^{-1}$). Thermionic emission is the primary source of free electrons in the MRI-active regions. See Section 5.3.2.1.

5.3.2.2. Non-thermal sources of ionization

In this section I present the full model of the inner disc which, in addition to the above, also includes non-thermal sources of ionization. Stellar X-rays, cosmic rays and radionuclides ionize hydrogen, producing atomic ions and free electrons. The resulting radial profiles of $\bar{\alpha}$, magnetic field strength, midplane temperature, density and ionization levels are shown in Fig. 5.9 as dotted lines. The radial profile of $\bar{\alpha}$ seems to overlap with the model where the hydrogen ionization rate was set to a negligibly low value. However, in the region where the MRI is dead in the previous models, $\bar{\alpha}$ is slightly higher than $\alpha_{\rm DZ}$ in this model. The magnetic field strength does not drop to zero in this model, further revealing that the MRI remains active here.

Fig. 5.11 shows the vertical structure of this disc at three different orbital radii. It reveals that at large orbital radii where the MRI was quenched in previous models, it remains active at large heights (see also Fig. 5.12). The plots of ionization levels show that the fraction of free electrons and atomic ions increase immensely towards disc surface where the non-thermal ionization dominates. As a result of the high electron fraction in the upper layers, potassium ions are depleted there, as they can easily recombine with the abundant electrons. Note that the fraction of free electrons and atomic ions become extremely large near the disc surface, i.e. higher than the total Mg abundance, and even total C, O abundances. The plots are cut off at $n_{\rm j}/n_{\rm H_2} > 10^{-3}$. Clearly, above this level the chemical model is not applicable, as the ionized hydrogen would become an important species and the assumption of a constant hydrogen number density would be invalid. Nevertheless, this is above the MRI active zone at all orbital radii, and it is thus not pertaining to the conclusions.

Fig. 5.13 compares the contributions from stellar X-rays and cosmic rays to the hydrogen ionization rate. As expected, the un-attenuated ionization rate due to X-rays is higher at the disc surface, but cosmic rays can penetrate deeper in the disc. Nevertheless, the MRI-active region in the upper disc layers is mostly ionized by X-rays.

Finally, note that in this outer region the gas accretes primarily through the dead zone, since the density at the dead disc midplane is much higher than in the X-ray-ionized MRI-active layer. This is why the vertically-averaged viscosity parameter is close to the dead-zone value, $\bar{\alpha} \sim \alpha_{\rm DZ}$ outwards from the pressure maximum.

5.3.2.3. Multiple solutions for the vertical disc structure

In the results presented so far, for the fiducial choice of disc and dust parameters, solutions for the vertical disc structure (i.e., solutions in disc height $z_{\rm surf}$) appear to be unique at any given orbital radius, grazing angle ϕ and magnetic field strength B. In general, there is also a single peak in the vertically-averaged viscosity $\bar{\alpha}$ as a function of magnetic field strength B (which determines the choice for B). The exception is the vicinity of the orbital radius at which the MRI becomes quenched at disc midplane (see Fig. 5.12). There, $\bar{\alpha}(B)$ has two peaks, one

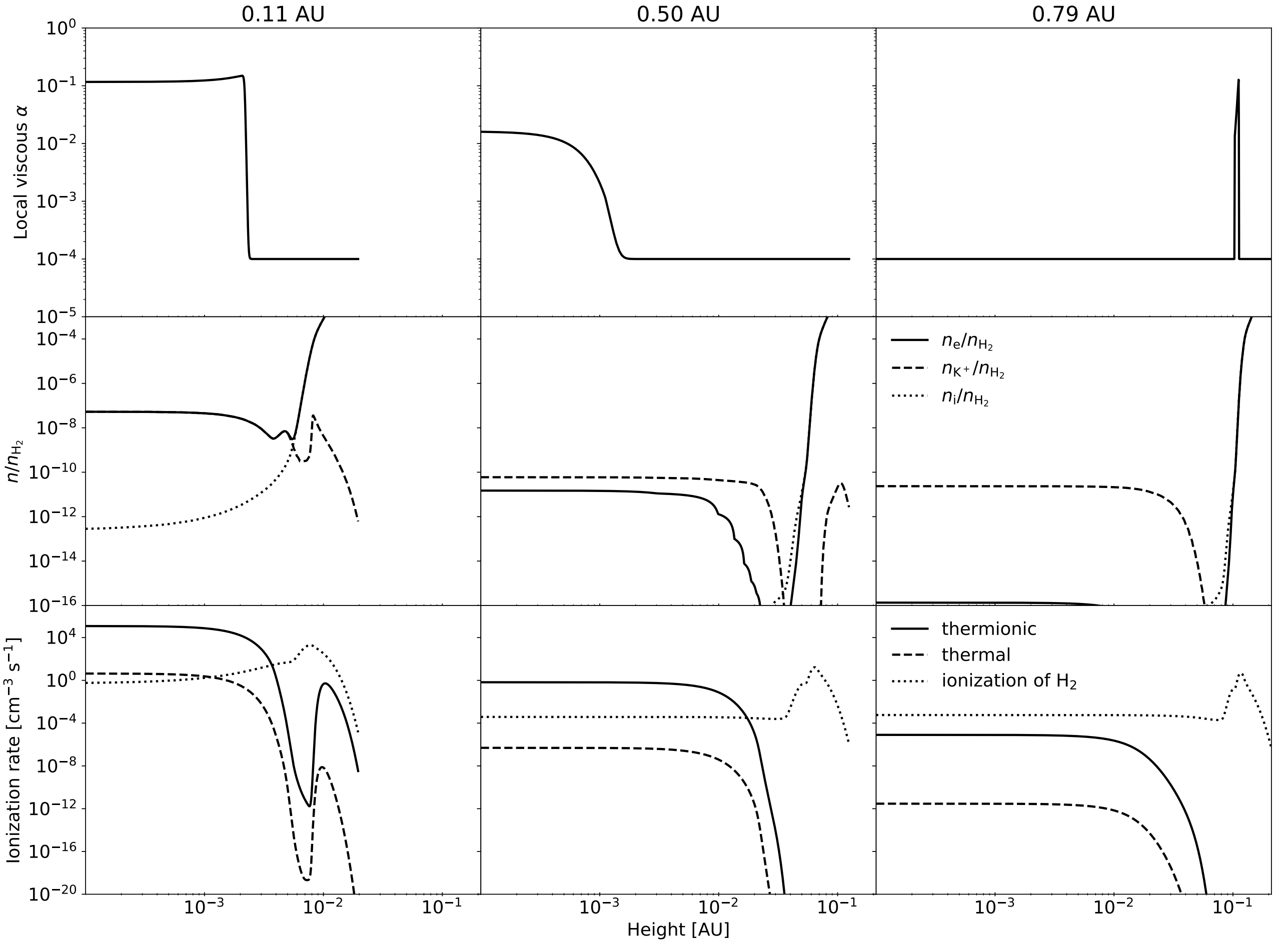


Figure 5.11: Local viscous parameter α (top), ionization fraction $n_\mathrm{e}/n_{\mathrm{H}_2}$ (middle) and ionization rates (bottom; thermionic $\mathcal{R}_\mathrm{therm}$, thermal $k_2 n_{\mathrm{H}_2} n_{\mathrm{K}^0}$ and non-thermal ζn_{H_2}) at three different radii (as indicated in panel titles) for the model with all sources of thermal and non-thermal ionization. Non-thermal ionization produces an MRI-active region high above disc midplane at larger radii (see the top right panel). See Section 5.3.2.2.

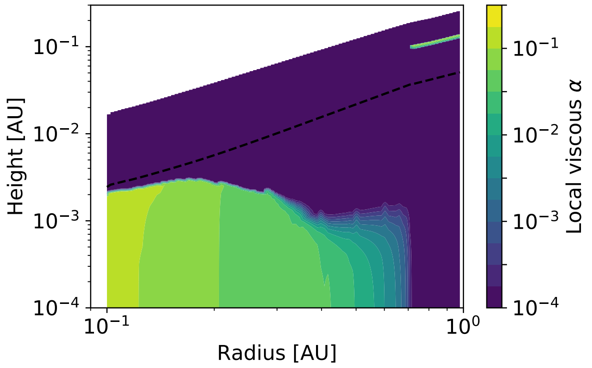


Figure 5.12: Local viscous parameter α as a function of location in the disc for the model with all sources of thermal and non-thermal ionization. In the innermost disc, thermionic and ion emission ionize the dense regions around the disc midplane, producing the high MRI-driven α there. At larger radii, the MRI is active in a thin layer high above the disc midplane, dominated by non-thermal sources of ionization. See Section 5.3.2.2.

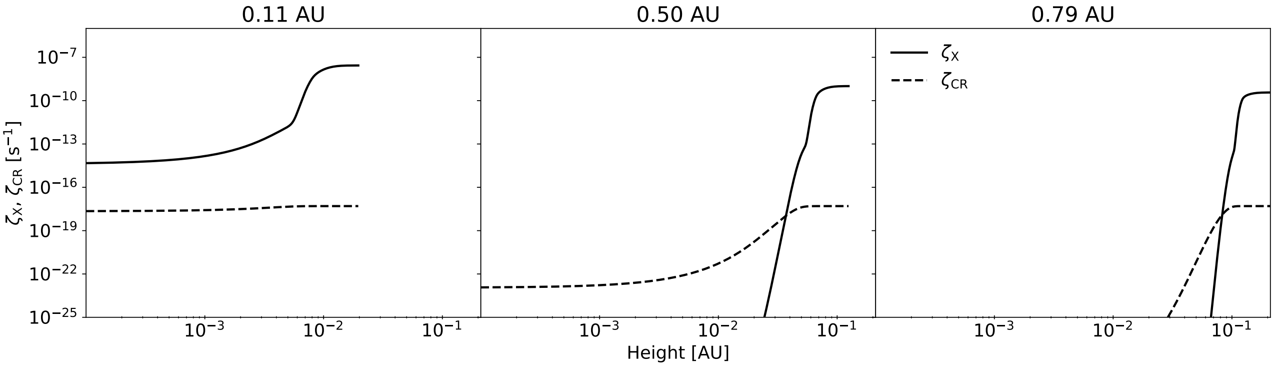


Figure 5.13: Ionization rates of molecular hydrogen due to stellar X-rays (ζ_X) and due to cosmic rays (ζ_{CR}) for the model with all sources of thermal and non-thermal ionization. Stellar X-rays are the dominant source of ionization in the disc upper layers. See Section 5.3.2.2.

corresponding to the solution where the MRI is active at disc midplane, and the other to the solution where the MRI is active in the upper disc layers, mostly ionized by stellar X-rays. As discussed in Section 5.2.5, B is chosen such that $\bar{\alpha}$ is maximized in this case as well.

Note that there could be, in principle, a thin MRI-active layer high up in the disc at shorter orbital radii as well, in addition to the active layer at midplane. Here, this does not appear due to the assumption that the magnetic field strength B is vertically constant. At the high B necessary to drive the efficient accretion at midplane, the MRI is quenched in the low-density disc atmosphere due to ambipolar diffusion. Therefore, it is only when the temperature drops and the high-temperature ionization cannot drive the accretion efficiently that the model features the thin active layer high up above the disc midplane.

For a different choice of parameters, e.g. if the maximum dust grain size is $a_{\max} = 100\,\mu$m, there may exist an additional range of orbital radii where there are multiple peaks in $\bar{\alpha}(B)$ and also multiple solutions for the disc vertical structure (for $z_{\rm surf}$) at a fixed value of the magnetic field strength B. Similarly to the above, this issue arises due to competing effects of high-temperature sources of ionization (thermal ionization, and thermionic and ion emission) and X-rays. As these sources of free electrons depend quite differently on the disc structure (density, temperature, column density), their combination leads to non-monotonous variations in the electron number density as a function of height above disc midplane. Since the viscous dissipation due to the MRI is sensitive to the ionization levels, the total dissipation can be a non-monotonous function of $z_{\rm surf}$. Since the solution for the vertical disc structure is determined by an equilibrium between an input and an output total heat, this can lead to multiple solutions in $z_{\rm surf}$.

To illustrate this issue Fig. 5.14 shows an example of three thermally-stable solutions for the vertical disc structure at a fixed value of magnetic field strength B that appears in the model for a maximum grain size $a_{\max} = 100\,\mu$m (we can ignore thermally unstable solutions, which are not shown here). Note that the dependence of the overall disc structure and the location of the pressure maximum on the dust grain size are presented and thoroughly discussed in the next

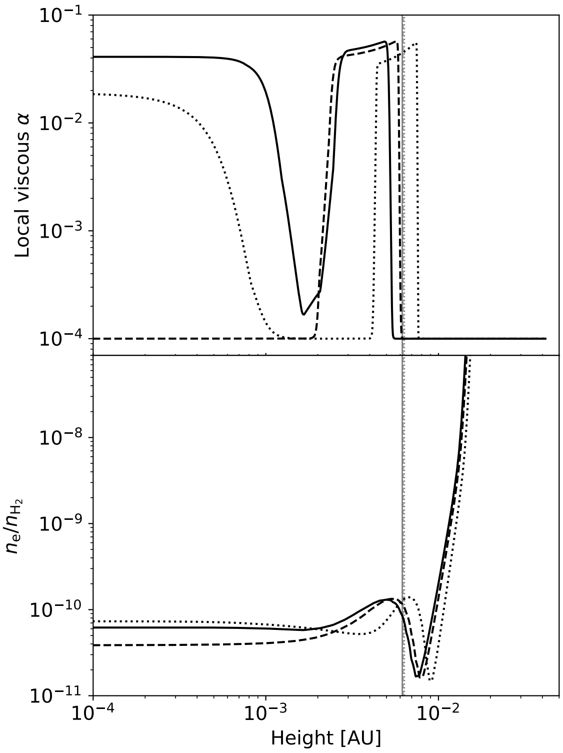


Figure 5.14: Degeneracy in the vertical disc structure at a fixed magnetic field strength in a model where the maximum dust grain size is $a_{\max} = 100\,\mu$m: local viscous parameter α as a function of height (top), and the fractional electron number density (n_e/n_{H_2}; bottom) as functions of height in three different equilibrium solutions. The different solutions arise from vertical variations in the viscous α; the lengthscales of these variations are much smaller than the disc pressure scale height (shown for each solution by the vertical grey lines). See Section 5.3.2.3.

chapter. Here, I only discuss the issue of the multiple solutions. Fig. 5.14 shows the viscosity α as a function of height in the top panel and the ratio n_e/n_{H_2} in the bottom.

Evidently, small variations in the free electron number density correspond to large variations in the viscosity α, all at heights below one disc pressure scale height (indicated by grey lines). This implies that the difference between these solutions is likely unphysical for two reasons. First, the viscosity α is in reality driven by turbulence, and turbulent motions should not abruptly change over lengthscales much smaller than a single pressure scale height. Second, chemical species can also be expected to be spatially mixed by turbulence, and so such vertical variations in the ionization levels might be smoothed over in reality. Since resolving these issues is out of the scope of the models, a solution with minimum $z_{\rm surf}$ is always chosen, which also appears to always correspond to a maximum $\bar{\alpha}$ at the given magnetic field strength.

5.4. Discussion

5.4.1. Effects of dust

A key feature of the inner disc is that the MRI drives high viscosity in the innermost regions, close to the star, but becomes largely suppressed at larger orbital distances (in the so-called dead zone). This leads to a formation of a local gas pressure maximum that may play a key role in planet formation at short orbital distances (Chatterjee and Tan 2014). This decrease in viscosity is expected to arise because the innermost regions are hot enough ($> 1000\,\mathrm{K}$) to thermally ionize potassium (coupling the gas to the magnetic field), but further out temperature and ionization levels decrease substantially (Gammie 1996). In Chapter 2, we have seen that in a thermally-ionized disc coupled self-consistently to an MRI viscosity, the inner edge of the dead zone lies at a few tenths of an AU.

One of the key differences of this work to Chapter 2 is that here the effects of dust on disc ionization levels are also taken into account. Small dust grains present in the disc sweep up free electrons and ions from the gas, and these recombine quickly on the grain surfaces. In the bulk of the protoplanetary disc dust grains efficiently lower the ionization levels, decoupling the magnetic field from the gas and suppressing the MRI (Sano et al. 2000; Ilgner and Nelson 2006; Wardle 2007; Salmeron and Wardle 2008; Bai and Goodman 2009). In the inner regions of protoplanetary discs dust grains also act to increase the ionization levels, as at high temperatures they can also emit electrons and ions into the gas (Desch and Turner 2015). The balance between thermal ionization and these processes then determines how well ionized the inner disc is, and thus, the extent of the high-viscosity region and the location of the dead zone inner edge.

The top panel of Fig. 5.9 shows that addition of dust only weakly affects the vertically-averaged viscosity $\bar{\alpha}$ in the inner disc. At a given orbital radius, $\bar{\alpha}$ is even slightly higher than in the model with no dust, implying that thermionic and ion emission are important sources of ionization. In fact, as Desch and Turner (2015) showed, thermionic emission can become the main source of free electrons at high temperatures. For the disc model discussed here this can be seen in the bottom panel of Fig. 5.11, which shows that at hot disc midplane thermionic emission dominates over other sources of ionization. Clearly then, for the chosen parameters, the adsorption of charges onto grains is more than counteracted by the expulsion of charges from hot grain surfaces.

Similarity in the resulting disc structure in the models with and without dust grains can be explained by the similar dependency on temperature that thermal and thermionic emission have (and which can also be deduced from the bottom panel of Fig. 5.11). As discussed by Desch and Turner (2015), thermal ionization of potassium becomes efficient at temperatures above $\sim 1000\,\mathrm{K}$ in accordance with its ionization potential $\mathrm{IP} = 4.34\,\mathrm{eV}$. The temperature at which thermionic emission becomes important is determined by the work function W of the material out which dust grains are made, and for silicates $W \sim 5\,\mathrm{eV}$. This alone would imply that thermionic emission becomes efficient at temperatures closer to 2000 K. However, above ∼1000 K potassium-bearing minerals start evaporating from grain surfaces (Lodders 2003), and a fraction of potassium atoms leaves the grain surface as ions (since $W \sim \mathrm{IP}$). Dust grains then become negatively charged, which reduces the effective potential that electrons ought to overcome for thermionic emission (the effective work function W_{eff}).

The above results could change significantly as dust grains grow or as they accumulate in the inner disc (as needed for the formation of solid planet cores). Dust adsorption of free charges, for example, becomes much less efficient for larger grains, since the total grain surface area decreases (Sano et al. 2000; Ilgner and Nelson 2006). I investigate how dust growth and varying dust-to-gas ratio affect the inner disc structure in the next chapter.

5.4.2. Importance of stellar irradiation

The absorbed flux of stellar irradiation is many times higher than the heat flux generated by accretion at any given radius, yet irradiation has a very small effect on the disc midplane temperature. Consequently, the ionization levels and the MRI-driven viscosity are similar in the models with and without stellar irradiation.

Essentially, this is because the stellar irradiation heats the disc optically-thin regions, from which heat escapes easily. Accretion heat is generated deep in the disc, where the optical depth is much higher. In the absence of stellar irradiation, the midplane temperature in the optically thick disc is $\sigma_{\mathrm{SB}} T_{\mathrm{mid}}^4 \sim F_{\mathrm{acc}} \tau_{\mathrm{mid}}$ (Hubeny 1990, though in this work the midplane temperature is somewhat lower due to convection). If on top of a viscously-heated layer of optical thickness $\tau_{\mathrm{mid}} \gg 1$ there is an irradiation-heated layer of optical thickness τ_{upper}, from eq. (5.8) it follows that $\sigma_{\mathrm{SB}} T_{\mathrm{mid}}^4 \sim F_{\mathrm{acc}} \tau_{\mathrm{mid}} + F_{\mathrm{irr}} \tau_{\mathrm{upper}}$ (again, neglecting convection). Here τ_{upper} is the optical depth of the disc to its own radiation down to a height at which the disc becomes optically thick to stellar irradiation. Then, if τ_{mid} is sufficiently larger than τ_{upper}, the midplane temperature is determined by viscous dissipation.

The results are consistent with those of D'Alessio et al. (1998), who also found that models with and without stellar irradiation yield roughly the same midplane temperatures in the optically-thick inner disc. Similarly, Flock et al. (2019) considered 2D static radiation-hydrodynamics models of the inner disc heated by stellar irradiation only, but found that the midplane temperature (and the orbital radius of the dead zone inner edge) would increase appreciably if accretion heat were included, on the condition that the accretion heat is released near the optically-thick midplane.

5.4.3. Convective instability in the inner disc

In Section 5.3.1 we have seen that a large region of the inner disc is convectively unstable. This is also the case even when the opacities are constant, i.e., a super-linear growth of the opacity with temperature (Lin and Papaloizou 1980) is not needed. Here, the high temperature gradient is established because the heat is deposited deep within the optically-thick disc. In the presented models most of the viscous dissipation happens near midplane, where the MRI is active, but the same is also true for a vertically-constant viscosity α. This result can also be confirmed analytically. We can consider a simplified problem of radiative transfer in the optically thick limit, where the temperature is given by $\sigma T^4 = \frac{3}{4}\tau F(z_{\rm surf})$ (Hubeny 1990). Assuming a constant disc opacity, the equation of hydrostatic equilibrium can be re-written as $\frac{dP}{d\tau} = \frac{\Omega^2 z}{\kappa_{\rm R}}$. The temperature gradient is then given by

$$\nabla = \frac{d\ln T}{d\ln P} = \frac{\kappa_{\rm R} P}{4\Omega^2 z\tau}.$$

The appropriate upper boundary condition for this problem is the disc photosphere ($\tau = 2/3$), where the gas pressure is given by $P_{\rm surf} = \Omega^2 z_{\rm surf}\tau_{\rm surf}/\kappa$ (assuming that the disc is vertically isothermal above the photosphere, Papaloizou and Terquem 1999). At the photosphere, given the chosen boundary condition, we have $\nabla = 1/4$. Near the midplane the optical depth is $\tau_{\rm MID} = \frac{1}{2}\kappa\Sigma$, and we may estimate the midplane pressure as

$$P_{\rm MID} = \rho_{\rm MID} c_{\rm s,MID} = \frac{\Sigma}{2H} c_{\rm s,MID} = \frac{1}{2}\Omega^2 \Sigma H,$$

where the disc scale height H is related to the midplane temperature. Substituting $\tau_{\rm MID}$ and $P_{\rm MID}$ into the expression for the temperature gradient, we have

$$\nabla_{\rm MID} = \frac{1}{4}\frac{H}{z},$$

implying that such a disc should become convectively unstable a bit below one scale height. However, we can further estimate the gradient ∇ near one scale height, by assuming that there $P_H \sim P_{\rm MID} e^{-1/2} \sim 0.6 P_{\rm MID}$, and $\tau_H \sim 0.3\tau_{\rm MID}$, as follows from vertically isothermal, Gaussian profiles of pressure and density. Thus, near $z = H$, we have

$$\nabla_H = \frac{1}{2}\frac{H}{z},$$

showing that the disc should become convectively unstable above one scale height.

Furthermore, in the convectively unstable regions I have used a simple approximation that convection is efficient and the temperature gradient is isentropic. More detailed calculations would yield an answer in which the temperature gradient lies between the isentropic one and the gradient given by the radiative transport. As it turns out, the result would not differ much from what is obtained here. In the optically-thick limit considered above, the difference in the temperature profile when the entire flux is transported by radiation and when the entire flux is transported by convection is a very weak function of optical depth, and remains small for rather large optical depths (Cassen 1993).

Limiting the temperature gradient is the only role of convection in the simple viscous model discussed here. In real discs, convection might interact with the MHD turbulence induced by the MRI. For example, in simulations of ideal MHD Bodo et al. (2013) and Hirose et al. (2014) find that convection can increase the angular momentum transport driven by the MRI by increasing the magnetic field strength, although it appears that this is only the case when convection is particularly strong (Hirose 2015). Concurrently, it is found that the relationship between the induced stress and the magnetic field strength are not modified. The consequences for the non-ideal MHD regime, relevant in protoplanetary discs, are not clear. In the solutions presented in this work the value of the vertically-averaged MRI-driven viscosity would decrease with both a decrease and an increase in the magnetic field strength due to non-ideal effects. Convection itself is not expected to drive the angular momentum transport at a level comparable to the MRI (e.g. Lesur and Papaloizou 2010; Held and Latter 2018), and in any case it is not self-sustainable, i.e., it requires additional source of heat near disc midplane to establish the high temperature gradient.

5.4.4. Energy transport by turbulent elements

Ruediger et al. (1988) considered vertical structure models in which the turbulent elements driving the angular momentum transport may also transport energy. Such turbulent energy transport flux would be analogous to convection, transporting energy down the entropy gradient, the difference being that turbulent elements may persist at sub-adiabatic temperature gradients (as they are driven by other instabilities) in which case they transport energy from cooler to hotter regions (Balbus 2000). We can expect that including this mode of energy transport into these calculations would not change appreciably any of the results presented here. D'Alessio et al. (1998) found that the turbulent energy transport accounts for less then 20% of the total energy flux at any given orbital radius. In this work, the convectively stable upper layers of the disc are MRI-dead, and thus the thermal diffusivity due to turbulence is likely very low. In the regions of the disc where the radiative flux alone would yield super-adiabatic temperature gradient, it is already assumed that convection efficiently establishes the adiabat. The proportion of turbulent energy flux could be higher than the convective energy flux in such regions, but the temperature would not change significantly.

5.4.5. Ambipolar diffusion in the strong-coupling regime

The criterion for the ambipolar diffusion to quench the MRI that is employed here is valid in the strong-coupling regime (Bai and Stone 2011). Strong coupling requires that the ionization equilibrium is achieved on a timescale t_{chem} shorter than the dynamical timescale $t_{\mathrm{dyn}} = 2\pi/\Omega$. Previously, in Chapter 2 it was reported that this condition is not fulfilled in most of the inner disc, as slow radiative recombinations make the chemical equilibrium timescale long. However, even in the absence of dust, three-body recombinations (recombinations through collisions

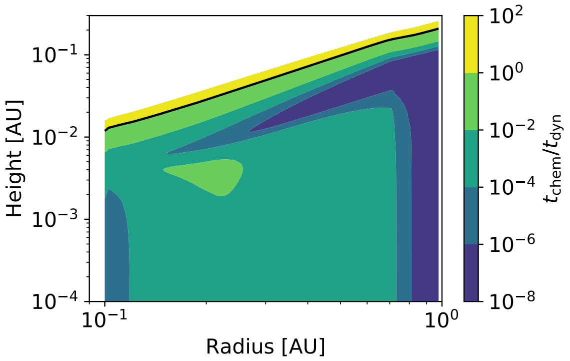


Figure 5.15: Ratio of the shortest recombination timescale ($t_{\rm chem}$) to the dynamical timescale ($t_{\rm dyn}$) as a function of the location in the disc. The solid line indicates where $t_{\rm chem}/t_{\rm dyn} = 1$. Everywhere below this line the chemical equilibrium timescale is shorter than the dynamical timescale, justifying the assumption of the strong-coupling regime. See Section 5.4.5.

with the abundant molecular hydrogen) are much faster than radiative recombinations, and adsorption onto grains is even faster (Desch and Turner 2015).

Since I solve directly for the equilibrium ionization state, I do not have access to the timescale $t_{\rm chem}$. However, we can estimate it as $t_{\rm chem} = n_{\rm e}/\mathcal{R}$, where $\mathcal{R}$ is the fastest of the above three recombination rates (in general, but not always, that is adsorption onto dust grains). Fig. 5.15 shows, for the fully self-consistent model with the full chemical network, that $t_{\rm chem}/t_{\rm dyn} < 1$ everywhere except in the uppermost, lowest-density layers of the disc. It can thus be concluded that the use of the ambipolar diffusion criterion in the strong-coupling regime is justified.

5.5. Conclusions

I present a steady-state model of the inner protoplanetary disc which accretes viscously, primarily due to the MRI. In this model, the disc is heated by viscous dissipation and stellar irradiation, and cools radiatively and convectively. The disc is ionized by thermal ionization, thermionic and ion emission from dust grains and by stellar X-rays, cosmic rays and radionuclides, and I also account for adsorption of charges onto dust grains. The disc structure (density, temperature), viscosity due to the MRI, disc opacity and ionization state are calculated self-consistently everywhere in the disc.

I investigate how these various processes affect the structure of the inner disc and the extent to which the MRI can drive efficient accretion, i.e., the location of the inner edge of the dead zone. I find that, since the inner disc is optically thick, stellar irradiation weakly affects the temperature at midplane, and therefore weakly affects the location of the dead zone inner edge. Furthermore, the inner disc is largely convectively unstable, which is shown to be a property of any optically-thick disc in which viscous heating is released near the midplane.

Dust controls the ionization state of the inner disc, and thus the onset of the MRI. Thermal ionization plays a secondary role, as thermionic and ion emission from dust grains ionize the hot dense regions. High above disc midplane stellar X-rays produce a thin MRI-active layer. However, this changes very little the overall viscosity at short orbital distances.

While there are some qualitative differences in the shape of the dead zone inner edge as a function of radius and height above disc midplane compared to the work discussed in Chapter 2, its radial location remains at roughly the same orbital distance of $\sim 0.7\,\mathrm{AU}$ for the fiducial parameters. This is a consequence of the similarity between the ionization potential of potassium and the work function of the dust grains. How these results depend on the model parameters, including dust grain size and dust-to-gas ratio, is explored in the next chapter.

6 DEPENDENCE ON DUST, DISC AND STELLAR PARAMETERS

6.1. Introduction

A key first step to forming the super-Earths in the inner regions of protoplanetary discs, near their present orbits, is enhancement of the solid component of the inner disc. The inner disc can be enriched by pebbles from the outer disc (Hansen and Murray 2013; Boley and Ford 2013; Chatterjee and Tan 2014; Hu et al. 2018), as pebbles are prone to inwards radial drift due to gas drag (Weidenschilling 1977). It has been hypothesized (Chatterjee and Tan 2014) that the radial drift of pebbles could be stopped at a local gas pressure maximum in the inner disc. Over time the pressure maximum could accumulate enough material to form a super-Earth-sized planet.

The pressure maximum will only trap pebbles which are prone to radial drift relative to the gas. Smaller dust grains that are well coupled to the gas may be advected and diffused through the pressure maximum by the gas accreting onto the star. In the inner disc, the size of dust grains is limited by fragmentation due to relative turbulent velocities (Birnstiel et al. 2010; 2012; Drazkowska et al. 2016). Pebbles that radially drift from the outer to the inner disc become smaller due to fragmentation, and the effect of radial drift weakens. In Chapter 3 it was shown that, in an inner disc in which the gas accretion and the grain turbulent velocities are driven by the MRI, the grains can become small enough to escape the pressure trap, through advection and radial mixing by the turbulent gas. Additionally, it was found that this leads to an accumulation of small dust grains throughout the inner disc, interior to the pressure maximum.

In Chapter 3 the effects of dust onto gas were not taken into account, and it was speculated that the accumulation of dust would quench the MRI, lowering the levels of turbulence and allowing some grain growth. Dust grains are expected to quench the MRI by adsorbing free charges from the gas phase (Sano et al. 2000; Ilgner and Nelson 2006; Wardle 2007; Salmeron and Wardle 2008; Bai and Goodman 2009; Mohanty et al. 2013). However, this would concurrently push the MRI-active region and the pressure maximum inwards, possibly eliminating it from the inner disc. Evidently, the outcome is a function of the size and the abundance of the dust grains.

In Chapter 5 I presented a model of a steady-state viscously accreting disc which includes both the MRI-driven viscosity and the effects of dust onto gas. This accounts for the adsorption of free charges onto dust grains, and also for the electron (thermionic) and ion emission from dust grains into the gas phase. The thermionic and ion emission become important at temperatures above $\sim$1000 K (so at the temperatures present in the inner disc) and act to increase the ionization fraction of the gas (Desch and Turner 2015). It was shown that for 1 μm grains, comprising 1% of the disc mass, these dust effects balance out and result in a pressure maximum at roughly the same location as predicted from thermal ionization. Additionally, this model also self-consistently considers the disc opacity due to dust grains, thus taking into account the effect of dust on the disc thermal structure.

In this chapter I investigate how the inner disc structure changes with dust-to-gas ratio, dust grain size, and other disc and stellar parameters, in order to narrow down the region of parameter space where the formation of planetary cores inside the ice line is more likely. I briefly overview the disc model in Section 6.2 and present the results in Section 6.3. Section 6.4 focuses on the existence and the location of the gas pressure maximum as a function of the above parameters, exploring the entire parameter space in detail. In Section 6.5 I discuss the implications of these results for the formation of the super-Earths and Section 6.6 summarizes the conclusions.

6.2. Methods

The disc model used here is presented in Chapter 5. Here, I only summarize the main points. It is assumed that the viscously-accreting disc is in steady-state, i.e., that the gas accretion rate $\dot{M}$ is radially constant, on the basis that the gas accretion rate evolves on a long (Myr) timescale. The disc structure is calculated self-consistently with disc opacities, ionization state and the viscosity due to the MRI. The disc is assumed to be in vertical hydrostatic and thermal equilibrium, heated by viscous dissipation and stellar irradiation, and cooled radiatively and/or via convection.

I consider disc opacities due to silicate dust grains. The opacities are calculated for a power-law grain size distribution (Mathis et al. 1977), with a minimum grain size $a_{\rm min} = 0.1\,\mu$m and a maximum grain size $a_{\rm max}$, using optical constants from Draine (2003, see Chapter 5 for details). Fig. 6.1 shows the opacities per unit mass of gas for a set of maximum grain sizes $a_{\rm max}$, assuming a dust-to-gas ratio $f_{\rm dg} = 0.01$, bulk density of dust grains $\rho_{\rm s} = 3.3\,{\rm g\,cm^{-3}}$, and stellar effective temperature $T_* = 4400\,{\rm K}$. Additionally, Fig. 6.2 shows the Planck-mean and Rosseland-mean opacities at a fixed temperature of 500 K, as functions of maximum grain size. The structure of the disc is considered only beyond the silicate sublimation line, and so the opacities due to gas molecular and atomic lines are neglected. Additionally, the contribution from the water ice and carbonaceous grains is neglected, since their sublimation temperatures are much lower than the temperatures expected in the hot ($> 1000\,\mathrm{K}$) MRI-active regions (e.g. Pollack et al. 1994).

Furthermore, the disc ionization state is calculated using a simple chemical network (Desch and Turner 2015) that includes thermal (collisional) ionization of potassium; ionization of molecular hydrogen by stellar X-rays, cosmic rays and radionuclides (producing metal (magnesium) ions by charge transfer); gas-phase recombinations; adsorption onto dust grains and thermionic and ion emission from dust grains. Only a single dust grain species of size $a_{\mathrm{gr}} = 0.1\,\mu\mathrm{m}$ is considered in the chemical network, but an effective dust-to-gas ratio f_{eff} is chosen to mimic the full size distribution stated above (see Chapter 5 for details).

Finally, the viscosity due to the MRI, parametrized using the Shakura and Sunyaev (1973) α parameter, is calculated using a prescription based on the results of magnetohydrodynamic simulations (see Chapter 5). This accounts for the suppression of the MRI by Ohmic and ambipolar diffusion. In the MRI-dead zones (where the MRI is suppressed), I assume the gas can still accrete due to a small constant viscosity parameter α_{DZ}, induced either by the adjacent MRI-active zone or by purely hydrodynamical instabilities. The dead-zone viscosity parameter α_{DZ} plays an important role in the results presented here. The viscous α is calculated both as a function of radius and height, and a vertically-averaged viscosity parameter is defined as

$$\bar{\alpha} \equiv \frac{\int_0^{z_{\mathrm{surf}}} \alpha P dz}{\int_0^{z_{\mathrm{surf}}} P dz}, \tag{6.1}$$

where z_{surf} is the disc height, defined as a height above which the gas pressure falls below some small constant value.

6.3. Results

As the fiducial model, I consider a disc with a gas accretion rate $\dot{M} = 10^{-8}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, stellar mass $M_* = 1\,\mathrm{M}_{\odot}$, stellar radius $R_* = 3\,\mathrm{R}_{\odot}$, effective stellar temperature $T_* = 4400\,\mathrm{K}$[1], viscosity parameter in the MRI-dead zone $\alpha_{\mathrm{DZ}} = 10^{-4}$, dust-to-gas ratio $f_{\mathrm{dg}} = 10^{-2}$, and maximum dust grain size $a_{\mathrm{max}} = 10^{-4}\,\mathrm{cm}$. In Chapter 5 I used this fiducial model to discuss the impact of various physical and chemical processes on the inner disc structure. Here, I explore the effects of varying these parameters on the inner disc structure. First, Section 6.3.1 details the effects of varying the dust-to-gas ratio f_{dg} and maximum dust grain size a_{max}. Then, Section 6.3.2 details the effects of varying the gas accretion rate $\dot{M}$, stellar mass M_*, and the dead-zone viscosity α_{DZ}.

---

[1] These specific stellar parameters correspond to a solar-mass star at an age of 0.5 Myr in the stellar evolution models of Baraffe et al. (2015). Over the first 5 Myr luminosity of a solar-mass star decreases by a factor of 5 in these models, and so the adopted parameters are roughly valid throughout the disc lifetime. Additionally, in Chapter 5 it was shown that the stellar irradiation has a weak effect on the extent of the MRI-active zone, since the inner disc is optically thick, and so this decrease in the luminosity is unimportant.

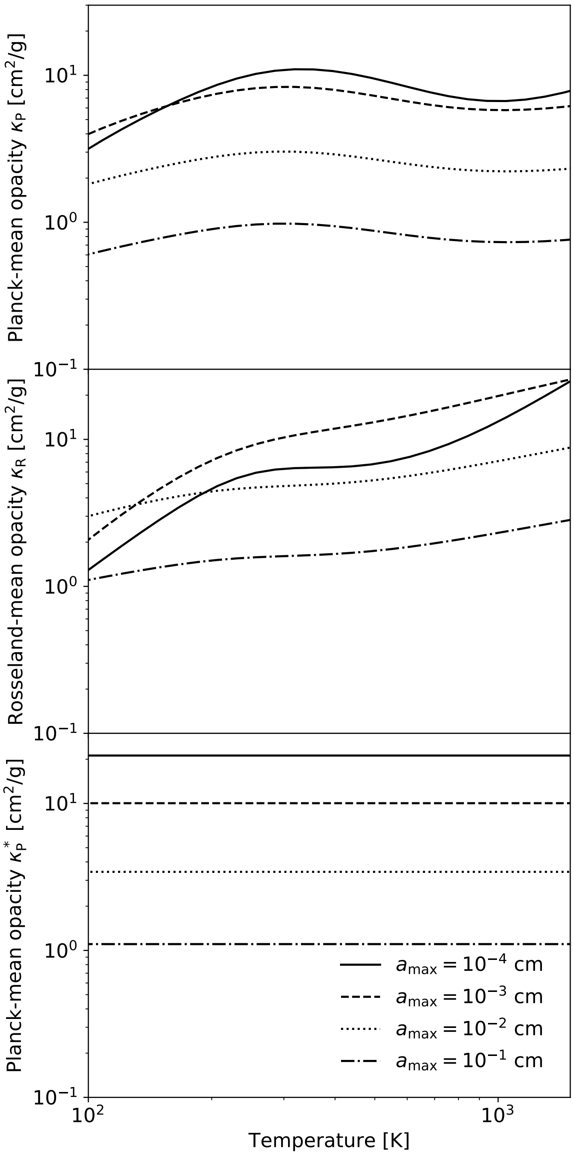


Figure 6.1: Planck-mean opacity κ_{P} (top), Rosseland-mean opacity κ_{R} (middle), and Planck-mean opacity at the stellar effective temperature κ_{P}^{*} (i.e., the absorption coefficient for the stellar irradiation; bottom), as functions of disc temperature, for different maximum grain sizes a_{max} as indicated in plot legend, assuming a dust-to-gas ratio $f_{\mathrm{dg}} = 0.01$ and stellar effective temperature $T_{*} = 4400\,\mathrm{K}$. Absorption is dominated by small grains, and so the Planck-mean opacities decrease with increasing maximum grain size. The Rosseland-mean opacity is a non-monotonous function of grain size for $a_{\mathrm{max}} \lesssim 10^{-2}\,\mathrm{cm}$.

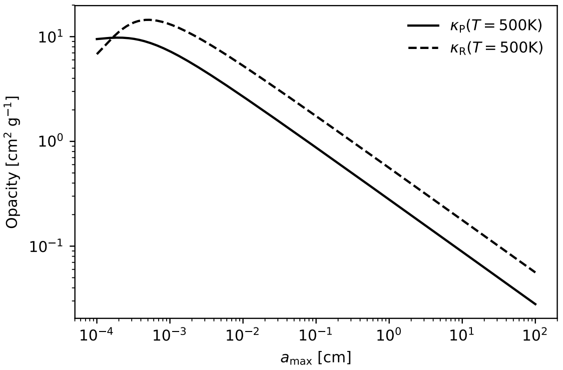


Figure 6.2: Planck-mean opacity κ_P and Rosseland-mean opacity κ_R at a temperature of 500 K, as functions of maximum grain size a_{max}, assuming a dust-to-gas ratio $f_{dg} = 0.01$. The Rosseland-mean opacity peaks around $a_{max} \sim 10^{-3}$ cm at this temperature.

6.3.1. Dust-to-gas ratio and dust size

Dust has two effects on the disc structure in this model: it determines opacities in the disc, and it affects the disc ionization state. To better understand the results of varying dust properties, I first consider only the dust opacities. That is, in Section 6.3.1.1, I consider a model with a vastly simplified chemical network, in which the only source of ionization is thermal ionization of potassium and free charges recombine only in the gas phase. Then, in Section 6.3.1.2 I present the results of the full model which also accounts for the adsorption of charges onto dust grains, thermionic and ion emission from dust grains, and ionization of molecular hydrogen by stellar X-rays, cosmic rays and radionuclides.

6.3.1.1. Thermally-ionized model

In this section I consider a model which includes dust opacities, but does not include dust effects on the disc chemistry, nor ionization of molecular hydrogen. Ionization levels are set by thermal ionization of potassium and calculated using the Saha equation. The results of varying the dust-to-gas ratio f_{dg} in this simplified model are shown in the left column of Fig. 6.3.

The top left panel shows the radial profile of the vertically-averaged viscosity parameter $\bar{\alpha}$, defined in eq. (6.1), for three different values of the dust-to-gas ratio, $f_{dg} = 10^{-4}$, 10^{-2} and 1, for a constant maximum grain size $a_{max} = 10^{-4}$ cm. In this simplified, thermally-ionized disc model, the MRI is active only at short radii, around the hot disc midplane. Therefore, the viscosity is highest in the innermost region where the midplane ionization fraction and the midplane temperature (shown in the second and third row, respectively) are highest, and it decreases with distance from the star. At some radius ionization fraction drops below that needed to sustain the MRI, and the viscosity parameter $\bar{\alpha}$ falls to the minimum, dead-zone value α_{DZ}. That determines the location of the local gas pressure maximum, shown in the bottom panel.

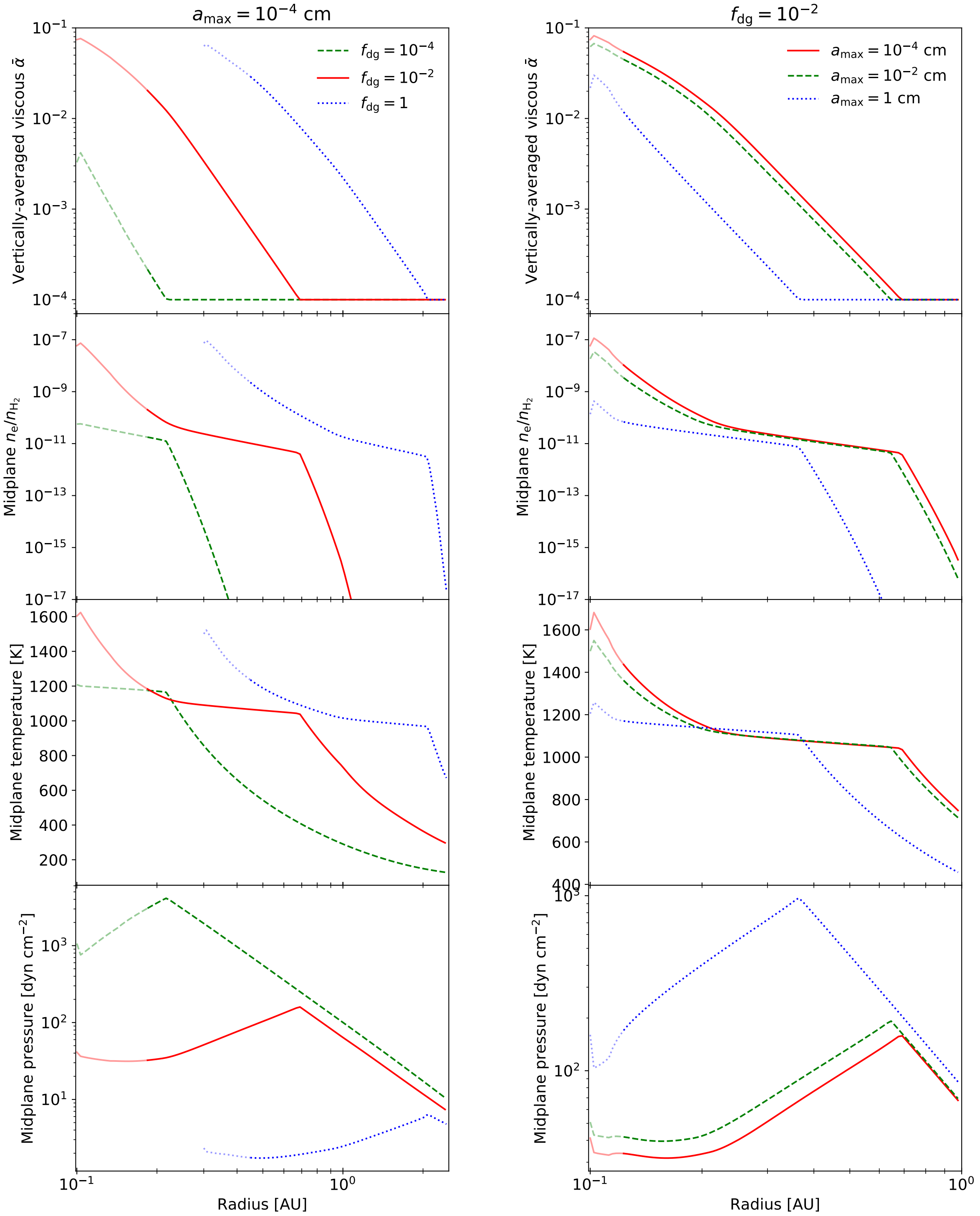


Figure 6.3: Results for the thermally-ionized disc (where the only source of ionization is thermal ionization of potassium, and dust effects on the disc chemistry are not included). The left column shows models with a constant maximum grain size $a_{\max} = 10^{-4}$ cm and varying dust-to-gas ratio $f_{\rm dg}$ as indicated in plot legend. The right column shows models with a constant dust-to-gas ratio $f_{\rm dg} = 10^{-2}$ and varying maximum grain size $a_{\max}$ as indicated in plot legend. The rows show radial profiles of (from top to bottom) vertically-averaged viscosity parameter $\bar{\alpha}$, midplane free electron fraction $n_{\rm e}/n_{\rm H_2}$, midplane temperature and midplane pressure. The inner edge of the $f_{\rm dg} = 1$ model is set to 0.3 AU, since radially inwards temperature increases above the sublimation temperature of silicates. The light-coloured lines indicate the regions affected by the inner boundary condition (see Chapter 5). See Section 6.3.1.1.

Fig. 6.3 shows that a higher dust-to-gas ratio results in a larger MRI-active zone (the region where $\bar{\alpha} > \alpha_{\rm DZ}$). This is because the inner disc is optically-thick, so the disc midplane temperature is set by the accretion heat released near midplane and the optical depth of the disc to its own radiation (with a caveat that the vertical temperature gradient is additionally limited by convection; see Chapter 5). The disc's opacity is directly proportional to the dust-to-gas ratio, and so a disc with more dust is more optically-thick, trapping more heat and making the disc midplane hotter. This makes the midplane more ionized, leading to a higher MRI-induced viscosity at a given radius.

Furthermore, the right column of Fig. 6.3 shows models with a constant dust-to-gas ratio of $f_{\rm dg} = 10^{-2}$, but varying the maximum dust grain size $a_{\rm max}$ in the range $10^{-4} - 1\,{\rm cm}$. The radial profiles of the viscosity parameter $\bar{\alpha}$ (the top right panel) show that the MRI-active region becomes smaller if dust grains are larger, at fixed dust-to-gas ratio. This is because larger dust grains have lower opacities (see Fig. 6.1), making the inner disc less optically thick. Same as in a disc with a lower dust-to-gas ratio, this makes the disc midplane cooler and less ionized. The effect on the disc temperature is clearly seen in the outer regions in which the MRI is dead and the viscosity parameter is constant, $\bar{\alpha} = \alpha_{\rm DZ}$ (the right-hand side panel in the third row). Going inwards, the MRI will become active roughly where the midplane temperature reaches $\sim 1000\,{\rm K}$, as necessary for the thermal ionization of potassium. Therefore, in these simplified, thermally-ionized models, dust growth results in the MRI-active zone edge being pushed inwards. Note that it is the effects of dust opacities alone that lead to these significant changes in the extent of the MRI-active zone. The effects of including dust grains in the chemical network are explored next.

6.3.1.2. *Full model*

In this section, I consider the full model that additionally includes (direct) effects of dust on the ionization fraction (adsorption of free charges onto dust grains, thermionic and ion emission), and also ionization of molecular hydrogen. Three sources of ionization are considered for the latter (stellar X-rays, cosmic rays and radionuclides), the X-rays being the most important (see Chapter 5). The results of varying the dust-to-gas ratio and maximum dust grain size in this full model are shown in the left and the right column of Fig. 6.4, respectively. From top to bottom, the figure shows the radial profiles of the vertically-averaged viscosity parameter, midplane free electron fraction, midplane temperature and midplane pressure.

As in the simplified, thermally-ionized models discussed above, in the innermost regions the viscosity parameter decreases with distance from the star. Here, however, the viscosity parameter reaches a minimum value close to the dead-zone viscosity, and then increases again radially outwards, due to ionization by stellar X-rays (this is true in all models, even if not always evident in the plots). Note that in the full model, in the innermost regions the main sources of ionization are thermionic and ion emission from dust grains, and free charges mostly

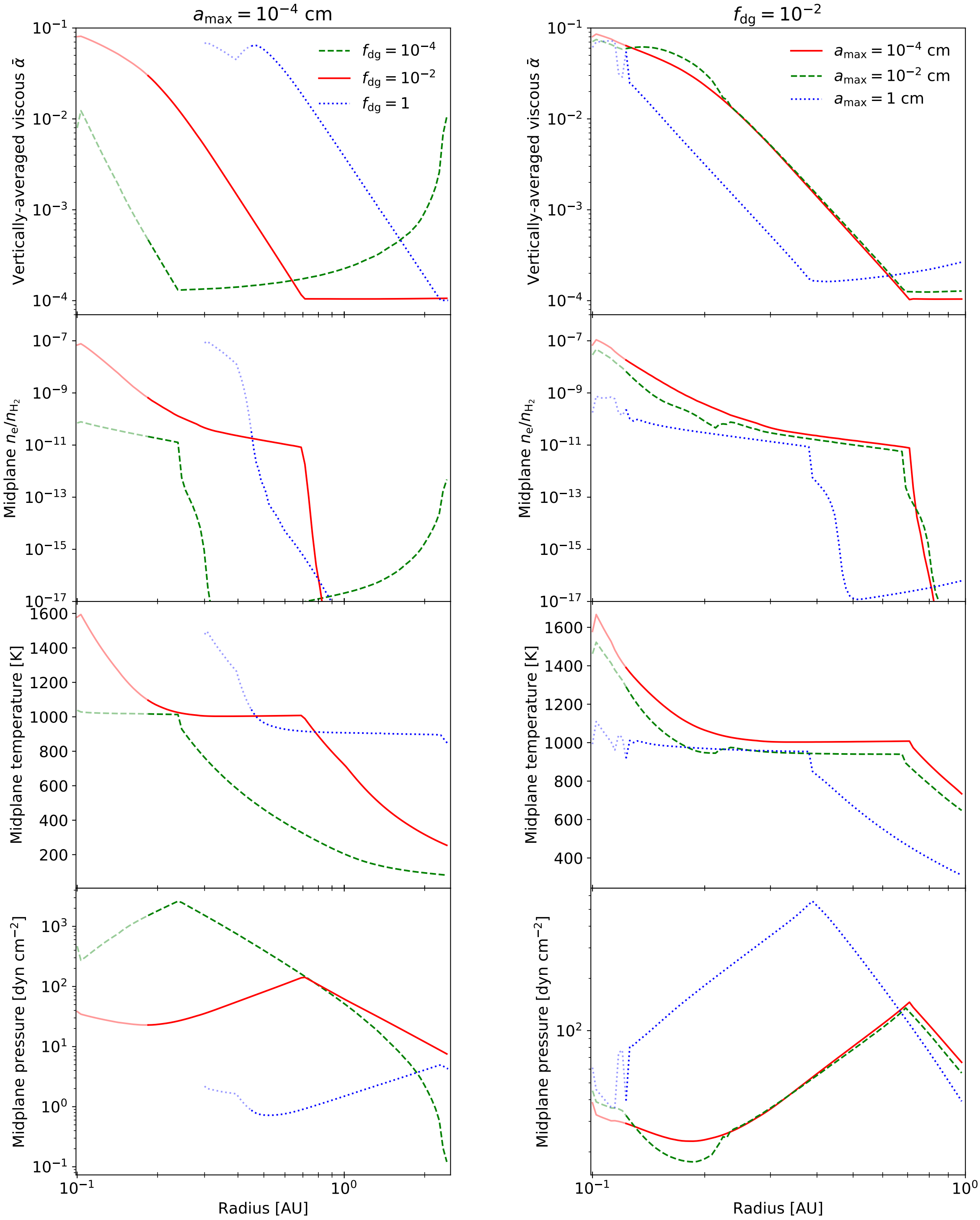


Figure 6.4: Results for the full model which includes dust effects on disc chemistry and ionization of molecular hydrogen. The left column shows models with a constant maximum grain size $a_{\max} = 10^{-4}$ cm and varying dust-to-gas ratio f_{dg} as indicated in plot legend. The right column shows models with a constant dust-to-gas ratio $f_{\mathrm{dg}} = 10^{-2}$ and varying maximum grain size $a_{\max}$ as indicated in plot legend. The rows show radial profiles of (from top to bottom) vertically-averaged viscosity parameter $\bar{\alpha}$, midplane free electron fraction $n_{\mathrm{e}}/n_{\mathrm{H}_2}$, midplane temperature and midplane pressure. The light-coloured lines indicate the regions affected by the inner boundary condition (see Chapter 5). The radius of the pressure maximum is larger for a larger dust-to-gas ratio; it is approximately the same for the maximum grain size of 10^{-4} cm and 10^{-2} cm, but quite smaller for the maximum grain size of 1 cm. See Section 6.3.1.2.

recombine by adsorption onto dust grains (see Chapter 5; Desch and Turner 2015). While this is fundamentally different from the gas-phase thermal ionization, ionization levels as a function of temperature and density are quantitatively similar. As a result, the viscosity parameter in these innermost regions (at high temperatures) is similar to the models with no dust. In Chapter 5 I discussed this for the fiducial maximum grain size $a_{\rm max} = 10^{-4}\,{\rm cm}$. The results presented in this work show that this conclusion holds at a wide range of dust-to-gas ratios and grain sizes (e.g., compare Fig. 6.4 with Fig. 6.3).

An exception to this is the case of high dust-to-gas ratio of $f_{\rm dg} = 1$. In this model, the midplane free electron fraction decreases substantially already at the distance of $\sim$0.4 AU (see the left-hand panel on the second row in Fig. 6.4). However, the viscous $\bar{\alpha}$ remains high out to $\sim$2 AU in this model (the top left panel in Fig. 6.4). In this high-$\bar{\alpha}$ region, the MRI is indeed active at the disc midplane (as shown in the left-hand panel of Fig. 6.5), despite the large decrease in the midplane electron number density. What drives the MRI in this case? The right-hand panel of Fig. 6.5 shows that between $\sim$0.4 AU and $\sim$2 AU the main ionized species are the potassium ions and the dust grains. Evidently, the number density of electrons decreases with increasing dust-to-gas ratio (keeping other parameters fixed), but the opposite is true for the number density of potassium ions evaporating from dust grains, above $\sim$900 K (see Desch and Turner 2015). In the resulting disc ionization state, due to charge conservation, the total charge of potassium ions equals the total charge on dust grains. However, since the dust grains have a much higher inertia than potassium ions, it is the potassium ions that couple the gas to the magnetic field. Overall, these results show that emission of potassium from dust grains is sufficient to sustain the MRI out to large radii, at high dust-to-gas ratios. However, also note that for a dust-to-gas ratio of unity, the dynamical back-reaction of dust on the gas (which is not considered in this work) would become important.

Furthermore, as noted above, in the full model the viscosity parameter $\bar{\alpha}$ reaches a minimum value, outwards from which it increases with radius. The minimum in the viscosity parameter corresponds to the location of the gas pressure maximum, shown in the bottom panels of Fig. 6.4. Outwards of the pressure maximum, temperature at the disc midplane is too low for efficient ionization, and so the MRI is only active in a thin X-ray-ionized layer high above disc midplane (seen near $\sim$2 AU in the left-hand panel of Fig. 6.5). In these outer regions, the viscosity parameter increases with decreasing dust-to-gas ratio and increasing dust grain size. This is because the main source of ionization are the stellar X-rays, and dust only acts as a recombination pathway. For lower dust-to-gas ratios and, equivalently, higher maximum grain size, the total grain surface area (onto which free charges adsorb) decreases, leading to higher ionization fraction and higher MRI-driven viscosity (Sano et al. 2000; Ilgner and Nelson 2006). Nevertheless, the contribution from the thin X-ray-ionized layer is low in all cases and the viscosity parameter $\bar{\alpha} \sim \alpha_{\rm DZ}$ in this region. In other words, in these outer regions the gas primarily accretes through the dead zone, and the stellar X-rays do not affect strongly the location of the pressure maximum.

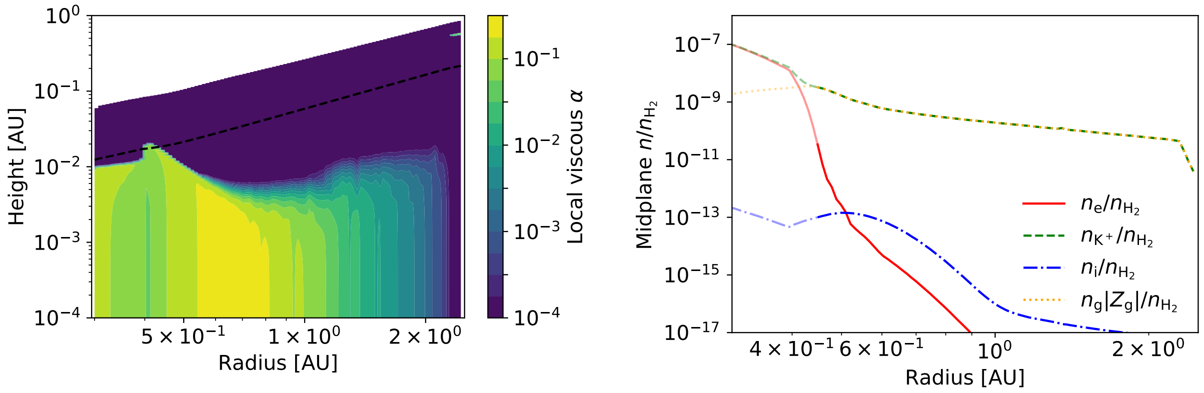


Figure 6.5: Local viscosity parameter α as a function of location in the disc (left) and the fractional number densities of various ionized species at the disc midplane as a function of radius (right), in a model with a dust-to-gas ratio of $f_{\mathrm{dg}} = 1$. See Section 6.3.1.2.

Overall, the extent of the high-viscosity inner region and the location of the pressure maximum seem to be dictated by the dependence of disc opacities on the dust-to-gas ratio and dust grain size through the effects discussed in the previous section. Note, however, that in the full model the location of the pressure maximum is approximately the same for $a_{\max} = 10^{-4}$ cm and $a_{\max} = 10^{-2}$ cm (see the top right panel in Fig. 6.4). This is because a larger grain size results in a lower effective dust-to-gas ratio in the chemical network. This somewhat decreases the critical temperature at which the thermionic and ion emission make the gas sufficiently ionized to start the MRI at disc midplane. In addition, larger grain sizes increase the minimum value of the viscosity parameter (set by the dead-zone value and the stellar X-rays), which moves the pressure maximum inwards. Concurrently, a decrease in the disc's opacity with grain size is moderate for $a_{\max} \lesssim 10^{-2}$ cm (see Fig. 6.1). The location of the pressure maximum as a function of dust grain size is considered in more detail in Section 6.4.

6.3.2. Gas accretion rate, stellar mass and dead-zone viscosity

In this section I investigate how the structure of the inner disc changes with varying gas accretion rate, stellar mass and dead-zone viscosity. Fig. 6.6 shows the results of the fiducial model and three other models in which these three parameters are varied. The different panels show, from top to bottom, the vertically-averaged viscosity parameter ($\bar{\alpha}$), midplane free electron fraction ($n_{\mathrm{e}}/n_{\mathrm{H_2}}$), midplane temperature and midplane pressure, as functions of radius.

In each panel the dashed line shows a model with a gas accretion rate $\dot{M} = 10^{-9}\,\mathrm{M_\odot\,yr^{-1}}$, lower than in the fiducial model with $\dot{M} = 10^{-8}\,\mathrm{M_\odot\,yr^{-1}}$, shown by the solid line. The lower gas accretion rate results in a smaller high-viscosity inner region, and a gas pressure maximum at a shorter radius. This is because the gas accretion rate determines the total viscous dissipation at any given radius. In the optically-thick inner disc, this sets the midplane temperature, the midplane ionization fraction and the viscosity. From Fig. 6.6, the radius of the gas pressure maximum scales with the gas accretion rate approximately as $r_{P_{\rm max}} \propto \dot{M}^{1/2}$. This is, in fact, the same scaling previously found in Chapter 2, where heating by stellar irradiation, dust effects and ionization of molecular hydrogen were neglected. While stellar irradiation and ionization of molecular hydrogen are indeed unimportant for the model parameters chosen here, the dust effects are not. However, while the chemistry setting the ionization state of the disc is qualitatively different in those simple models (thermal ionization of potassium) and in the models presented here (thermionic and ion emission), in both cases the ionization fraction increases sharply above roughly the same temperature ($\sim$1000 K), yielding the same approximate scaling.

Note also that the ionization of molecular hydrogen (predominantly by the stellar X-rays) is not entirely negligible. X-rays activate the MRI in a thin layer high above the disc midplane outwards of the pressure maximum, increasing the viscosity parameter $\bar{\alpha}$ compared to the dead-zone value $\alpha_{\rm DZ}$. The top panel of Fig. 6.6 shows that this X-ray-activated accretion becomes more important in the case of the lower gas accretion rate, increasing the viscosity parameter outwards of the pressure maximum by a factor of 2. This is because in these outer regions a lower gas accretion rate results in lower gas surface densities, increasing the relative contribution of the accretion through the X-ray-activated layer relative to the accretion through the dead zone.

Furthermore, the dotted line in Fig. 6.6 shows a model with a dead-zone viscosity $\alpha_{\rm DZ} = 10^{-3}$ (higher than the fiducial $\alpha_{\rm DZ} = 10^{-4}$). As in the simple models discussed in Chapter 2, the exact value of $\alpha_{\rm DZ}$ is unimportant in the innermost, well-ionized region. In the outer regions the accretion stress is dominated by that in the dead zone, and so $\alpha_{\rm DZ}$ determines the minimum vertically-averaged viscosity parameter in the disc. Therefore, a larger $\alpha_{\rm DZ}$ results in the pressure maximum moving radially inwards compared to the fiducial model.

Finally, the dash-dotted line shows a model with a stellar mass of $M_* = 0.1\,{\rm M}_\odot$. Here I adopt a stellar radius of $R_* = 1\,{\rm R}_\odot$ and an effective temperature of $T_* = 2925\,{\rm K}$ [2]. Keeping the other parameters constant, structure of the inner disc surrounding a $M_* = 0.1\,{\rm M}_\odot$ star is merely shifted radially inwards compared to the fiducial model with $M_* = 1\,{\rm M}_\odot$. This is predominantly due to the Keplerian angular velocity being smaller at a given orbital radius around a less massive star, which reduces the total viscous dissipation at the given radius. The resulting approximate scaling $r_{P_{\rm max}} \propto M_*^{-1/3}$ is consistent with the simple models from Chapter 2, stressing again that (for the chosen $\dot{M}$, dust parameters etc.) stellar irradiation and ionization by stellar X-rays are unimportant in setting the location of the pressure maximum.

---

[2] These stellar parameters are taken from the stellar evolution models of Baraffe et al. (2015), adopting a stellar age of 0.5 Myr. For $M_* = 0.1\,{\rm M}_\odot$, the stellar luminosity is predicted to decrease only by a factor of 3 over the first 5 Myr in these models. It is not expected that this decrease is important, since stellar irradiation plays a minor role in the structure of the inner disc even at the adopted maximum value of the luminosity (see the discussion in Chapter 5).

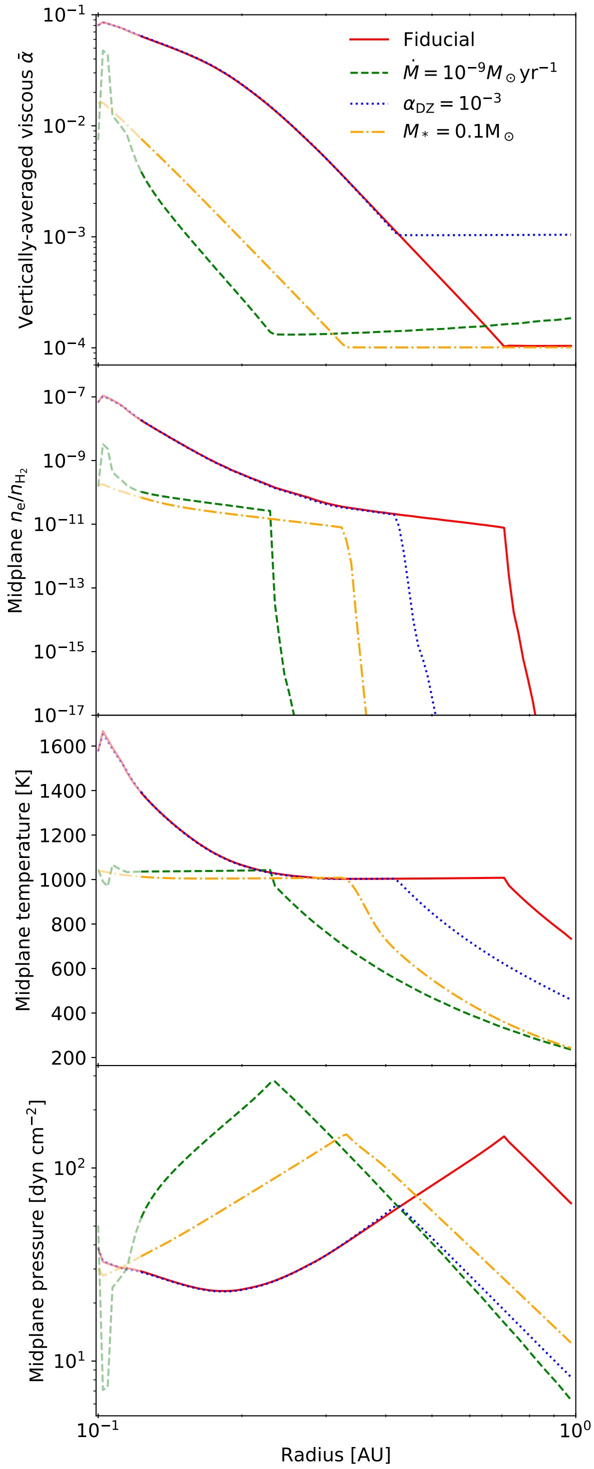


Figure 6.6: Results for the full model which includes dust effects on disc chemistry and ionization of molecular hydrogen. The different panels show radial profiles of (from top to bottom) vertically-averaged viscosity parameter $\bar{\alpha}$, midplane free electron fraction $n_\mathrm{e}/n_{\mathrm{H}_2}$, midplane temperature and midplane pressure. Fiducial model (solid lines) has a gas accretion rate $\dot{M} = 10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, stellar mass $M_* = 1\,\mathrm{M}_\odot$ and dead-zone viscosity $\alpha_{\mathrm{DZ}} = 10^{-4}$. Models with a lower gas accretion rate (dashed line), a smaller stellar mass (dash-dotted line) or a larger dead-zone viscosity (dotted line) all yield the gas pressure maximum at a smaller radius. The light-coloured lines indicate the regions affected by the inner boundary condition (see Chapter 5). See Section 6.3.2.

6.4. Location of the pressure maximum

The above results show that, for a wide range of disc, stellar and dust parameters, the inner disc features a high-viscosity inner region, a low-viscosity outer region, and a gas pressure maximum at the transition between the two regions. This gas pressure maximum has been hypothesized to have a key role in the formation of the super-Earths inside the water ice line (Chatterjee and Tan 2014). Based on the models discussed so far, increasing dust grain size and dead-zone viscosity parameter, and decreasing gas accretion rate and stellar mass result in the pressure maximum moving inwards, towards the star. In this section, I examine in more detail the dependence of the radius of the pressure maximum (and its existence) on these parameters. In order to reduce the computational complexity of the problem and to be able to explore the parameter space in detail, here I neglect heating of the disc by stellar irradiation (but still consider the ionization of the disc by the stellar X-rays). This has been shown to be a good approximation if the inner disc is optically thick (see Chapter 5), but also see the discussion at the end of this section.

Fig. 6.7 shows the radius of the pressure maximum (or pressure bump) as a function of the maximum dust grain size, for the stellar mass $1\,\mathrm{M}_\odot$ (the left-hand panel) and $0.1\,\mathrm{M}_\odot$ (the right-hand panel). In all models shown here, dust-to-gas ratio has the fiducial value of 10^{-2}. The solid, dashed and dotted lines correspond to different values of the dead-zone viscosity parameter ($\alpha_{\mathrm{DZ}} = 10^{-5}$, 10^{-4} and 10^{-3}, respectively). For the different stellar masses different ranges of the gas accretion rate $\dot{M}$ are explored, as indicated in plot legends next to each panel. The chosen ranges are motivated by observational studies, which find that for a solar-mass star typically $\dot{M} \sim 10^{-8}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ and for the stellar mass of $0.1\,\mathrm{M}_\odot$, typically $\dot{M} \sim 10^{-10}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ (e.g. Mohanty et al. 2005; Manara et al. 2012; Alcalá et al. 2014; 2017; Manara et al. 2017). There is a significant spread both in the reported mean values in these studies (± 1 dex for the stellar mass of $0.1\,\mathrm{M}_\odot$, and somewhat less for a solar-mass star) and within the observed samples in each study (up to 2 dex). However, the correlation with the stellar mass appears robust, and so we can adopt the above typical values as mean values and vary the gas accretion by ± 1 dex for each stellar mass.

The radius of the pressure maximum as a function of the maximum dust grain size (a_{max}) shows a similar trend across virtually all values of the stellar mass, accretion rate and dead-zone viscosity: it weakly increases with increasing a_{max} for small grains, peaks at about $a_{\mathrm{max}} \sim 10^{-2}$ cm, and steadily decreases for larger dust grains. The factors causing this have already been briefly discussed in Section 6.3.1. First, recall that a decrease in the disc opacity means that accretion heat can escape more easily, making the disc midplane colder and pushing the pressure maximum inwards. For small dust grains, larger dust grain size leads to a moderate increase in the opacity (here, the relevant opacity is the opacity of the disc to its own radiation in the optically-thick regions, i.e., the Rosseland-mean opacity, see Fig. 6.1). At the same time, the increase in dust grain size reduces the critical temperature at which ionization fraction rises due

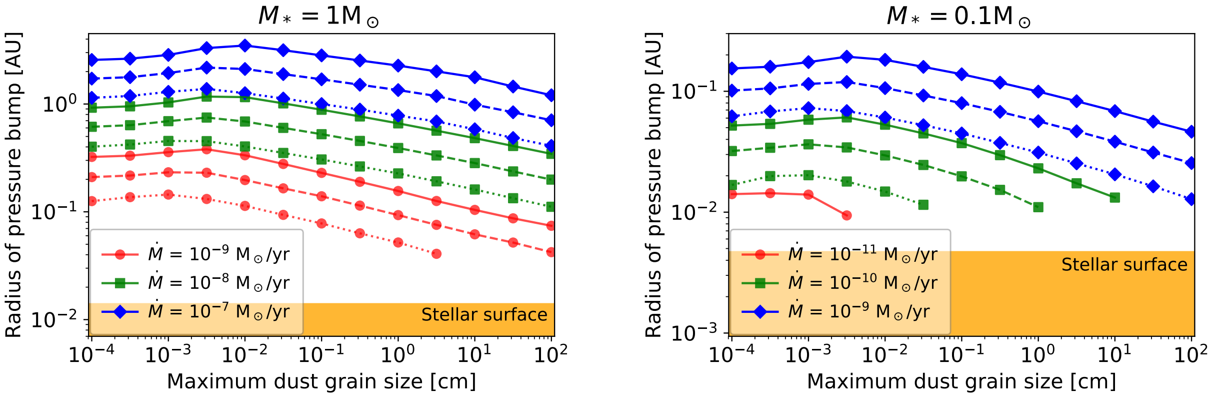


Figure 6.7: Radius of the gas pressure bump as a function of maximum dust grain size. Left-hand panel shows the results for a stellar mass $M_* = 1\,\mathrm{M}_\odot$, right-hand panel for $M_* = 0.1\,\mathrm{M}_\odot$. In all models the dust-to-gas ratio is 10^{-2}. Solid, dashed and dotted lines show results for the dead-zone viscosity parameter $\alpha_{\mathrm{DZ}} = 10^{-5}$, 10^{-4} and 10^{-3}, respectively. Blue, green and red lines indicate different gas accretion rates $\dot{M}$ as indicated in plot legends for each panel. Keeping other parameters fixed, the radius of the pressure bump is a non-monotonic function of the maximum grain size, and peaks at around $\sim 10^{-2}$ cm. For $M_* = 0.1\,\mathrm{M}_\odot$, there is no pressure maximum for the low accretion rate and $\alpha_{\mathrm{DZ}} > 10^{-5}$, for the stellar radius adopted here (indicated by the orange region). See Section 6.4.

to thermionic and ion emission (as the increase in dust grain size is equivalent to a reduction in the effective dust-to-gas ratio in the chemical network; see also Desch and Turner 2015), pushing the pressure maximum outwards. Concurrently, the minimum value of the viscosity parameter is larger for larger grains, pushing the pressure maximum inwards. When all these factors are compounded, for small grains, the radius of the pressure maximum increases with a_{max}. However, if the grains grow beyond $a_{\mathrm{max}} \sim 10^{-2}$ cm, disc opacities are severely reduced and the net effect is a decrease in the radius of the pressure maximum. Additionally, note that the exact value of a_{max} at which the radius of the pressure bump peaks somewhat varies with $\dot{M}$ and α_{DZ}; this can be expected, since both determine the gas surface density and thus the optical depth at disc midplane and the relative importance of the above three effects. It is also important to note that including heating due to stellar irradiation seems to somewhat modify this trend, as can be seen from the comparison of the models discussed in the previous section (specifically, in Fig. 6.4, where for the fiducial disc and stellar parameters, micron-size grains result in a pressure maximum at a larger radius than 100-micron grains) and Fig. 6.7. The resulting differences in the radial location of the pressure maximum are, however, small.

For the solar-mass star, within the observationally-motivated range of $\dot{M}$ and a wide range of α_{DZ}, the pressure maximum is found to exist at $r_{P_{\mathrm{max}}} \gtrsim 0.04$ AU for a wide range of grain sizes. On the other hand, for $M_* = 0.1\,\mathrm{M}_\odot$ both the smaller orbital velocity and the observationally-determined lower gas accretion rates lead to significantly lower $r_{P_{\mathrm{max}}}$. For the gas accretion rate $\dot{M} = 10^{-11}$ and $\alpha_{\mathrm{DZ}} = 10^{-5}$ the pressure maximum does not exist if grains grow larger than few $\times 10^{-3}$ cm. For the same gas accretion rate and higher α_{DZ}, no pressure maximum exists at all.

Finally, recall that in the models shown in Fig. 6.7 heating by stellar irradiation has been neglected. One the one hand, including stellar irradiation could increase the disc midplane temperature by a small amount in these optically-thick solutions, and move the pressure maximum slightly outwards. This would be a small effect, as can be verified for the sets of parameters that were also considered in the previous section (e.g., results including stellar irradiation are shown for the fiducial disc and stellar parameters, and maximum grain sizes up to 1 cm, in Fig. 6.4). On the other hand, for some parameters, including stellar irradiation could also lead to additional or modified solutions that are optically thin. Specifically, for lower gas accretion rates, an equilibrium solution could exist with a low gas surface density if the density was low enough for the X-rays to ionize the entire disc column and produce a much higher accretion efficiency. This is especially the case if the dust grains are larger, as this lowers the ion recombination rate and increases the ionization fraction. Since the accretion rate is proportional to the product of the gas surface density and the viscosity, such a solution also requires the disc temperature to be high enough to match the required accretion stress. In the viscously-heated solutions considered in this section such solutions are not found (in the vicinity of the pressure maximum), since the temperature in a viscously-heated optically-thin disc is low. However, in an optically thin disc, the temperature would be set by stellar irradiation. If the temperature due to irradiation is high enough, for some parameters, a disc in steady state would adapt to this X-ray-ionized solution, and no pressure maximum would exist. Therefore, this issue should be studied further in the future.

6.5. Discussion

In this work I have investigated how the structure of the inner disc, accreting primarily through the MRI, changes with various disc, stellar and dust parameters. Of particular interest are the existence and the location of a local gas pressure maximum and a highly-turbulent region inwards of it, which could accumulate dust grains radially drifting from the outer disc, possibly leading to the formation of planetary cores (Chatterjee and Tan 2014; Hu et al. 2018). The models presented in this work are steady-state models, each with a distribution of dust grains that is fixed throughout the disc. However, as discussed in the rest of this section, these models provide us with important insights into how the inner disc could evolve as the dust grains grow, if and how the dust will accumulate, how this accumulation could feedback on the gas structure, and the disc parameters that are favourable for the formation of planetary cores.

6.5.1. Dust growth

Dust growth to a few tens of microns increases the extent of the high-viscosity inner region and the radius of the pressure maximum, as an increase in dust grain size leads to a moderate increase in the disc opacity and a decrease in the threshold temperature at which thermionic and ion emission become efficient. Growth beyond that size has the opposite effect, as it leads to a significant decrease in the disc opacity, making the disc midplane colder, and thus less ionized. Therefore, in the inner disc, if dust grows larger than $\sim 10^{-2}$ cm sizes, the dead-zone inner edge expands inwards.

Note that this is the opposite of what happens in the outer regions of protoplanetary discs. The outer regions are ionized primarily by the stellar X-rays and cosmic rays. These sources of ionization become more important further away from the star, as the disc column density decreases and the high-temperature effects become unimportant. These regions are expected to be optically thin to their own radiation, and the primary source of heat is stellar irradiation. Therefore, the dust acts primarily to lower the ionization fraction by adsorbing free charges from the gas. Because of this, in the outer regions the dead zone is expected to shrink as the dust grains grow (Sano et al. 2000; Ilgner and Nelson 2006).

In the inner disc, dust growth is limited by collisional fragmentation of dust grains due to relative turbulent velocities (Birnstiel et al. 2010; 2012; Drazkowska et al. 2016). We can consider the location of the pressure maximum under an assumption that the maximum dust grain size has reached this limit, with the relative grain velocities induced by the MRI-driven turbulence. In a turbulent disc, typical collisional relative velocity between dust grains is given by $V_{\mathrm{dd}}^2 \approx 3V_{\mathrm{g}}^2\mathrm{St}$ (for $\mathrm{St} < 1$, Ormel and Cuzzi 2007b), where V_{g} is the typical turbulent gas velocity (given by $V_{\mathrm{g}}^2 = \alpha c_{\mathrm{s}}^2$) and St is the dust particle Stokes number (the ratio between the particle stopping time and the eddy turnover time, assumed to equal the orbital timescale $1/\Omega$; i.e., the non-dimensional stopping time). There is a critical velocity u_{frag} above which a collision between dust grains results in their fragmentation, rather than sticking/growth. For silicate grains of similar size, $u_{\mathrm{frag}} \sim 1\,\mathrm{m\,s^{-1}}$ (Blum and Münch 1993; Beitz et al. 2011; Schräpler et al. 2012; Bukhari Syed et al. 2017, although note that grains might become more sticky at the high temperatures present in the inner disc, Demirci et al. (2019)). Since Stokes number St is directly related to the grain size, and the collision velocity to St, fragmentation imposes an upper limit on dust growth. At the fragmentation limit (Birnstiel et al. 2009; 2012),

$$\mathrm{St}_{\mathrm{frag}} = \frac{u_{\mathrm{frag}}^2}{3\alpha c_{\mathrm{s}}^2}. \tag{6.2}$$

The exact relationship between the Stokes number and the particle size depends on the relevant drag law (Weidenschilling 1977). Typically, the dust grains in protoplanetary discs are smaller than the mean free path of gas molecules, and therefore couple to the gas according to the Epstein drag law. However, due to the high densities in the inner disc, dust grains may enter the Stokes regime. Importantly, the above approximate expression for the turbulent relative velocity between dust grains (V_{dd}) has been derived under an assumption that St does not depend on the relative velocity between the dust grain and the gas, V_{dg}. This assumption is true for grains in the Epstein drag regime. In the Stokes regime, it is true only if the Reynolds number Re of the particle is less than unity. Therefore, it is necessary to always check that this condition is fulfilled for the particles in the Stokes regime, and that we may employ the above expression for $V_{\rm dd}$. The Reynolds number of a particle itself depends on the velocity $V_{\rm dg}$, for which we can adopt another approximate expression, $V_{\rm dg}^2 = V_{\rm g}^2{\rm St}/(1+{\rm St})$ (Cuzzi and Hogan 2003, note that this expression was derived analytically for $\rm St \ll 1$, but also shown to be applicable for a wide range of St through a comparison with numerical simulations).

The location of the pressure maximum in a disc in which grain growth is limited by fragmentation is calculated as follows. At the location of the pressure maximum for various combinations of stellar mass, accretion rate, dead-zone viscosity parameter and maximum dust grain size (i.e., for every point in Fig. 6.7), the fragmentation limit for the particle Stokes number, $\rm St_{frag}$, is calculated (assuming $\alpha = \bar{\alpha}$), as well as the corresponding grain size, $a_{\rm frag}$ (for an appropriate drag law). Then, for every combination of the stellar mass, the accretion rate and the dead-zone viscosity parameter, using linear interpolation, the grain size is found such that $a_{\rm frag}(a_{\rm max}) = a_{\rm max}$. This yields a corresponding radius of the pressure maximum.

Note that this calculation utilizes models in which the maximum dust grain size is assumed to be constant everywhere in the disc, and so the obtained solutions also formally correspond to models in which the maximum dust grain size is radially constant (and equal to the fragmentation limit $a_{\rm frag}$ at the pressure maximum). In a real disc, the fragmentation limit to which particles can grow would be a function of the turbulence levels and other parameters which vary as functions of radius. While this calculation does not take this radial variation of dust size into account, the fragmentation limit at the pressure maximum and the location of the pressure maximum would remain the same as in the solutions found here. In particular, note that radially inwards from the pressure maximum, $a_{\rm frag}$ should decrease compared to the value at the pressure maximum, as the viscosity parameter $\bar{\alpha}$ increases. As we will find, the solution for $a_{\rm frag}$ at the pressure maximum is always higher than $a_{\rm max}$ at which the radius of the pressure maximum peaks (as a function of $a_{\rm max}$, in Fig. 6.7). That is, at a fixed radius, a decrease in particle size inwards of the pressure maximum would yield an increase in the viscosity parameter, compared to the value calculated using the fragmentation limit at the pressure maximum. Accounting for the radial variation in particle size would then make the radial gradient of the viscosity parameter steeper inwards of the pressure maximum, but it would not change the location of its minimum, and therefore not the location of the pressure maximum obtained here.

The results for the radius of the pressure maximum and the grain size are shown in Fig. 6.8, as functions of the gas accretion rate, for different values of stellar mass and dead-zone viscosity parameter. The maximum grain size limited by turbulent fragmentation (middle panel) is sensitive to the dead-zone viscosity (and turbulence) parameter $\alpha_{\rm DZ}$ (results for $\alpha_{\rm DZ} = 10^{-5}$, 10^{-4} and 10^{-3} are shown by the solid, dashed and dotted lines, respectively). This is because the location of the pressure maximum corresponds to the location of the minimum in the vertically-averaged viscosity (and turbulence) parameter $\bar{\alpha}$; since outwards from the pressure maximum the disc primarily accretes through the dense MRI-dead region around the midplane, at the pressure maximum $\bar{\alpha} \sim \alpha_{\rm DZ}$.

Remarkably, despite the sensitivity of the grain size to $\alpha_{\rm DZ}$, the radius of the pressure maximum depends very weakly on this parameter. Basically, the fragmentation-limited solution for the grain size is always larger than the grain size at which the radius of the pressure maximum peaks as a function of the grain size (see Fig. 6.7). Therefore, in this region of the parameter space, a larger grain size yields a smaller radius of the pressure bump. Larger (fragmentation-limited) grain size is obtained for lower values of the dead-zone viscosity parameter, as discussed above. Concurrently, at a fixed grain size (and other parameters), a lower dead-zone viscosity parameter yields a larger radius of the pressure bump (see Section 6.3.2). Evidently, compounding these inter-dependencies results in a weakly-varying radius of the pressure bump for a wide range of values for the dead-zone viscosity parameter.

Furthermore, for the solar-mass star, the pressure maximum is located at radii $\gtrsim 0.1\,{\rm AU}$ in the entire range of the observationally-motivated gas accretion rates considered here. Compounding this with the result that a higher dust-to-gas ratio (resulting from accumulation of dust, see also the next section) would move the pressure maximum outwards, this places the potentially planet-forming region within the range of the observed orbital distances of the super-Earths.

For a lower-mass star, $M_* = 0.1\,{\rm M}_\odot$, at a fixed gas accretion rate, dust grains may grow to somewhat larger sizes than for the solar-mass star. Overall, the radius of the pressure maximum is expected to be smaller, due to the larger grain sizes, but also (mostly) due to the lower viscous dissipation at a given accretion rate, and the lower observed gas accretion rates (see the discussions in sections 6.3.2 and 6.4). As the radius of the pressure maximum moves inwards for lower gas accretion rates, no high-viscosity inner region and no pressure maximum are found at the low end of the observed accretion rates (i.e., for $\dot{M} = 10^{-11}\,{\rm M}_\odot\,{\rm yr}^{-1}$). This implies that planet formation around these stars is more likely earlier in the disc lifetime.

6.5.2. Dust accumulation

For dust grains to become trapped within the pressure maximum, the outwards radial drift velocity of dust grains just inwards of the pressure maximum should be higher than the velocity with which the accreting gas advects the grains inwards. The ratio of the radial drift and the gas advection velocities is roughly equal to the ratio of the particle Stokes number and the viscous α (Jacquet et al. 2012). Therefore, for the particle radial drift to overcome advection with the gas inwards of the pressure maximum, it is required that ${\rm St}/\alpha > 1$. To check whether this condition is fulfilled, we can consider the ratio between the Stokes number at the fragmentation limit ${\rm St}_{\rm frag}$ and the viscosity and turbulence parameter $\bar{\alpha}$. The results are shown in the bottom panel of Fig. 6.8.

The ratio ${\rm St}_{\rm frag}/\bar{\alpha}$ is most sensitive to the value of the dead-zone viscosity parameter $\alpha_{\rm DZ}$.

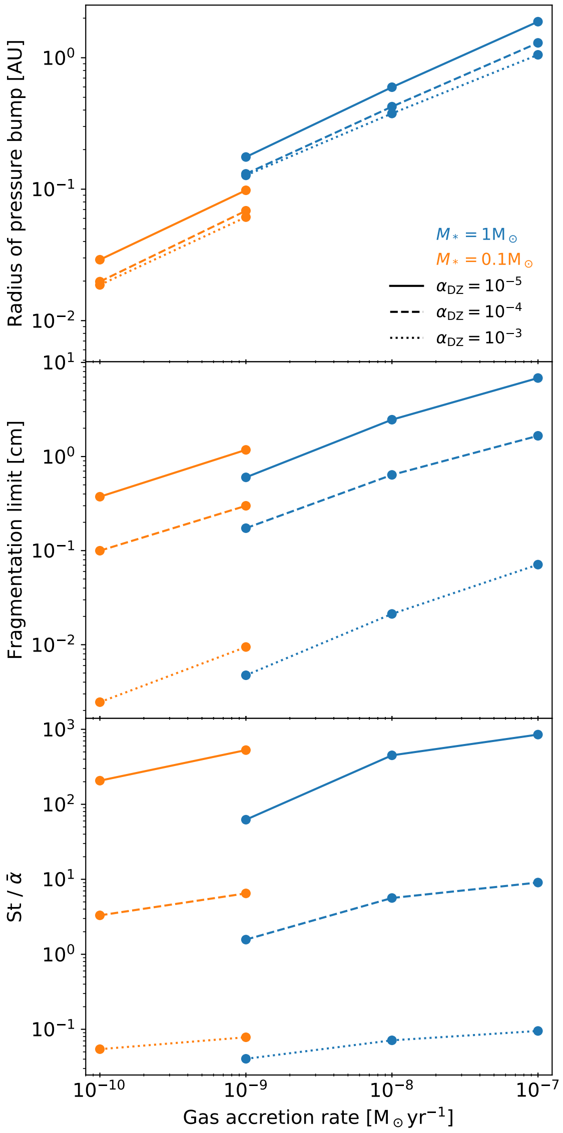


Figure 6.8: Radius of the pressure bump (top), maximum dust grain size (middle) and ratio of the Stokes number to the viscosity parameter ($\mathrm{St}/\bar{\alpha}$) at the pressure bump (bottom) as functions of the gas accretion rate, for a maximum dust grain size that corresponds to the grain growth limit due to turbulent fragmentation. Solid, dashed and dotted lines correspond to different values of the dead-zone viscosity parameter α_{DZ} as shown in the plot legend. The dust grain size and $\mathrm{St}/\bar{\alpha}$ are most sensitive to the dead-zone viscosity parameter α_{DZ}. Concurrently, varying α_{DZ} weakly affects the location of the pressure bump. Blue lines show the results for a stellar mass $M_* = 1\,\mathrm{M}_\odot$, orange lines for $M_* = 0.1\,\mathrm{M}_\odot$. There is no solution for $M_* = 0.1\,\mathrm{M}_\odot$ at the accretion rate of $\dot{M} = 10^{-11}\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, which is at the lower end of the observed rates for this stellar mass. See Sections 6.5.1 and 6.5.2.

The Stokes number at the fragmentation limit $\mathrm{St}_{\mathrm{frag}}$ is given by eq. (6.2). At a fixed critical fragmentation velocity u_{frag}, it is a function only of the viscosity parameter $\bar{\alpha}$ and the speed of sound c_{s} (i.e., the temperature) at the pressure maximum. The viscosity parameter $\bar{\alpha}$ at this location is determined by the accretion through the dead zone at disc midplane and the accretion through a thin X-ray ionized layer high above the midplane. The former dominates due to the density at midplane being much higher, and so $\bar{\alpha} \sim \alpha_{\mathrm{DZ}}$. Therefore, $\mathrm{St}_{\mathrm{frag}}/\bar{\alpha} \propto \alpha_{\mathrm{DZ}}^{-2}$.

Furthermore, the temperature at the pressure maximum is expected to be roughly ~1000 K, as above this temperature thermionic and ion emission from dust grains lead to a sharp increase in the ionization fraction and the onset of the MRI at the disc midplane. However, the exact value varies as a function of dust grain properties, gas density at the pressure maximum, as well as the highly non-linear MRI criteria. The dependence of $\mathrm{St}_{\mathrm{frag}}/\bar{\alpha}$ on $\dot{M}$ comes mostly from the variations in this critical temperature.

Overall, whether the dust grains become trapped in the pressure maximum depends on an assumed value of α_{DZ}. We may consider how the disc might evolve forward, depending on this value. First, if the dead-zone viscosity parameter is low, $\alpha_{\mathrm{DZ}} = 10^{-5}$, $\mathrm{St}_{\mathrm{frag}}/\bar{\alpha} \gg 1$. In this case, dust grains could readily accumulate at the pressure maximum. This accumulation might lead to an unstable configuration, as an increase in the dust-to-gas ratio leads to an increase in $\bar{\alpha}$ at a given radius (see Fig. 6.4). As the dust-to-gas ratio would increase at the pressure maximum, and decrease both inwards and outwards, this might lead to an emergence of an additional minimum in $\bar{\alpha}$, and thus to a formation of an additional pressure trap. Time-dependent simulations are needed to examine further evolution of the disc, and the possibility of planetesimal formation in this case.

Second, consider a case in which the dead-zone viscosity parameter is closer to the middle of the plausible range, $\alpha_{\mathrm{DZ}} = 10^{-4}$. For all considered gas accretion rates $\mathrm{St}_{\mathrm{frag}}/\bar{\alpha} \sim 1$. This is the case considered in Chapter 3, where it was shown that the gas pressure maximum does not trap large amounts of dust. This is because $\bar{\alpha}$ increases inwards of the pressure maximum, and $\mathrm{St}_{\mathrm{frag}}$ decreases. This limits the radial width of the pressure trap, as defined above. Dust advection with the accreting gas is not sufficient to remove the dust grains from the trap; however, the grains are also mixed radially by the turbulence. The dust-to-gas ratio at the pressure maximum is then limited by the corresponding radial diffusion term. Nevertheless, dust would still accumulate in the entire region interior to the pressure maximum, as in the highly-turbulent innermost region dust grains become small enough to couple to the gas, reducing the radial drift relative to the outer disc.

In this chapter, it is also found that a higher dust-to-gas ratio yields a larger extent of the high-viscosity inner region (see Section 6.3.1.2). This is highly beneficial for planet formation in the inner disc, as it implies that the accumulation of dust is not only sustainable, but also leads to a radial expansion of the high-viscosity, high-turbulence region inside of which the dust accumulates. In particular, this expansion is beneficial for the growth of the small fragmentation-limited dust grains into larger, more rigid solid bodies, i.e., into planetesimals.

Specifically, planetesimals may form out of small grains through a combination of the streaming instabilities (SI) and the gravitational instability (Youdin and Goodman 2005; Johansen and Youdin 2007; Bai and Stone 2010; Johansen et al. 2012; Simon et al. 2016; 2017; Schäfer et al. 2017). Under certain conditions, the SI leads to localized concentrations of dust grains. If these concentrations are susceptible to gravitational instability, this may lead to the formation of planetesimals. It was pointed out in Chapter 3 that this process is unlikely in the inner disc if the pressure (and the density) maximum is located at very short orbital distances, as too close to the star the tidal effect of the star prevents the gravitational collapse. Therefore, the shift of the pressure maximum to larger orbital distances due to the accumulation of dust could potentially help to overcome this barrier and form planetesimals. Note that, for the same reasons, planetesimal formation near the pressure maximum in the inner disc favours larger gas accretion rates, and therefore formation at earlier times in the disc lifetime.

Finally, if the dead-zone viscosity parameter is high, $\alpha_{\rm DZ} = 10^{-3}$, $\mathrm{St}_{\rm frag}/\bar{\alpha} \ll 1$. In this case, dust grains are so small and well-coupled to the gas that they are advected through the pressure maximum inwards. Dust may still accumulate interior to the pressure maximum, as a consequence of fragmentation in the innermost regions, as noted above. However, it is unlikely that this could lead to the formation of larger solid bodies. While the exact value of the dead-zone viscosity is unimportant for the location of the pressure maximum (see Fig. 6.8 and Section 6.5.1), in this case the grains would be too small, too well-coupled to the gas to start the streaming instabilities (Carrera et al. 2015; Yang et al. 2017).

6.6. Conclusions

I have explored how the structure of the MRI-accreting inner regions of protoplanetary discs changes as a function of the dust-to-gas ratio, dust grain size, and other disc and stellar parameters. I have especially focused on the location of the gas pressure maximum arising at the boundary between the highly-viscous innermost region and the low-viscosity outer region. The existence and the location of the pressure maximum, and the disc structure in its vicinity, are key to the formation of the super-Earths inside the water ice line.

At fixed dust parameters, the radius of the pressure maximum is directly related to the stellar mass M_* and the gas accretion rate $\dot{M}$. This is because the stellar mass and the accretion rate determine the total viscous dissipation at a given radius, and thus the temperature and the ionization fraction at disc midplane. The radius of the pressure maximum is inversely related to the assumed viscosity parameter in the MRI-dead zone $\alpha_{\rm DZ}$. The location of the pressure maximum corresponds to a minimum in the viscosity parameter. Even though in the model discussed here there is an MRI-active layer at all radii, in the outer regions this is a thin (X-ray ionized) layer high above the disc midplane, and the disc primarily accretes through the dense MRI-dead regions around the midplane. Therefore, the minimum viscosity parameter is close in value to the dead-zone α_{DZ}.

At fixed stellar and disc parameters, the location of the pressure maximum moves radially outwards as the dust grains grow to few tens of microns. Grain growth to still larger sizes results in the pressure maximum moving inwards, towards the star. This behaviour is primarily driven by the effects of dust opacities on the disc thermal structure.

I calculate the location of the pressure maximum for the case of dust growth being limited by turbulent fragmentation. For a solar-mass star and gas accretion rates in the range $10^{-9} - 10^{-7}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$, this always places the pressure maximum outwards of 0.1 AU. In this fragmentation-limited regime, the radius of the pressure maximum depends very weakly on the dead-zone viscosity parameter, and it is most sensitive to the gas accretion rate. For a stellar mass of $0.1\,\mathrm{M}_{\odot}$, no pressure maximum exists at the lower end of the observed gas accretion rates ($10^{-11}\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$). This suggests that planet formation in the inner disc is more likely early in the disc lifetime.

The fragmentation-limited dust grain size and its Stokes number are most sensitive to the value of the viscosity (and turbulence) parameter at the pressure maximum. As noted above, this roughly equals the assumed value of the viscosity parameter in the MRI-dead zone (α_{DZ}). Therefore, whether the dust grains can become trapped in the pressure maximum is determined by this uncertain parameter. Dust trapping is likely for the lower end of plausible values ($\alpha_{\mathrm{DZ}} = 10^{-5}$) and will not happen for the higher end ($\alpha_{\mathrm{DZ}} = 10^{-3}$).

Importantly, the pressure maximum moves outwards for higher dust-to-gas ratios. This suggests that accumulation of dust near the pressure maximum and/or inwards of it results in an expansion of the dust-enriched region. However, time-dependent simulations are needed to further study the potential outcomes, and the viability of planetesimal formation in the inner disc.

7 SUMMARY AND OUTLOOK

In this thesis I have studied the structure of the inner regions of protoplanetary discs, and whether these regions provide favourable conditions for the formation of the abundant close-in super-Earths. The key question that I have explored is whether dust can accumulate in the inner disc, possibly at a local gas pressure maximum that is expected to form if the innermost regions of the disc are accreting due to the MRI.

7.1. Summary

In Chapter 2 I have discussed a steady-state model of the inner disc in which the gas structure is governed by viscous accretion, primarily due to the MRI. Here, a simple model of the gas disc structure is self-consistently coupled to a detailed parametrization of the MRI-driven viscosity. This parametrization accounts for the suppression of the MRI by Ohmic and ambipolar diffusion. In line with theoretical expectations, the MRI is active in the innermost region due to thermal ionization of potassium. In the absence of non-thermal sources of ionization, the outer, colder region is MRI-dead. A local gas pressure maximum forms, as expected, at the transition between the two regions. It is shown that, for a wide range of disc and stellar parameters, the pressure maximum indeed occurs at the orbital distances at which the close-in super-Earths are observed.

Another important result of this study is that the steady-state structure of the inner disc is viscously unstable. At a given orbital distance the vertically-averaged viscosity increases with the gas accretion rate so much that, throughout most of the inner disc, the gas surface density decreases with increasing gas accretion rate. This implies that a small accumulation of gas will lead to a lower accretion rate, which will lead to a further accumulation of gas. Therefore, this instability might be important for the evolution of both gas and dust in the inner disc and deserves further study, as discussed further below.

Next, in Chapter 3, I have investigated the evolution of dust in this steady-state gas disc model, including dust grain growth, radial drift and fragmentation. In the inner disc the dust grain size is limited by fragmentation due to relative turbulent velocities (in this case, induced by the MRI). Grains that radially drift from the outer to the inner disc become smaller due to fragmentation, and the effects of gas drag weaken. As a result, the pressure maximum is not an efficient trap for the grains. However, in the innermost, highly-turbulent region, the grains fragment to such small sizes that they become almost completely coupled to the gas. As this effectively halts the radial drift, dust accumulates throughout the inner disc, interior to the pressure maximum.

Using the obtained structure of the dusty disc, I explore if the dust grains can overcome the fragmentation limit through a combination of streaming instabilities and gravitational collapse. Previous studies have shown that the combination of these processes is a promising mechanism to form planetesimals out of dust grains, if certain conditions are met. However, in the inner disc, these processes appear to be viable in a restricted region of parameter space. There are two main limitations. First, the streaming instabilities require the dust grains to settle vertically towards the disc midplane. However, due to their small sizes, grains are efficiently stirred by the turbulence. Second, even if the criteria to start the streaming instabilities are fulfilled, and if it is assumed that this leads to localized concentrations of dust, gravitational collapse into planetesimals is difficult due to the tidal effect of the star. The required critical density of dust increases steeply radially inwards and presents a major obstacle for planetesimal formation in the inner disc. Note that, due to the preference for higher densities, the pressure (and the density) maximum is still the most favourable location, despite the dust being enhanced everywhere interior to the pressure maximum.

However, even without knowing how exactly solid planet cores form in the inner disc, it is possible (and fruitful) to study later stages of planet formation. In particular, for the super-Earths that have an atmosphere, it is known that their masses are dominated their solid cores, while their radii are largely determined by their atmospheres. This means that, for a given core mass, one can study formation of the gaseous envelope, and compare the results to the observations of planet radii. The formation of the envelope depends on the conditions in the gaseous disc in which the planet is embedded, creating a connection between theoretical disc models and observations of planet properties. Thus, in Chapter 4, I have studied the size of the atmospheres that super-Earth-mass cores can accrete, given the densities and temperatures present in the inner disc, as obtained from the model presented in Chapter 2. Furthermore, I considered how these atmospheres evolve to their present ages, including the atmospheric mass-loss due to high-energy stellar flux. I find that the MRI-accreting inner disc structure is favourable in avoiding runaway gas accretion onto planet cores (for typical super-Earth masses), while the predicted planetary atmospheres are still at least as large as observed. On the other hand, the predicted atmospheres are still often larger than the observed ones, suggesting that additional mass-loss mechanisms may be at work.

The inner disc model discussed in Chapter 2 proved extremely useful in identifying various benefits and barriers for planet formation at short orbital distances, as discussed in Chapters 3 and 4. However, there are a number of limitations of this model. Most importantly, the effects of dust grains on the disc ionization state are not accounted for, and the effect of dust on the disc thermal structure is only considered via a single constant value of the opacity. It is already pointed out in Chapter 2 that omitting these effects leads to a violation of some of the underlying assumptions about the criteria for the MRI, and that realistic dust opacities may deviate from the assumed constant value. Findings presented in Chapter 3 further stress the importance of considering the effects of dust on the gas structure and on the levels of turbulence. The model also features other simplifications, such as assuming that the disc is vertically isothermal, and neglecting heating and ionization of the disc by the stellar flux.

Therefore, in Chapter 5, I build an improved model of the inner disc structure, still coupled to the same detailed model of MRI-driven accretion. In this improved model, disc thermal structure is set by viscous heating, heating due to stellar irradiation, radiative and convective energy transport, and realistic dust opacities. I show that, for the fiducial disc and stellar parameters (including micron-size grains), stellar irradiation is largely unimportant, due to the inner disc being optically thick. The vertical temperature profile deviates from vertically isothermal, even causing the disc to become convectively unstable. Nevertheless, these changes only weakly affect the radial location of the gas pressure maximum.

Including the effects of dust on the disc ionization state also only weakly affects the location of the gas pressure maximum. This, on the other hand, is not because the effects of dust are negligible. On the contrary, thermionic and ion emission from dust grains are the primary sources of ionization at high temperatures when grains are included. However, similarly to thermal ionization, there is a threshold temperature at which these processes become efficient, and this threshold temperature is roughly the same as for thermal ionization of potassium. Therefore, the MRI-driven viscosity rises at roughly the same radial location in the disc. Additionally, in this improved model, ionization of molecular hydrogen by stellar X-rays activates the MRI at larger radii (where temperatures are too low for the thermal effects). However, in these outer regions (but still in the vicinity of the pressure maximum), the MRI is active only in a thin low-density layer high above the disc midplane. Gas accretion is still dominated by the dense, MRI-dead zone around the disc midplane.

Finally, in Chapter 6, I have used this improved model to investigate how the structure of the inner disc depends on various disc and stellar parameters. I have found that the dependence of the radial location of the pressure maximum on stellar mass, gas accretion rate and the dead-zone viscosity parameter is similar to the scalings identified in the simple model discussed in Chapter 2. This finding can be attributed to the fact that the basic mechanisms setting the disc structure remain the same, including there being a critical temperature above which disc becomes ionized enough to couple to the magnetic field.

Importantly, for fixed values of the above parameters, the inner disc structure changes as a function of the dust grain size and the dust-to-gas ratio. For maximum dust grain sizes larger than ∼100 microns, an increase in grain size results in a decrease in the extent of the high-viscosity inner region and thus in the radius of the local gas pressure maximum. Still, if grain growth is limited by turbulent fragmentation, in a disc surrounding a solar-mass star the gas pressure maximum should exist at few tenths of an AU. For a stellar mass of $0.1\,\mathrm{M}_{\odot}$, the allowed region of the parameter space for the existence of the pressure maximum is more constricted, as no pressure maximum is found for the low end of the observed gas accretion rates. Overall, these results confirm that the inner disc will feature a high-viscosity region and a pressure maximum that may accumulate the solids necessary to build super-Earths, except, perhaps, at later stages of disc evolution around small stars.

However, whether the formation of solid planet cores is likely in the inner disc depends on the uncertain dead-zone viscosity parameter. Grains are likely to accumulate, either at the pressure maximum (for lower values of the dead-zone viscosity) or inwards of it (as discussed in Chapter 3). However, for high values of the dead-zone viscosity parameter, turbulent fragmentation and stirring would inhibit processes necessary to form larger solid bodies. Here, an important result presented in this thesis is that dust accumulation (i.e., an increase in the dust-to-gas ratio) acts to increase the MRI-driven viscosity at a given radius (and thus the radial extent of the high-viscosity inner region). If grains accumulate at the pressure maximum, this behaviour might lead to non-trivial outcomes, whose further study requires time-dependent simulations, as discussed below.

7.2. Outlook

Throughout this work it is assumed that the MRI-accreting inner disc is in an equilibrium, steady state. Concurrently, the obtained results suggest that the disc is likely to evolve out of this state.

First, in Chapter 2 it was shown that the steady-state MRI-accreting inner disc is unstable to surface density perturbations. At a fixed radius, the MRI-driven accretion rate decreases with an increasing surface density (as calculated using the steady-state models). Therefore, a perturbation in the disc surface density might lead out of the steady-state by creating an increasing pile-up of mass at certain locations. This so-called viscous instability (Lightman and Eardley 1974; Pringle 1981) could have important consequences for the inner disc structure and planet formation. It is likely to produce rings and gaps in the gas structure on the viscous timescale. Gas rings may further concentrate dust, while also allowing dust growth due to decreased turbulence levels, creating favourable conditions for planetesimal formation.

Second, joint evolution of gas and dust might not produce a steady state. In particular, it is unclear how the disc would evolve if dust grains accumulate in the pressure maximum. Higher dust-to-gas ratio would lead to a local increase of the viscosity. It is unclear whether further evolution of gas and dust might lead to a modified equilibrium state, or creation of multiple viscosity minima (and pressure maxima).

Evidently, the inner disc will evolve in time, and the possible outcomes are crucial to determine the degree to which dust grains accumulate in the inner disc, and whether they can overcome the fragmentation barrier and build solid planet cores. Therefore, further study is required to understand the time-dependence of the inner disc structure, using simulations that self-consistently couple the gas and the dust. Such simulations would rely directly on the model presented in this thesis, which would be used to tabulate the value of the viscosity parameter as a function of gas surface density, dust grain size, dust-to-gas ratio etc., as an input for the time-dependent evolution.

This proposed future study should also account for the dynamical effects of dust on the gas, which have been neglected in this work. If dust grains drift radially through the disc due to gas drag, there is also a back-reaction on the gas (Nakagawa et al. 1986). This becomes increasingly important at high dust-to-gas ratios. The drag back-reaction acts to flatten the radial gas pressure profile, and so it would affect if and how much dust may accumulate in the pressure maximum (Taki et al. 2016).

These time-dependent semi-analytic models will undoubtedly be useful for future studies of the early stages of planet formation in the inner disc. However, the semi-analytic parametrization of gas accretion presented in this thesis should also be examined with magnetohydrodynamic simulations suited for the conditions in the inner disc, and accounting for both Ohmic and ambipolar diffusion. For example, the fact that the semi-analytic steady-state solution is viscously unstable, even in the absence of any dust growth or spatial evolution, deserves more detailed study. More generally, magnetohydrodynamic simulations are important in order to check the unverified assumptions of the semi-analytic model, such as the assumption that the magnetic field always evolves as to maximize the accretion stress in the inner disc. Another application would be to investigate the effect of convective motions on the evolution of the MRI, which has been shown to be important in the non-ideal limit (Bodo et al. 2013; Hirose et al. 2014).

Finally, it must be stressed that the presented picture of the MRI-driven accretion in the inner disc is incomplete. Out of the three non-ideal MHD effects relevant in protoplanetary discs, only two have been accounted for. In addition to Ohmic and ambipolar diffusion, Hall effect may also be important. The Hall effect can drive angular momentum transport in the dead zone (Lesur et al. 2014), although this is sensitive to the alignment between the vertical magnetic field and the disc (Simon et al. 2015; Bai 2014). The simple calculation of the Hall Elsasser number in Chapter 2 shows that a “Hall zone” might also overlap with an MRI-active region. Although the innermost MRI-active region should still persist when the Hall effect is included, the Hall effect can significantly affect planet formation in this region simply by pushing the pressure maximum inwards, and thus increasing the critical density for the gravitational collapse of grains into planetesimals.

BIBLIOGRAPHY

Adams, F. C., Lada, C. J., and Shu, F. H. (1987). Spectral Evolution of Young Stellar Objects. ApJ, 312:788.

Agol, E., Steffen, J., Sari, R., and Clarkson, W. (2005). On detecting terrestrial planets with timing of giant planet transits. MNRAS, 359(2):567–579.

Alcalá, J. M., Manara, C. F., Natta, A., Frasca, A., Testi, L., Nisini, B., Stelzer, B., Williams, J. P., Antoniucci, S., Biazzo, K., Covino, E., Esposito, M., Getman, F., and Rigliaco, E. (2017). X-shooter spectroscopy of young stellar objects in Lupus. Accretion properties of class II and transitional objects. A&A, 600:A20.

Alcalá, J. M., Natta, A., Manara, C. F., Spezzi, L., Stelzer, B., Frasca, A., Biazzo, K., Covino, E., Randich, S., Rigliaco, E., Testi, L., Comerón, F., Cupani, G., and D'Elia, V. (2014). X-shooter spectroscopy of young stellar objects. IV. Accretion in low-mass stars and substellar objects in Lupus. A&A, 561:A2.

ALMA Partnership, Brogan, C. L., Pérez, L. M., Hunter, T. R., Dent, W. R. F., Hales, A. S., Hills, R. E., Corder, S., Fomalont, E. B., Vlahakis, C., Asaki, Y., Barkats, D., Hirota, A., Hodge, J. A., Impellizzeri, C. M. V., Kneissl, R., Liuzzo, E., Lucas, R., Marcelino, N., Matsushita, S., Nakanishi, K., Phillips, N., Richards, A. M. S., Toledo, I., Aladro, R., Broguiere, D., Cortes, J. R., Cortes, P. C., Espada, D., Galarza, F., Garcia-Appadoo, D., Guzman-Ramirez, L., Humphreys, E. M., Jung, T., Kameno, S., Laing, R. A., Leon, S., Marconi, G., Mignano, A., Nikolic, B., Nyman, L. A., Radiszcz, M., Remijan, A., Rodón, J. A., Sawada, T., Takahashi, S., Tilanus, R. P. J., Vila Vilaro, B., Watson, L. C., Wiklind, T., Akiyama, E., Chapillon, E., de Gregorio-Monsalvo, I., Di Francesco, J., Gueth, F., Kawamura, A., Lee, C. F., Nguyen Luong, Q., Mangum, J., Pietu, V., Sanhueza, P., Saigo, K., Takakuwa, S., Ubach, C., van Kempen, T., Wootten, A., Castro-Carrizo, A., Francke, H., Gallardo, J., Garcia, J., Gonzalez, S., Hill, T., Kaminski, T., Kurono, Y., Liu, H. Y., Lopez, C., Morales, F., Plarre, K., Schieven, G., Testi, L., Videla, L., Villard, E., Andreani, P., Hibbard, J. E., and Tatematsu, K. (2015). The 2014 ALMA Long Baseline Campaign: First Results from High Angular Resolution Observations toward the HL Tau Region. ApJ, 808(1):L3.

Andrews, S. M., Huang, J., Pérez, L. M., Isella, A., Dullemond, C. P., Kurtovic, N. T., Guzmán, V. V., Carpenter, J. M., Wilner, D. J., Zhang, S., Zhu, Z., Birnstiel, T., Bai, X.-N., Benisty, M., Hughes, A. M., Öberg, K. I., and Ricci, L. (2018). The Disk Substructures at High Angular Resolution Project (DSHARP). I. Motivation, Sample, Calibration, and Overview. ApJ, 869(2):L41.

Andrews, S. M. and Williams, J. P. (2005). Circumstellar Dust Disks in Taurus-Auriga: The Submillimeter Perspective. ApJ, 631(2):1134–1160.

Andrews, S. M. and Williams, J. P. (2007). A Submillimeter View of Circumstellar Dust Disks in ρ Ophiuchi. ApJ, 671(2):1800–1812.

Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., Rosenfeld, K. A., Öberg, K. I., Birnstiel, T., Espaillat, C., Cieza, L. A., Williams, J. P., Lin, S.-Y., and Ho, P. T. P. (2012). The TW Hya Disk at 870 μm: Comparison of CO and Dust Radial Structures. ApJ, 744:162.

Bai, X.-N. (2011a). Magnetorotational-instability-driven Accretion in Protoplanetary Disks. ApJ, 739(1):50.

Bai, X.-N. (2011b). The Role of Tiny Grains on the Accretion Process in Protoplanetary Disks. ApJ, 739(1):51.

Bai, X.-N. (2014). Hall-effect-Controlled Gas Dynamics in Protoplanetary Disks. I. Wind Solutions at the Inner Disk. ApJ, 791(2):137.

Bai, X.-N. (2017). Global Simulations of the Inner Regions of Protoplanetary Disks with Comprehensive Disk Microphysics. ApJ, 845(1):75.

Bai, X.-N. and Goodman, J. (2009). Heat and Dust in Active Layers of Protostellar Disks. ApJ, 701(1):737–755.

Bai, X.-N. and Stone, J. M. (2010). Dynamics of Solids in the Midplane of Protoplanetary Disks: Implications for Planetesimal Formation. ApJ, 722(2):1437–1459.

Bai, X.-N. and Stone, J. M. (2011). Effect of Ambipolar Diffusion on the Nonlinear Evolution of Magnetorotational Instability in Weakly Ionized Disks. ApJ, 736(2):144.

Bai, X.-N. and Stone, J. M. (2013). Local Study of Accretion Disks with a Strong Vertical Magnetic Field: Magnetorotational Instability and Disk Outflow. ApJ, 767(1):30.

Balbus, S. A. (2000). Stability, Instability, and "Backward" Transport in Stratified Fluids. ApJ, 534(1):420–427.

Balbus, S. A. (2011). Magnetohydrodynamics of Protostellar Disks. In Garcia, P. J. V., editor, *Physical Processes in Circumstellar Disks around Young Stars*, pages 237–282. University of Chicago Press, Chicago, IL.

Balbus, S. A. and Hawley, J. F. (1991). A powerful local shear instability in weakly magnetized disks. I - Linear analysis. II - Nonlinear evolution. ApJ, 376:214–233.

Balbus, S. A. and Hawley, J. F. (1998). Instability, turbulence, and enhanced transport in accretion disks. *Reviews of Modern Physics*, 70(1):1–53.

Baraffe, I., Chabrier, G., Allard, F., and Hauschildt, P. H. (1998). Evolutionary models for solar metallicity low-mass stars: mass-magnitude relationships and color-magnitude diagrams. A&A, 337:403–412.

Baraffe, I., Homeier, D., Allard, F., and Chabrier, G. (2015). New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit. A&A, 577:A42.

Beckwith, S. V. W. and Sargent, A. I. (1991). Particle Emissivity in Circumstellar Disks. ApJ, 381:250.

Beckwith, S. V. W., Sargent, A. I., Chini, R. S., and Guesten, R. (1990). A Survey for Circumstellar Disks around Young Stellar Objects. AJ, 99:924.

Beitz, E., Güttler, C., Blum, J., Meisner, T., Teiser, J., and Wurm, G. (2011). Low-velocity Collisions of Centimeter-sized Dust Aggregates. ApJ, 736:34.

Birnstiel, T., Dullemond, C. P., and Brauer, F. (2009). Dust retention in protoplanetary disks. A&A, 503(1):L5–L8.

Birnstiel, T., Dullemond, C. P., and Brauer, F. (2010). Gas- and dust evolution in protoplanetary disks. A&A, 513:A79.

Birnstiel, T., Klahr, H., and Ercolano, B. (2012). A simple model for the evolution of the dust population in protoplanetary disks. A&A, 539:A148.

Birnstiel, T., Ormel, C. W., and Dullemond, C. P. (2011). Dust size distributions in coagulation/fragmentation equilibrium: numerical solutions and analytical fits. A&A, 525:A11.

Blum, J. and Münch, M. (1993). Experimental Investigations on Aggregate-Aggregate Collisions in the Early Solar Nebula. *Icarus*, 106(1):151–167.

Bodo, G., Cattaneo, F., Mignone, A., and Rossi, P. (2013). Fully Convective Magnetorotational Turbulence in Stratified Shearing Boxes. ApJ, 771(2):L23.

Boley, A. C. and Ford, E. B. (2013). The Formation of Systems with Tightly-packed Inner Planets (STIPs) via Aerodynamic Drift. *arXiv e-prints*, page arXiv:1306.0566.

Borucki, W. J., Koch, D., and Kepler Science Team (2010). Kepler Planet Detection Mission: Highlights of the First Results. In *AAS/Division for Planetary Sciences Meeting Abstracts #42*, AAS/Division for Planetary Sciences Meeting Abstracts, page 47.03.

Borucki, W. J., Koch, D. G., Basri, G., Batalha, N., Brown, T. M., Bryson, S. T., Caldwell, D., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Gautier, Thomas N., I., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Latham, D. W., Lissauer, J. J., Marcy, G. W., Rowe, J., Sasselov, D., Boss, A., Charbonneau, D., Ciardi, D., Doyle, L., Dupree, A. K., Ford, E. B., Fortney, J., Holman, M. J., Seager, S., Steffen, J. H., Tarter, J., Welsh, W. F., Allen, C., Buchhave, L. A., Christiansen, J. L., Clarke, B. D., Das, S., Désert, J.-M., Endl, M., Fabrycky, D., Fressin, F., Haas, M., Horch, E., Howard, A., Isaacson, H., Kjeldsen, H., Kolodziejczak, J., Kulesa, C., Li, J., Lucas, P. W., Machalek, P., McCarthy, D., MacQueen, P., Meibom, S., Miquel, T., Prsa, A., Quinn, S. N., Quintana, E. V., Ragozzine, D., Sherry, W., Shporer, A., Tenenbaum, P., Torres, G., Twicken, J. D., Van Cleve, J., Walkowicz, L., Witteborn, F. C., and Still, M. (2011). Characteristics of Planetary Candidates Observed by Kepler. II. Analysis of the First Four Months of Data. ApJ, 736(1):19.

Bouvier, J., Alencar, S. H. P., Harries, T. J., Johns-Krull, C. M., and Romanova, M. M. (2007). Magnetospheric Accretion in Classical T Tauri Stars. In Reipurth, B., Jewitt, D., and Keil, K., editors, *Protostars and Planets V*, page 479.

Brandenburg, A., Nordlund, A., Stein, R. F., and Torkelsson, U. (1995). Dynamo-generated Turbulence and Large-Scale Magnetic Fields in a Keplerian Shear Flow. ApJ, 446:741.

Brauer, F., Dullemond, C. P., and Henning, T. (2008). Coagulation, fragmentation and radial motion of solid particles in protoplanetary disks. A&A, 480:859–877.

Bukhari Syed, M., Blum, J., Wahlberg Jansson, K., and Johansen, A. (2017). The Role of Pebble Fragmentation in Planetesimal Formation. I. Experimental Study. ApJ, 834:145.

Burke, C. J., Bryson, S. T., Mullally, F., Rowe, J. F., Christiansen, J. L., Thompson, S. E., Coughlin, J. L., Haas, M. R., Batalha, N. M., Caldwell, D. A., Jenkins, J. M., Still, M., Barclay, T., Borucki, W. J., Chaplin, W. J., Ciardi, D. R., Clarke, B. D., Cochran, W. D., Demory, B.-O., Esquerdo, G. A., Gautier, Thomas N., I., Gilliland , R. L., Girouard, F. R., Havel, M., Henze, C. E., Howell, S. B., Huber, D., Latham, D. W., Li, J., Morehead, R. C., Morton, T. D., Pepper, J., Quintana, E., Ragozzine, D., Seader, S. E., Shah, Y., Shporer, A., Tenenbaum, P., Twicken, J. D., and Wolfgang, A. (2014). Planetary Candidates Observed by Kepler IV: Planet Sample from Q1-Q8 (22 Months). ApJS, 210(2):19.

Calvet, N., Magris, G. C., Patino, A., and D'Alessio, P. (1992). Irradiation of Accretion Disks around Young Objects. II. Continuum Energy Distribution. Rev. Mexicana Astron. Astrofis., 24:27.

Carrera, D., Gorti, U., Johansen, A., and Davies, M. B. (2017). Planetesimal Formation by the Streaming Instability in a Photoevaporating Disk. ApJ, 839(1):16.

Carrera, D., Johansen, A., and Davies, M. B. (2015). How to form planetesimals from mm-sized chondrules and chondrule aggregates. A&A, 579:A43.

Cassen, P. (1993). Why Convective Heat Transport in the Solar Nebula was Inefficient. In *Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 261.

Chatterjee, S. and Tan, J. C. (2014). Inside-out Planet Formation. ApJ, 780(1):53.

Chatterjee, S. and Tan, J. C. (2015). Vulcan Planets: Inside-out Formation of the Innermost Super-Earths. ApJ, 798(2):L32.

Chiang, E. and Laughlin, G. (2013). The minimum-mass extrasolar nebula: in situ formation of close-in super-Earths. MNRAS, 431(4):3444–3455.

Chiang, E. I. and Goldreich, P. (1997). Spectral Energy Distributions of T Tauri Stars with Passive Circumstellar Disks. ApJ, 490(1):368–376.

Chiang, E. I., Joung, M. K., Creech-Eakman, M. J., Qi, C., Kessler, J. E., Blake, G. A., and van Dishoeck, E. F. (2001). Spectral Energy Distributions of Passive T Tauri and Herbig Ae Disks: Grain Mineralogy, Parameter Dependences, and Comparison with Infrared Space Observatory LWS Observations. ApJ, 547(2):1077–1089.

Cimerman, N. P., Kuiper, R., and Ormel, C. W. (2017). Hydrodynamics of embedded planets' first atmospheres - III. The role of radiation transport for super-Earth planets. MNRAS, 471:4662–4676.

Coleman, G. A. L. and Nelson, R. P. (2014). On the formation of planetary systems via oligarchic growth in thermally evolving viscous discs. MNRAS, 445(1):479–499.

Coleman, G. A. L. and Nelson, R. P. (2016). On the formation of compact planetary systems via concurrent core accretion and migration. MNRAS, 457(3):2480–2500.

Cossou, C., Raymond, S. N., Hersant, F., and Pierens, A. (2014). Hot super-Earths and giant planet cores from different migration histories. A&A, 569:A56.

Cresswell, P. and Nelson, R. P. (2006). On the evolution of multiple protoplanets embedded in a protostellar disc. A&A, 450(2):833–853.

Cuzzi, J. N. and Hogan, R. C. (2003). Blowing in the wind. I. Velocities of chondrule-sized particles in a turbulent protoplanetary nebula. *Icarus*, 164(1):127–138.

D'Alessio, P., Calvet, N., and Hartmann, L. (2001). Accretion Disks around Young Objects. III. Grain Growth. ApJ, 553(1):321–334.

D'Alessio, P., Calvet, N., Hartmann, L., Lizano, S., and Cantó, J. (1999). Accretion Disks around Young Objects. II. Tests of Well-mixed Models with ISM Dust. ApJ, 527(2):893–909.

D'Alessio, P., Cantö, J., Calvet, N., and Lizano, S. (1998). Accretion Disks around Young Objects. I. The Detailed Vertical Structure. ApJ, 500(1):411–427.

Davis, S. W., Stone, J. M., and Pessah, M. E. (2010). Sustained Magnetorotational Turbulence in Local Simulations of Stratified Disks with Zero Net Magnetic Flux. ApJ, 713(1):52–65.

Demirci, T., Krause, C., Teiser, J., and Wurm, G. (2019). Onset of planet formation in the warm inner disk. Colliding dust aggregates at high temperatures. A&A, 629:A66.

Desch, S. J. and Turner, N. J. (2015). High-temperature Ionization in Protoplanetary Disks. ApJ, 811(2):156.

Draine, B. T. (2003). Scattering by Interstellar Dust Grains. I. Optical and Ultraviolet. ApJ, 598(2):1017–1025.

Draine, B. T. and Sutin, B. (1987). Collisional charging of interstellar grains. ApJ, 320:803–817.

Drazkowska, J. and Alibert, Y. (2017). Planetesimal formation starts at the snow line. A&A, 608:A92.

Drazkowska, J., Alibert, Y., and Moore, B. (2016). Close-in planetesimal formation by pile-up of drifting pebbles. A&A, 594:A105.

Drazkowska, J., Windmark, F., and Dullemond, C. P. (2013). Planetesimal formation via sweep-up growth at the inner edge of dead zones. A&A, 556:A37.

Dressing, C. D. and Charbonneau, D. (2013). The Occurrence Rate of Small Planets around Small Stars. ApJ, 767(1):95.

Dressing, C. D. and Charbonneau, D. (2015). The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the Full Kepler Dataset and an Empirical Measurement of the Detection Sensitivity. ApJ, 807(1):45.

Dressing, C. D., Charbonneau, D., Dumusque, X., Gettel, S., Pepe, F., Collier Cameron, A., Latham, D. W., Molinari, E., Udry, S., Affer, L., Bonomo, A. S., Buchhave, L. A., Cosentino, R., Figueira, P., Fiorenzano, A. F. M., Harutyunyan, A., Haywood, R. D., Johnson, J. A., Lopez-Morales, M., Lovis, C., Malavolta, L., Mayor, M., Micela, G., Motalebi, F., Nascimbeni, V., Phillips, D. F., Piotto, G., Pollacco, D., Queloz, D., Rice, K., Sasselov, D., Ségransan, D., Sozzetti, A., Szentgyorgyi, A., and Watson, C. (2015). The Mass of Kepler-93b and The Composition of Terrestrial Planets. ApJ, 800(2):135.

Dubrulle, B., Morfill, G., and Sterzik, M. (1995). The dust subdisk in the protoplanetary nebula. *Icarus*, 114(2):237–246.

Dullemond, C. P. (2002). The 2-D structure of dusty disks around Herbig Ae/Be stars. I. Models with grey opacities. A&A, 395:853–862.

Dullemond, C. P., Juhasz, A., Pohl, A., Sereshti, F., Shetty, R., Peters, T., Commercon, B., and Flock, M. (2012). RADMC-3D: A multi-purpose radiative transfer tool. *Astrophysics Source Code Library*, page ascl:1202.015.

Dutrey, A., Guilloteau, S., Duvert, G., Prato, L., Simon, M., Schuster, K., and Menard, F. (1996). Dust and gas distribution around T Tauri stars in Taurus-Auriga. I. Interferometric 2.7mm continuum and $\hat{1}3\hat{C}$O J=1-0 observations. A&A, 309:493–504.

Dzyurkevich, N., Flock, M., Turner, N. J., Klahr, H., and Henning, T. (2010). Trapping solids at the inner edge of the dead zone: 3-D global MHD simulations. A&A, 515:A70.

Ercolano, B. and Glassgold, A. E. (2013). X-ray ionization rates in protoplanetary discs. MNRAS, 436(4):3446–3450.

Fedele, D., van den Ancker, M. E., Henning, T., Jayawardhana, R., and Oliveira, J. M. (2010). Timescale of mass accretion in pre-main-sequence stars. A&A, 510:A72.

Fleming, T. P., Stone, J. M., and Hawley, J. F. (2000). The Effect of Resistivity on the Nonlinear Stage of the Magnetorotational Instability in Accretion Disks. ApJ, 530(1):464–477.

Flock, M., Fromang, S., Turner, N. J., and Benisty, M. (2017). 3D Radiation Nonideal Magnetohydrodynamical Simulations of the Inner Rim in Protoplanetary Disks. ApJ, 835(2):230.

Flock, M., Turner, N. J., Mulders, G. D., Hasegawa, Y., Nelson, R. P., and Bitsch, B. (2019). Planet formation and migration near the silicate sublimation front in protoplanetary disks. A&A, 630:A147.

Fortney, J. J., Marley, M. S., and Barnes, J. W. (2007). Planetary Radii across Five Orders of Magnitude in Mass and Stellar Insolation: Application to Transits. ApJ, 659(2):1661–1672.

Frank, J., King, A., and Raine, D. J. (2002). *Accretion Power in Astrophysics: Third Edition*. Cambridge University Press, Cambridge, UK.

Fressin, F., Torres, G., Charbonneau, D., Bryson, S. T., Christiansen, J., Dressing, C. D., Jenkins, J. M., Walkowicz, L. M., and Batalha, N. M. (2013). The False Positive Rate of Kepler and the Occurrence of Planets. ApJ, 766(2):81.

Fromang, S., Latter, H., Lesur, G., and Ogilvie, G. I. (2013). Local outflows from turbulent accretion disks. A&A, 552:A71.

Fulton, B. J. and Petigura, E. A. (2018). The California Kepler Survey VII. Precise Planet Radii Leveraging Gaia DR2 Reveal the Stellar Mass Dependence of the Planet Radius Gap. *ArXiv e-prints*.

Fulton, B. J., Petigura, E. A., Howard, A. W., Isaacson, H., Marcy, G. W., Cargile, P. A., Hebb, L., Weiss, L. M., Johnson, J. A., Morton, T. D., Sinukoff, E., Crossfield, I. J. M., and Hirsch, L. A. (2017). The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets. AJ, 154(3):109.

Fung, J., Artymowicz, P., and Wu, Y. (2015). The 3D Flow Field Around an Embedded Planet. ApJ, 811:101.

Fung, J. and Lee, E. J. (2018). Inner Super-Earths, Outer Gas Giants: How Pebble Isolation and Migration Feedback Keep Jupiters Cold. ApJ, 859:126.

Gammie, C. F. (1996). Layered Accretion in T Tauri Disks. ApJ, 457:355.

Garaud, P. and Lin, D. N. C. (2007). The Effect of Internal Dissipation and Surface Irradiation on the Structure of Disks and the Location of the Snow Line around Sun-like Stars. ApJ, 654(1):606–624.

Garaud, P., Meru, F., Galvagni, M., and Olczak, C. (2013). From Dust to Planetesimals: An Improved Model for Collisional Growth in Protoplanetary Disks. ApJ, 764(2):146.

Ginzburg, S., Schlichting, H. E., and Sari, R. (2018). Core-powered mass-loss and the radius distribution of small exoplanets. MNRAS, 476:759–765.

Glassgold, A. E., Najita, J., and Igea, J. (1997). X-Ray Ionization of Protoplanetary Disks. ApJ, 480(1):344–350.

Goldreich, P. and Tremaine, S. (1979). The excitation of density waves at the Lindblad and corotation resonances by an external potential. ApJ, 233:857–871.

Goldreich, P. and Tremaine, S. (1980). Disk-satellite interactions. ApJ, 241:425–441.

Grasset, O., Schneider, J., and Sotin, C. (2009). A Study of the Accuracy of Mass-Radius Relationships for Silicate-Rich and Ice-Rich Planets up to 100 Earth Masses. ApJ, 693(1):722–733.

Gullbring, E., Hartmann, L., Briceño, C., and Calvet, N. (1998). Disk Accretion Rates for T Tauri Stars. ApJ, 492(1):323–341.

Haghighipour, N. and Boss, A. P. (2003). On Pressure Gradients and Rapid Migration of Solids in a Nonuniform Solar Nebula. ApJ, 583(2):996–1003.

Haisch, Karl E., J., Lada, E. A., and Lada, C. J. (2001). Disk Frequencies and Lifetimes in Young Clusters. ApJ, 553(2):L153–L156.

Hansen, B. M. S. and Murray, N. (2012). Migration Then Assembly: Formation of Neptune-mass Planets inside 1 AU. ApJ, 751(2):158.

Hansen, B. M. S. and Murray, N. (2013). Testing in Situ Assembly with the Kepler Planet Candidate Sample. ApJ, 775(1):53.

Hartmann, L., Calvet, N., Gullbring, E., and D'Alessio, P. (1998). Accretion and the Evolution of T Tauri Disks. ApJ, 495(1):385–400.

Hawley, J. F. and Balbus, S. A. (1992). A Powerful Local Shear Instability in Weakly Magnetized Disks. III. Long-Term Evolution in a Shearing Sheet. ApJ, 400:595.

Hawley, J. F., Gammie, C. F., and Balbus, S. A. (1995). Local Three-dimensional Magnetohydrodynamic Simulations of Accretion Disks. ApJ, 440:742.

Hawley, J. F., Gammie, C. F., and Balbus, S. A. (1996). Local Three-dimensional Simulations of an Accretion Disk Hydromagnetic Dynamo. ApJ, 464:690.

Hawley, J. F. and Stone, J. M. (1998). Nonlinear Evolution of the Magnetorotational Instability in Ion-Neutral Disks. ApJ, 501(2):758–771.

Hayashi, C. (1981). Structure of the Solar Nebula, Growth and Decay of Magnetic Fields and Effects of Magnetic and Turbulent Viscosities on the Nebula. *Progress of Theoretical Physics Supplement*, 70:35–53.

Held, L. E. and Latter, H. N. (2018). Hydrodynamic convection in accretion discs. MNRAS, 480(4):4797–4816.

Hillenbrand, L. A., Strom, S. E., Vrba, F. J., and Keene, J. (1992). Herbig Ae/Be Stars: Intermediate-Mass Stars Surrounded by Massive Circumstellar Accretion Disks. ApJ, 397:613.

Hirose, S. (2015). Magnetic turbulence and thermodynamics in the inner region of protoplanetary discs. MNRAS, 448(4):3105–3120.

Hirose, S., Blaes, O., Krolik, J. H., Coleman, M. S. B., and Sano, T. (2014). Convection Causes Enhanced Magnetic Turbulence in Accretion Disks in Outburst. ApJ, 787(1):1.

Holman, M. J. and Murray, N. W. (2005). The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets. *Science*, 307(5713):1288–1291.

Hsu, D. C., Ford, E. B., Ragozzine, D., and Ashby, K. (2019). Occurrence Rates of Planets Orbiting FGK Stars: Combining Kepler DR25, Gaia DR2, and Bayesian Inference. AJ, 158(3):109.

Hsu, D. C., Ford, E. B., and Terrien, R. (2020). Occurrence Rates of Planets Orbiting M Stars: Applying ABC to Kepler DR25, Gaia DR2, and 2MASS Data. *arXiv e-prints*, page arXiv:2002.02573.

Hu, X., Tan, J. C., Zhu, Z., Chatterjee, S., Birnstiel, T., Youdin, A. N., and Mohanty, S. (2018). Inside-out Planet Formation. IV. Pebble Evolution and Planet Formation Timescales. ApJ, 857(1):20.

Hu, X., Zhu, Z., Tan, J. C., and Chatterjee, S. (2016). Inside-out Planet Formation. III. Planet-Disk Interaction at the Dead Zone Inner Boundary. ApJ, 816(1):19.

Hubeny, I. (1990). Vertical Structure of Accretion Disks: A Simplified Analytical Model. ApJ, 351:632.

Ida, S. and Guillot, T. (2016). Formation of dust-rich planetesimals from sublimated pebbles inside of the snow line. A&A, 596:L3.

Ida, S. and Lin, D. N. C. (2010). Toward a Deterministic Model of Planetary Formation. VI. Dynamical Interaction and Coagulation of Multiple Rocky Embryos and Super-Earth Systems around Solar-type Stars. ApJ, 719(1):810–830.

Igea, J. and Glassgold, A. E. (1999). X-Ray Ionization of the Disks of Young Stellar Objects. ApJ, 518(2):848–858.

Ikoma, M. and Hori, Y. (2012). In Situ Accretion of Hydrogen-rich Atmospheres on Short-period Super-Earths: Implications for the Kepler-11 Planets. ApJ, 753:66.

Ilgner, M. and Nelson, R. P. (2006). On the ionisation fraction in protoplanetary disks. I. Comparing different reaction networks. A&A, 445:205–222.

Inamdar, N. K. and Schlichting, H. E. (2015). The formation of super-Earths and mini-Neptunes with giant impacts. MNRAS, 448(2):1751–1760.

Inamdar, N. K. and Schlichting, H. E. (2016). Stealing the Gas: Giant Impacts and the Large Diversity in Exoplanet Densities. ApJ, 817:L13.

Izidoro, A., Bitsch, B., Raymond, S. N., Johansen, A., Morbidelli, A., Lambrechts, M., and Jacobson, S. A. (2019). Formation of planetary systems by pebble accretion and migration: Hot super-Earth systems from breaking compact resonant chains. *arXiv e-prints*, page arXiv:1902.08772.

Izidoro, A., Ogihara, M., Raymond, S. N., Morbidelli, A., Pierens, A., Bitsch, B., Cossou, C., and Hersant, F. (2017). Breaking the chains: hot super-Earth systems from migration and disruption of compact resonant chains. MNRAS, 470(2):1750–1770.

Jackson, A. P., Davis, T. A., and Wheatley, P. J. (2012). The coronal X-ray-age relation and its implications for the evaporation of exoplanets. MNRAS, 422:2024–2043.

Jacquet, E., Gounelle, M., and Fromang, S. (2012). On the aerodynamic redistribution of chondrite components in protoplanetary disks. *Icarus*, 220(1):162–173.

Jin, L. (1996). Damping of the Shear Instability in Magnetized Disks by Ohmic Diffusion. ApJ, 457:798.

Johansen, A. and Youdin, A. (2007). Protoplanetary Disk Turbulence Driven by the Streaming Instability: Nonlinear Saturation and Particle Concentration. ApJ, 662:627–641.

Johansen, A., Youdin, A., and Mac Low, M.-M. (2009). Particle Clumping and Planetesimal Formation Depend Strongly on Metallicity. ApJ, 704:L75–L79.

Johansen, A., Youdin, A. N., and Lithwick, Y. (2012). Adding particle collisions to the formation of asteroids and Kuiper belt objects via streaming instabilities. A&A, 537:A125.

Jontof-Hutter, D. (2019). The Compositional Diversity of Low-Mass Exoplanets. *Annual Review of Earth and Planetary Sciences*, 47:141–171.

Jontof-Hutter, D., Ford, E. B., Rowe, J. F., Lissauer, J. J., Fabrycky, D. C., Van Laerhoven, C., Agol, E., Deck, K. M., Holczer, T., and Mazeh, T. (2016). Secure Mass Measurements from Transit Timing: 10 Kepler Exoplanets between 3 and 8 M with Diverse Densities and Incident Fluxes. ApJ, 820:39.

Karkoschka, E. and Tomasko, M. G. (2011). The haze and methane distributions on Neptune from HST-STIS spectroscopy. *Icarus*, 211:780–797.

Keith, S. L. and Wardle, M. (2014). Accretion in giant planet circumplanetary discs. MNRAS, 440(1):89–105.

Kley, W. and Nelson, R. P. (2012). Planet-Disk Interaction and Orbital Evolution. ARA&A, 50:211–249.

Kothe, S., Güttler, C., and Blum, J. (2010). The Physics of Protoplanetesimal Dust Agglomerates. V. Multiple Impacts of Dusty Agglomerates at Velocities Above the Fragmentation Threshold. ApJ, 725:1242–1251.

Kretke, K. A. and Lin, D. N. C. (2007). Grain Retention and Formation of Planetesimals near the Snow Line in MRI-driven Turbulent Protoplanetary Disks. ApJ, 664(1):L55–L58.

Kretke, K. A. and Lin, D. N. C. (2010). Structure of Magnetorotational Instability Active Protoplanetary Disks. ApJ, 721(2):1585–1592.

Kretke, K. A., Lin, D. N. C., Garaud, P., and Turner, N. J. (2009). Assembling the Building Blocks of Giant Planets Around Intermediate-Mass Stars. ApJ, 690(1):407–415.

Lee, E. J. and Chiang, E. (2015). To Cool is to Accrete: Analytic Scalings for Nebular Accretion of Planetary Atmospheres. ApJ, 811:41.

Lee, E. J. and Chiang, E. (2016). Breeding Super-Earths and Birthing Super-puffs in Transitional Disks. ApJ, 817:90.

Lee, E. J., Chiang, E., and Ferguson, J. W. (2017). Optically Thin Core Accretion: How Planets Get Their Gas in Nearly Gas-Free Disks. *ArXiv e-prints*.

Lee, E. J., Chiang, E., and Ormel, C. W. (2014). Make Super-Earths, Not Jupiters: Accreting Nebular Gas onto Solid Cores at 0.1 AU and Beyond. ApJ, 797:95.

Lesur, G., Ferreira, J., and Ogilvie, G. I. (2013). The magnetorotational instability as a jet launching mechanism. A&A, 550:A61.

Lesur, G., Kunz, M. W., and Fromang, S. (2014). Thanatology in protoplanetary discs. The combined influence of Ohmic, Hall, and ambipolar diffusion on dead zones. A&A, 566:A56.

Lesur, G. and Papaloizou, J. C. B. (2010). The subcritical baroclinic instability in local accretion disc models. A&A, 513:A60.

Levermore, C. D. and Pomraning, G. C. (1981). A flux-limited diffusion theory. ApJ, 248:321–334.

Lightman, A. P. and Eardley, D. M. (1974). Black Holes in Binary Systems: Instability of Disk Accretion. ApJ, 187:L1.

Lin, D. N. C. and Papaloizou, J. (1980). On the structure and evolution of the primordial solar nebula. MNRAS, 191:37–48.

Liu, B., Zhang, X., Lin, D. N. C., and Aarseth, S. J. (2015). Migration and Growth of Protoplanetary Embryos. II. Emergence of Proto-Gas-Giant Cores versus Super Earth Progenitors. ApJ, 798(1):62.

Lodders, K. (2003). Solar System Abundances and Condensation Temperatures of the Elements. ApJ, 591(2):1220–1247.

Long, F., Herczeg, G. J., Harsono, D., Pinilla, P., Tazzari, M., Manara, C. F., Pascucci, I., Cabrit, S., Nisini, B., Johnstone, D., Edwards, S., Salyk, C., Menard, F., Lodato, G., Boehler, Y., Mace, G. N., Liu, Y., Mulders, G. D., Hendler, N., Ragusa, E., Fischer, W. J., Banzatti, A., Rigliaco, E., van de Plas, G., Dipierro, G., Gully-Santiago, M., and Lopez-Valdivia, R. (2019). Compact Disks in a High-resolution ALMA Survey of Dust Structures in the Taurus Molecular Cloud. ApJ, 882(1):49.

Lopez, E. D. and Fortney, J. J. (2013). The Role of Core Mass in Controlling Evaporation: The Kepler Radius Distribution and the Kepler-36 Density Dichotomy. ApJ, 776:2.

Lopez, E. D. and Fortney, J. J. (2014). Understanding the Mass-Radius Relation for Sub-neptunes: Radius as a Proxy for Composition. ApJ, 792:1.

Lundkvist, M. S., Kjeldsen, H., Albrecht, S., Davies, G. R., Basu, S., Huber, D., Justesen, A. B., Karoff, C., Silva Aguirre, V., van Eylen, V., Vang, C., Arentoft, T., Barclay, T., Bedding, T. R., Campante, T. L., Chaplin, W. J., Christensen-Dalsgaard, J., Elsworth, Y. P., Gilliland, R. L., Handberg, R., Hekker, S., Kawaler, S. D., Lund, M. N., Metcalfe, T. S., Miglio, A., Rowe, J. F., Stello, D., Tingley, B., and White, T. R. (2016). Hot super-Earths stripped by their host stars. *Nature Communications*, 7:11201.

Lynden-Bell, D. and Pringle, J. E. (1974). The evolution of viscous discs and the origin of the nebular variables. MNRAS, 168:603–637.

Lyra, W. and Umurhan, O. M. (2019). The Initial Conditions for Planet Formation: Turbulence Driven by Hydrodynamical Instabilities in Disks around Young Stars. PASP, 131(1001):072001.

Malygin, M. G., Kuiper, R., Klahr, H., Dullemond, C. P., and Henning, T. (2014). Mean gas opacity for circumstellar environments and equilibrium temperature degeneracy. A&A, 568:A91.

Mamajek, E. E. (2009). Initial Conditions of Planet Formation: Lifetimes of Primordial Disks. In Usuda, T., Tamura, M., and Ishii, M., editors, *American Institute of Physics Conference Series*, volume 1158, pages 3–10.

Manara, C. F., Robberto, M., Da Rio, N., Lodato, G., Hillenbrand, L. A., Stassun, K. G., and Soderblom, D. R. (2012). Hubble Space Telescope Measures of Mass Accretion Rates in the Orion Nebula Cluster. ApJ, 755:154.

Manara, C. F., Rosotti, G., Testi, L., Natta, A., Alcalá, J. M., Williams, J. P., Ansdell, M., Miotello, A., van der Marel, N., Tazzari, M., Carpenter, J., Guidi, G., Mathews, G. S., Oliveira, I., Prusti, T., and van Dishoeck, E. F. (2016). Evidence for a correlation between mass accretion rates onto young stars and the mass of their protoplanetary disks. A&A, 591:L3.

Manara, C. F., Testi, L., Herczeg, G. J., Pascucci, I., Alcalá, J. M., Natta, A., Antoniucci, S., Fedele, D., Mulders, G. D., Henning, T., Mohanty, S., Prusti, T., and Rigliaco, E. (2017). X-shooter study of accretion in Chamaeleon I. II. A steeper increase of accretion with stellar mass for very low-mass stars? A&A, 604:A127.

Marcy, G. W., Isaacson, H., Howard, A. W., Rowe, J. F., Jenkins, J. M., Bryson, S. T., Latham, D. W., Howell, S. B., Gautier, Thomas N., I., Batalha, N. M., Rogers, L., Ciardi, D., Fischer, D. A., Gilliland, R. L., Kjeldsen, H., Christensen-Dalsgaard, J., Huber, D., Chaplin, W. J., Basu, S., Buchhave, L. A., Quinn, S. N., Borucki, W. J., Koch, D. G., Hunter, R., Caldwell, D. A., Van Cleve, J., Kolbl, R., Weiss, L. M., Petigura, E., Seager, S., Morton, T., Johnson, J. A., Ballard, S., Burke, C., Cochran, W. D., Endl, M., MacQueen, P., Everett, M. E., Lissauer, J. J., Ford, E. B., Torres, G., Fressin, F., Brown, T. M., Steffen, J. H., Charbonneau, D., Basri, G. S., Sasselov, D. D., Winn, J., Sanchis-Ojeda, R., Christiansen, J., Adams, E., Henze, C., Dupree, A., Fabrycky, D. C., Fortney, J. J., Tarter, J., Holman, M. J., Tenenbaum, P., Shporer, A., Lucas, P. W., Welsh, W. F., Orosz, J. A., Bedding, T. R., Campante, T. L., Davies, G. R., Elsworth, Y., Handberg, R., Hekker, S., Karoff, C., Kawaler, S. D., Lund, M. N., Lundkvist, M., Metcalfe, T. S., Miglio, A., Silva Aguirre, V., Stello, D., White, T. R., Boss, A., Devore, E., Gould, A., Prsa, A., Agol, E., Barclay, T., Coughlin, J., Brugamyer, E., Mullally, F., Quintana, E. V., Still, M., Thompson, S. E., Morrison, D., Twicken, J. D., Désert, J.-M., Carter, J., Crepp, J. R., Hébrard, G., Santerne, A., Moutou, C., Sobeck, C., Hudgins, D., Haas, M. R., Robertson, P., Lillo-Box, J., and Barrado, D. (2014). Masses, Radii, and Orbits of Small Kepler Planets: The Transition from Gaseous to Rocky Planets. ApJS, 210(2):20.

Masset, F. S., Morbidelli, A., Crida, A., and Ferreira, J. (2006). Disk Surface Density Transitions as Protoplanet Traps. ApJ, 642(1):478–487.

Mathis, J. S., Rumpl, W., and Nordsieck, K. H. (1977). The size distribution of interstellar grains. ApJ, 217:425–433.

Mayor, M., Marmier, M., Lovis, C., Udry, S., Ségransan, D., Pepe, F., Benz, W., Bertaux, J. L., Bouchy, F., Dumusque, X., Lo Curto, G., Mordasini, C., Queloz, D., and Santos, N. C. (2011). The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets. *arXiv e-prints*, page arXiv:1109.2497.

McNally, C. P., Nelson, R. P., and Paardekooper, S.-J. (2019). Multiplanet systems in inviscid discs can avoid forming resonant chains. MNRAS, 489(1):L17–L21.

McNeil, D. S. and Nelson, R. P. (2010). On the formation of hot Neptunes and super-Earths. MNRAS, 401(3):1691–1708.

Meru, F., Geretshauser, R. J., Schäfer, C., Speith, R., and Kley, W. (2013). Growth and fragmentation of centimetre-sized dust aggregates: the dependence on aggregate size and porosity. MNRAS, 435:2371–2390.

Mizuno, H. (1980). Formation of the Giant Planets. *Progress of Theoretical Physics*, 64:544–557.

Mohanty, S., Ercolano, B., and Turner, N. J. (2013). Dead, Undead, and Zombie Zones in Protostellar Disks as a Function of Stellar Mass. ApJ, 764:65.

Mohanty, S., Jayawardhana, R., and Basri, G. (2005). The T Tauri Phase Down to Nearly Planetary Masses: Echelle Spectra of 82 Very Low Mass Stars and Brown Dwarfs. ApJ, 626(1):498–522.

Moll, R. (2012). Shearing box simulations of accretion disk winds. A&A, 548:A76.

Morbidelli, A., Chambers, J., Lunine, J. I., Petit, J. M., Robert, F., Valsecchi, G. B., and Cyr, K. E. (2000). Source regions and time scales for the delivery of water to Earth. *Meteoritics and Planetary Science*, 35(6):1309–1320.

Morbidelli, A., Lambrechts, M., Jacobson, S., and Bitsch, B. (2015). The great dichotomy of the Solar System: Small terrestrial embryos and massive giant planet cores. *Icarus*, 258:418–429.

Mulders, G. D., Pascucci, I., Apai, D., and Ciesla, F. J. (2018). The Exoplanet Population Observation Simulator. I. The Inner Edges of Planetary Systems. AJ, 156(1):24.

Muzerolle, J., Calvet, N., and Hartmann, L. (1998). Magnetospheric Accretion Models for the Hydrogen Emission Lines of T Tauri Stars. ApJ, 492(2):743–753.

Muzerolle, J., Calvet, N., and Hartmann, L. (2001). Emission-Line Diagnostics of T Tauri Magnetospheric Accretion. II. Improved Model Tests and Insights into Accretion Physics. ApJ, 550(2):944–961.

Nakagawa, Y., Sekiya, M., and Hayashi, C. (1986). Settling and growth of dust particles in a laminar phase of a low-mass solar nebula. *Icarus*, 67:375–390.

Natta, A., Prusti, T., Neri, R., Wooden, D., Grinin, V. P., and Mannings, V. (2001). A reconsideration of disk properties in Herbig Ae stars. A&A, 371:186–197.

Nelson, R. P. (2018). Planetary Migration in Protoplanetary Disks. In *Handbook of Exoplanets*, page 139. Springer International Publishing AG, part of Springer Nature.

Nelson, R. P., Gressel, O., and Umurhan, O. M. (2013). Linear and non-linear evolution of the vertical shear instability in accretion discs. MNRAS, 435(3):2610–2632.

Öberg, K. I., Murray-Clay, R., and Bergin, E. A. (2011). The Effects of Snowlines on C/O in Planetary Atmospheres. ApJ, 743:L16.

O'dell, C. R. and Wen, Z. (1994). Postrefurbishment Mission Hubble Space Telescope Images of the Core of the Orion Nebula: Proplyds, Herbig-Haro Objects, and Measurements of a Circumstellar Disk. ApJ, 436:194.

Ogihara, M. and Ida, S. (2009). N-Body Simulations of Planetary Accretion Around M Dwarf Stars. ApJ, 699(1):824–838.

Ormel, C. W. and Cuzzi, J. N. (2007a). Closed-form expressions for particle relative velocities induced by turbulence. A&A, 466:413–420.

Ormel, C. W. and Cuzzi, J. N. (2007b). Closed-form expressions for particle relative velocities induced by turbulence. A&A, 466(2):413–420.

Ormel, C. W., Shi, J.-M., and Kuiper, R. (2015). Hydrodynamics of embedded planets' first atmospheres - II. A rapid recycling of atmospheric gas. MNRAS, 447:3512–3525.

Owen, J. E. (2014). Accreting Planets as Dust Dams in "Transition" Disks. ApJ, 789:59.

Owen, J. E. and Wu, Y. (2013). Kepler Planets: A Tale of Evaporation. ApJ, 775(2):105.

Owen, J. E. and Wu, Y. (2016). Atmospheres of Low-mass Planets: The “Boil-off”. ApJ, 817:107.

Owen, J. E. and Wu, Y. (2017). The Evaporation Valley in the Kepler Planets. ApJ, 847(1):29.

Paardekooper, S. J. and Mellema, G. (2006). Halting type I planet migration in non-isothermal disks. A&A, 459(1):L17–L20.

Panić, O., Hogerheijde, M. R., Wilner, D., and Qi, C. (2009). A break in the gas and dust surface density of the disc around the T Tauri star IM Lupi. A&A, 501:269–278.

Papaloizou, J. C. B. and Terquem, C. (1999). Critical Protoplanetary Core Masses in Protoplanetary Disks and the Formation of Short-Period Giant Planets. ApJ, 521(2):823–838.

Pfeil, T. and Klahr, H. (2019). Mapping the Conditions for Hydrodynamic Instability on Steady-State Accretion Models of Protoplanetary Disks. ApJ, 871(2):150.

Pinilla, P., Birnstiel, T., Ricci, L., Dullemond, C. P., Uribe, A. L., Testi, L., and Natta, A. (2012). Trapping dust particles in the outer regions of protoplanetary disks. A&A, 538:A114.

Pollack, J. B., Hollenbach, D., Beckwith, S., Simonelli, D. P., Roush, T., and Fong, W. (1994). Composition and Radiative Properties of Grains in Molecular Clouds and Accretion Disks. ApJ, 421:615.

Poon, S. T. S., Nelson, R. P., Jacobson, S. A., and Morbidelli, A. (2020). Formation of compact systems of super-Earths via dynamical instabilities and giant impacts. MNRAS, 491(4):5595–5620.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2002). *Numerical Recipes in C++ : The Art of Scientific Computing.* Cambridge University Press, Cambridge, UK.

Pringle, J. E. (1981). Accretion discs in astrophysics. ARA&A, 19:137–162.

Rafikov, R. R. (2006). Atmospheres of Protoplanetary Cores: Critical Mass for Nucleated Instability. ApJ, 648:666–682.

Rafikov, R. R. (2007). Convective Cooling and Fragmentation of Gravitationally Unstable Disks. ApJ, 662(1):642–650.

Raymond, S. N., Barnes, R., and Mandell, A. M. (2008). Observable consequences of planet formation models in systems with close-in terrestrial planets. MNRAS, 384(2):663–674.

Rogers, L. A. (2015). Most 1.6 Earth-radius Planets are Not Rocky. ApJ, 801(1):41.

Rogers, L. A., Bodenheimer, P., Lissauer, J. J., and Seager, S. (2011). Formation and Structure of Low-density exo-Neptunes. ApJ, 738:59.

Rosenfeld, K. A., Andrews, S. M., Wilner, D. J., Kastner, J. H., and McClure, M. K. (2013). The Structure of the Evolved Circumbinary Disk around V4046 Sgr. ApJ, 775:136.

Rouse, C. A. (1961). Ionization Equilibrium Equation of State. ApJ, 134:435.

Ruediger, G., Elstner, D., and Tschaepe, R. (1988). Vertical structure of thin Keplerian disks. I - The role of the turbulent Prandtl number. Acta Astron., 38(4):299–314.

Salmeron, R. and Wardle, M. (2008). Magnetorotational instability in protoplanetary discs: the effect of dust grains. MNRAS, 388:1223–1238.

Sano, T., Inutsuka, S.-i., Turner, N. J., and Stone, J. M. (2004). Angular Momentum Transport by Magnetohydrodynamic Turbulence in Accretion Disks: Gas Pressure Dependence of the Saturation Level of the Magnetorotational Instability. ApJ, 605(1):321–339.

Sano, T., Miyama, S. M., Umebayashi, T., and Nakano, T. (2000). Magnetorotational Instability in Protoplanetary Disks. II. Ionization State and Unstable Regions. ApJ, 543:486–501.

Sano, T. and Stone, J. M. (2002). The Effect of the Hall Term on the Nonlinear Evolution of the Magnetorotational Instability. II. Saturation Level and Critical Magnetic Reynolds Number. ApJ, 577(1):534–553.

Santerne, A., Moutou, C., Tsantaki, M., Bouchy, F., Hébrard, G., Adibekyan, V., Almenara, J. M., Amard, L., Barros, S. C. C., Boisse, I., Bonomo, A. S., Bruno, G., Courcol, B., Deleuil, M., Demangeon, O., Díaz, R. F., Guillot, T., Havel, M., Montagnier, G., Rajpurohit, A. S., Rey, J., and Santos, N. C. (2016). SOPHIE velocimetry of Kepler transit candidates. XVII. The physical properties of giant exoplanets within 400 days of period. A&A, 587:A64.

Sargent, A. I. and Beckwith, S. (1987). Kinematics of the Circumstellar Gas of HL Tauri and R Monocerotis. ApJ, 323:294.

Schäfer, U., Yang, C.-C., and Johansen, A. (2017). Initial mass function of planetesimals formed by the streaming instability. A&A, 597:A69.

Schlichting, H. E. (2014). Formation of Close in Super-Earths and Mini-Neptunes: Required Disk Masses and their Implications. ApJ, 795(1):L15.

Schoonenberg, D., Liu, B., Ormel, C. W., and Dorn, C. (2019). Pebble-driven planet formation for TRAPPIST-1 and other compact systems. A&A, 627:A149.

Schoonenberg, D. and Ormel, C. W. (2017). Planetesimal formation near the snowline: in or out? A&A, 602:A21.

Schoonenberg, D., Ormel, C. W., and Krijt, S. (2018). A Lagrangian model for dust evolution in protoplanetary disks: formation of wet and dry planetesimals at different stellar masses. A&A, 620:A134.

Schräpler, R. and Blum, J. (2011). The Physics of Protoplanetesimal Dust Agglomerates. VI. Erosion of Large Aggregates as a Source of Micrometer-sized Particles. ApJ, 734:108.

Schräpler, R., Blum, J., Seizinger, A., and Kley, W. (2012). The Physics of Protoplanetesimal Dust Agglomerates. VII. The Low- velocity Collision Behavior of Large Dust Agglomerates. ApJ, 758:35.

Shakura, N. I. and Sunyaev, R. A. (1973). Reprint of 1973A&amp;A....24..337S. Black holes in binary systems. Observational appearance. A&A, 500:33–51.

Shi, J., Krolik, J. H., and Hirose, S. (2010). What is the Numerically Converged Amplitude of Magnetohydrodynamics Turbulence in Stratified Shearing Boxes? ApJ, 708(2):1716–1727.

Shu, F. H. (1992). *The Physics of Astrophysics. Volume II: Gas Dynamics.* University Science Books, Mill Valley, CA.

Simon, J. B., Armitage, P. J., Li, R., and Youdin, A. N. (2016). The Mass and Size Distribution of Planetesimals Formed by the Streaming Instability. I. The Role of Self-gravity. ApJ, 822(1):55.

Simon, J. B., Armitage, P. J., Youdin, A. N., and Li, R. (2017). Evidence for Universality in the Initial Planetesimal Mass Function. ApJ, 847(2):L12.

Simon, J. B., Hawley, J. F., and Beckwith, K. (2009). Simulations of Magnetorotational Turbulence with a Higher-Order Godunov Scheme. ApJ, 690(1):974–997.

Simon, J. B., Lesur, G., Kunz, M. W., and Armitage, P. J. (2015). Magnetically driven accretion in protoplanetary discs. MNRAS, 454(1):1117–1131.

Simon, M., Dutrey, A., and Guilloteau, S. (2000). Dynamical Masses of T Tauri Stars and Calibration of Pre-Main-Sequence Evolution. ApJ, 545(2):1034–1043.

Squire, J. and Hopkins, P. F. (2018). Resonant Drag Instability of Grains Streaming in Fluids. ApJ, 856:L15.

Stoll, M. H. R. and Kley, W. (2014). Vertical shear instability in accretion disc models with radiation transport. A&A, 572:A77.

Stone, J. M., Hawley, J. F., Gammie, C. F., and Balbus, S. A. (1996). Three-dimensional Magnetohydrodynamical Simulations of Vertically Stratified Accretion Disks. ApJ, 463:656.

Suzuki, T. K. and Inutsuka, S.-i. (2009). Disk Winds Driven by Magnetorotational Instability and Dispersal of Protoplanetary Disks. ApJ, 691(1):L49–L54.

Suzuki, T. K., Muto, T., and Inutsuka, S.-i. (2010). Protoplanetary Disk Winds via Magnetorotational Instability: Formation of an Inner Hole and a Crucial Assist for Planet Formation. ApJ, 718(2):1289–1304.

Takeuchi, T. and Lin, D. N. C. (2002). Radial Flow of Dust Particles in Accretion Disks. ApJ, 581:1344–1355.

Taki, T., Fujimoto, M., and Ida, S. (2016). Dust and gas density evolution at a radial pressure bump in protoplanetary disks. A&A, 591:A86.

Teague, R., Bae, J., Birnstiel, T., and Bergin, E. A. (2018). Evidence for a Vertical Dependence on the Pressure Structure in AS 209. ApJ, 868(2):113.

Terebey, S., Shu, F. H., and Cassen, P. (1984). The collapse of the cores of slowly rotating isothermal clouds. ApJ, 286:529–551.

Terquem, C. and Papaloizou, J. C. B. (2007). Migration and the Formation of Systems of Hot Super-Earths and Neptunes. ApJ, 654(2):1110–1120.

Terquem, C. E. J. M. L. J. (2008). New Composite Models of Partially Ionized Protoplanetary Disks. ApJ, 689(1):532–538.

Testi, L., Birnstiel, T., Ricci, L., Andrews, S., Blum, J., Carpenter, J., Dominik, C., Isella, A., Natta, A., Williams, J. P., and Wilner, D. J. (2014). Dust Evolution in Protoplanetary Disks. In Beuther, H., Klessen, R. S., Dullemond, C. P., and Henning, T., editors, *Protostars and Planets VI*, page 339.

Trapman, L., Ansdell, M., Hogerheijde, M. R., Facchini, S., Manara, C. F., Miotello, A., Williams, J. P., and Bruderer, S. (2020). Constraining the radial drift of millimeter-sized grains in the protoplanetary disks in Lupus. *arXiv e-prints*, page arXiv:2004.07257.

Tu, L., Johnstone, C. P., Güdel, M., and Lammer, H. (2015). The extreme ultraviolet and X-ray Sun in Time: High-energy evolutionary tracks of a solar-like star. A&A, 577:L3.

Turner, N. J., Sano, T., and Dziourkevitch, N. (2007). Turbulent Mixing and the Dead Zone in Protostellar Disks. ApJ, 659(1):729–737.

Umebayashi, T. and Nakano, T. (1981). Fluxes of Energetic Particles and the Ionization Rate in Very Dense Interstellar Clouds. PASJ, 33:617.

Umebayashi, T. and Nakano, T. (1988). Chapter 13. Ionization State and Magnetic Fields in the Solar Nebula. *Progress of Theoretical Physics Supplement*, 96:151–160.

Umebayashi, T. and Nakano, T. (2009). Effects of Radionuclides on the Ionization State of Protoplanetary Disks and Dense Cloud Cores. ApJ, 690(1):69–81.

Valencia, D., Ikoma, M., Guillot, T., and Nettelmann, N. (2010). Composition and fate of short-period super-Earths. The case of CoRoT-7b. A&A, 516:A20.

Van Eylen, V., Agentoft, C., Lundkvist, M. S., Kjeldsen, H., Owen, J. E., Fulton, B. J., Petigura, E., and Snellen, I. (2018). An asteroseismic view of the radius valley: stripped cores, not born rocky. MNRAS, 479(4):4786–4795.

Wardle, M. (1999). The Balbus-Hawley instability in weakly ionized discs. MNRAS, 307(4):849–856.

Wardle, M. (2007). Magnetic fields in protoplanetary disks. Ap&SS, 311:35–45.

Weidenschilling, S. J. (1977). Aerodynamics of solid bodies in the solar nebula. MNRAS, 180:57–70.

Weidling, R., Güttler, C., Blum, J., and Brauer, F. (2009). The Physics of Protoplanetesimal Dust Agglomerates. III. Compaction in Multiple Collisions. ApJ, 696:2036–2043.

Weiss, L. M. and Marcy, G. W. (2014). The Mass-Radius Relation for 65 Exoplanets Smaller than 4 Earth Radii. ApJ, 783(1):L6.

Williams, J. P. and Cieza, L. A. (2011). Protoplanetary Disks and Their Evolution. ARA&A, 49(1):67–117.

Windmark, F., Birnstiel, T., Güttler, C., Blum, J., Dullemond, C. P., and Henning, T. (2012a). Planetesimal formation by sweep-up: how the bouncing barrier can be beneficial to growth. A&A, 540:A73.

Windmark, F., Birnstiel, T., Ormel, C. W., and Dullemond, C. P. (2012b). Breaking through: The effects of a velocity distribution on barriers to dust growth. A&A, 544:L16.

Wolfgang, A. and Lopez, E. (2015). How Rocky Are They? The Composition Distribution of Kepler's Sub-Neptune Planet Candidates within 0.15 AU. ApJ, 806(2):183.

Wolfgang, A., Rogers, L. A., and Ford, E. B. (2016). Probabilistic Mass-Radius Relationship for Sub-Neptune-Sized Planets. ApJ, 825:19.

Wood, K., Wolff, M. J., Bjorkman, J. E., and Whitney, B. (2002). The Spectral Energy Distribution of HH 30 IRS: Constraining the Circumstellar Dust Size Distribution. ApJ, 564(2):887–895.

Wright, N. J., Drake, J. J., Mamajek, E. E., and Henry, G. W. (2011). The Stellar-activity-Rotation Relationship and the Evolution of Stellar Dynamos. ApJ, 743(1):48.

Wu, Y. (2019). Mass and Mass Scalings of Super-Earths. ApJ, 874(1):91.

Yang, C.-C., Johansen, A., and Carrera, D. (2017). Concentrating small particles in protoplanetary disks through the streaming instability. A&A, 606:A80.

Youdin, A. N. and Chiang, E. I. (2004). Particle Pileups and Planetesimal Formation. ApJ, 601:1109–1119.

Youdin, A. N. and Goodman, J. (2005). Streaming Instabilities in Protoplanetary Disks. ApJ, 620:459–469.

Youdin, A. N. and Lithwick, Y. (2007). Particle stirring in turbulent gas disks: Including orbital oscillations. *Icarus*, 192:588–604.

Youdin, A. N. and Shu, F. H. (2002). Planetesimal Formation by Gravitational Instability. ApJ, 580:494–505.

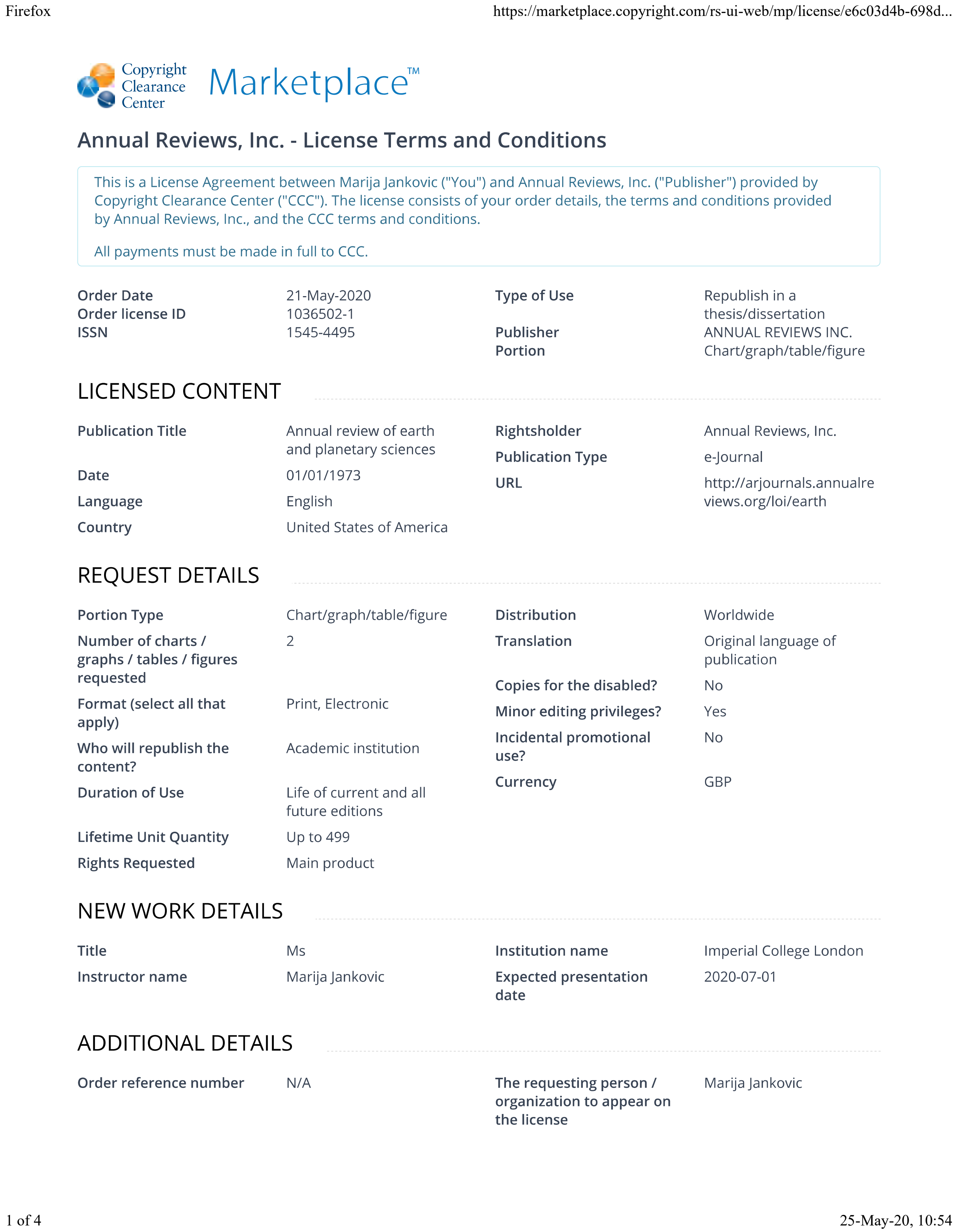
Zeng, L. and Sasselov, D. (2013). A Detailed Model Grid for Solid Planets from 0.1 through 100 Earth Masses. PASP, 125(925):227.

Zhu, Z., Hartmann, L., and Gammie, C. (2009). Nonsteady Accretion in Protostars. ApJ, 694(2):1045–1055.

Zsom, A., Ormel, C. W., Güttler, C., Blum, J., and Dullemond, C. P. (2010). The outcome of protoplanetary dust growth: pebbles, boulders, or planetesimals? II. Introducing the bouncing barrier. A&A, 513:A57.

Permissions

Page No.	Type of work:	Name of work	Source of work	Copyright holder and contact	permission requested on	I have permission yes /no	Permission note
13	figures	Fig. 5 Massradius diagram for exoplanets..., Fig. 7 Planetary massradius diagram...	D. Jontof-Hutter, The Compositional Diversity of Low-Mass Exoplanets, Annual Review of Earth and Planetary Sciences (2019), vol 47, page 155	Annual Reviews, Inc.; marketplace.copyright.com (Copyright Clearance Center)	21.05.20	yes	License obtained via Copyright Clearance Center
14	figure	Fig. 7 Top: completeness-corrected histogram of planet radii...	B. Fulton et al., The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets, The Astronomical Journal (2017), Volume 154, Issue 3, article id. 109, 19 pp.	AAS; permissions@ioppublishing.org	21.05.20	yes	Reproduced with the permission of B. Fulton and AAS



Firefox https://marketplace.copyright.com/rs-ui-web/mp/license/e6c03d4b-698d...

Copyright Clearance Center Marketplace™

Annual Reviews, Inc. - License Terms and Conditions

This is a License Agreement between Marija Jankovic ("You") and Annual Reviews, Inc. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Annual Reviews, Inc., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	21-May-2020	**Type of Use**	Republish in a thesis/dissertation
Order license ID	1036502-1	**Publisher**	ANNUAL REVIEWS INC.
ISSN	1545-4495	**Portion**	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Annual review of earth and planetary sciences	**Rightsholder**	Annual Reviews, Inc.
Date	01/01/1973	**Publication Type**	e-Journal
Language	English	**URL**	http://arjournals.annualreviews.org/loi/earth
Country	United States of America		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	**Distribution**	Worldwide
Number of charts / graphs / tables / figures requested	2	**Translation**	Original language of publication
Format (select all that apply)	Print, Electronic	**Copies for the disabled?**	No
Who will republish the content?	Academic institution	**Minor editing privileges?**	Yes
Duration of Use	Life of current and all future editions	**Incidental promotional use?**	No
Lifetime Unit Quantity	Up to 499	**Currency**	GBP
Rights Requested	Main product		

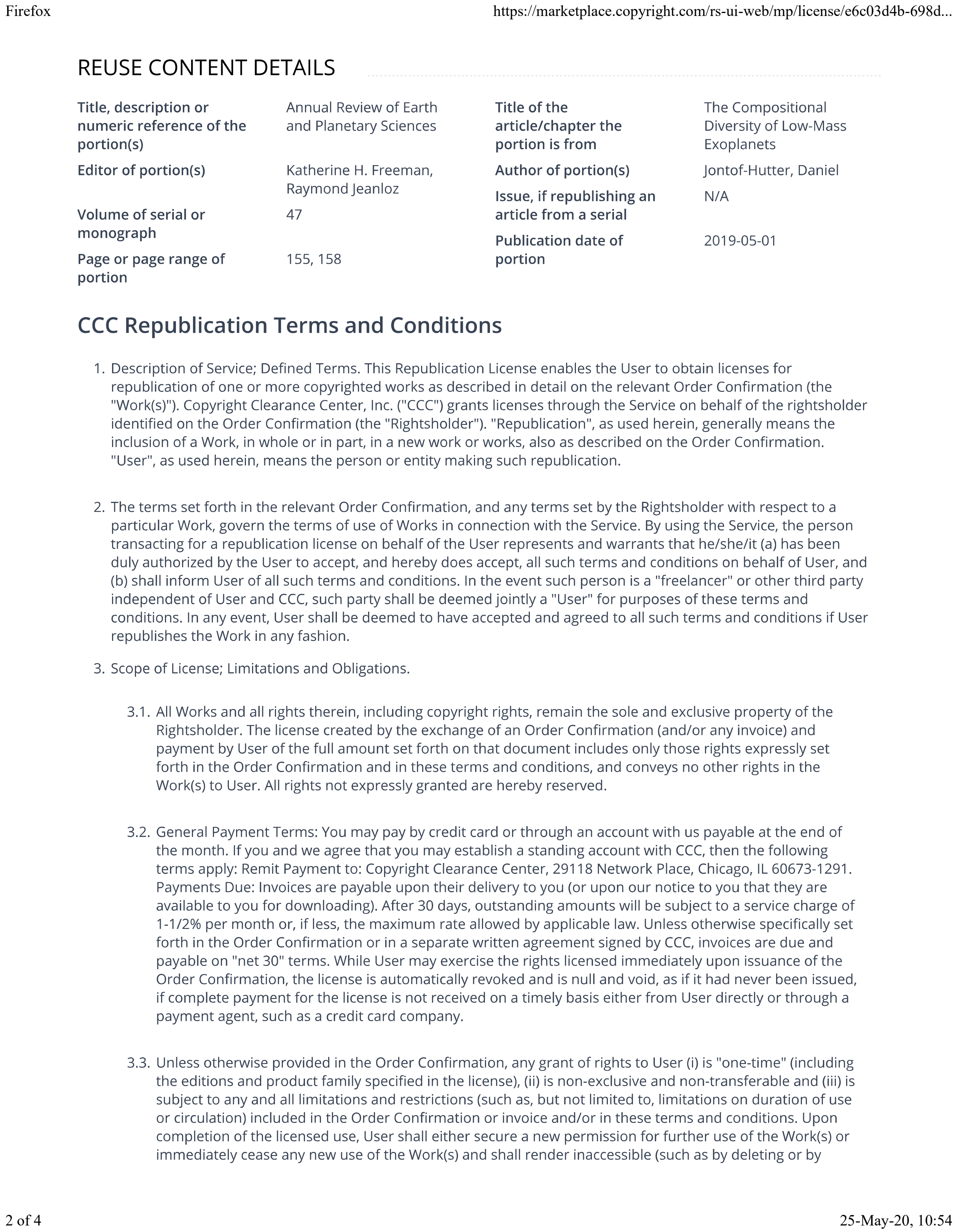
NEW WORK DETAILS

Title	Ms	**Institution name**	Imperial College London
Instructor name	Marija Jankovic	**Expected presentation date**	2020-07-01

ADDITIONAL DETAILS

Order reference number	N/A	**The requesting person / organization to appear on the license**	Marija Jankovic

1 of 4 25-May-20, 10:54



Firefox https://marketplace.copyright.com/rs-ui-web/mp/license/e6c03d4b-698d...

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Annual Review of Earth and Planetary Sciences	**Title of the article/chapter the portion is from**	The Compositional Diversity of Low-Mass Exoplanets
Editor of portion(s)	Katherine H. Freeman, Raymond Jeanloz	**Author of portion(s)**	Jontof-Hutter, Daniel
Volume of serial or monograph	47	**Issue, if republishing an article from a serial**	N/A
Page or page range of portion	155, 158	**Publication date of portion**	2019-05-01

CCC Republication Terms and Conditions

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

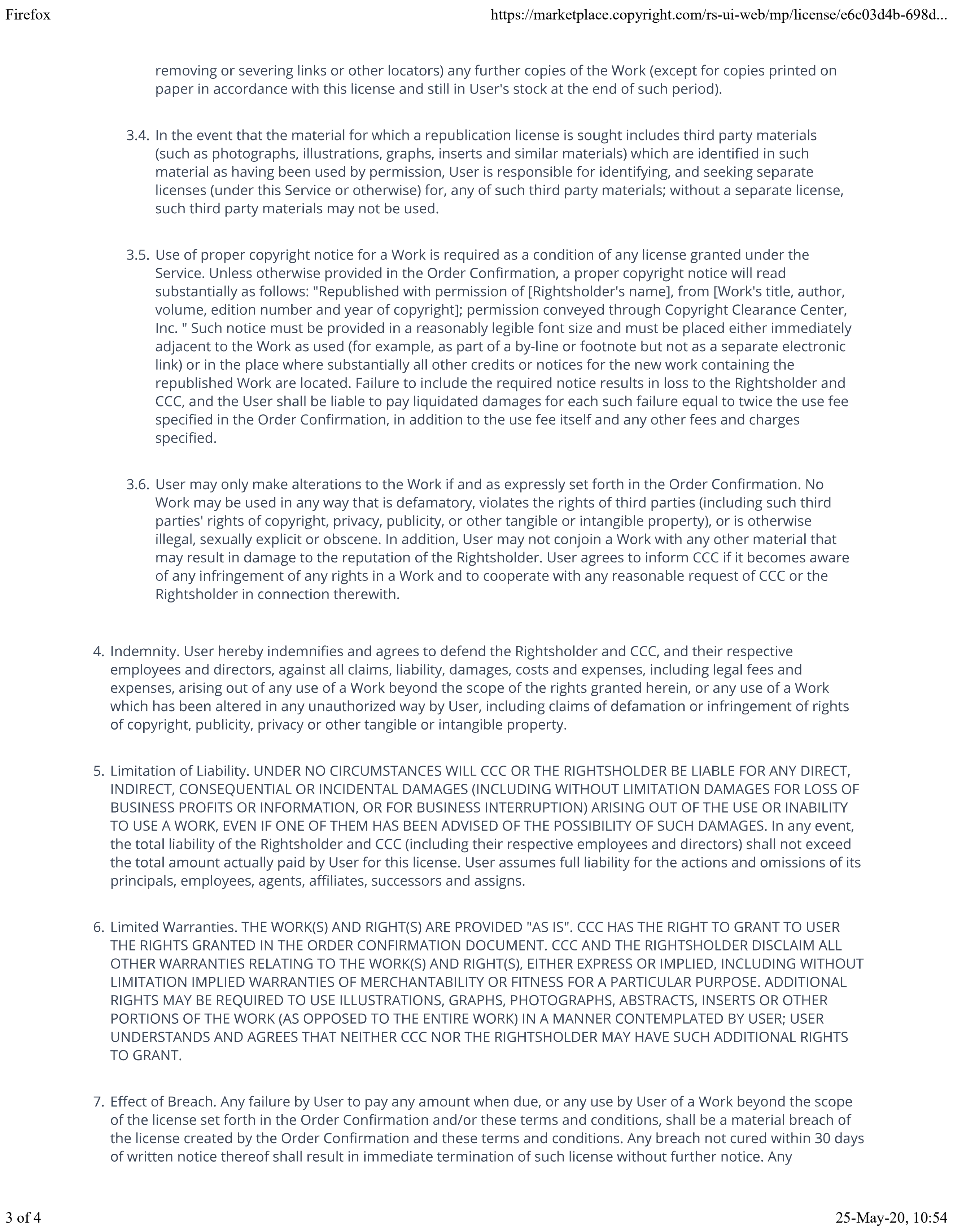
3. Scope of License; Limitations and Obligations.

    3.1. All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

    3.2. General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 29118 Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

    3.3. Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by

2 of 4 25-May-20, 10:54



removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4. In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5. Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

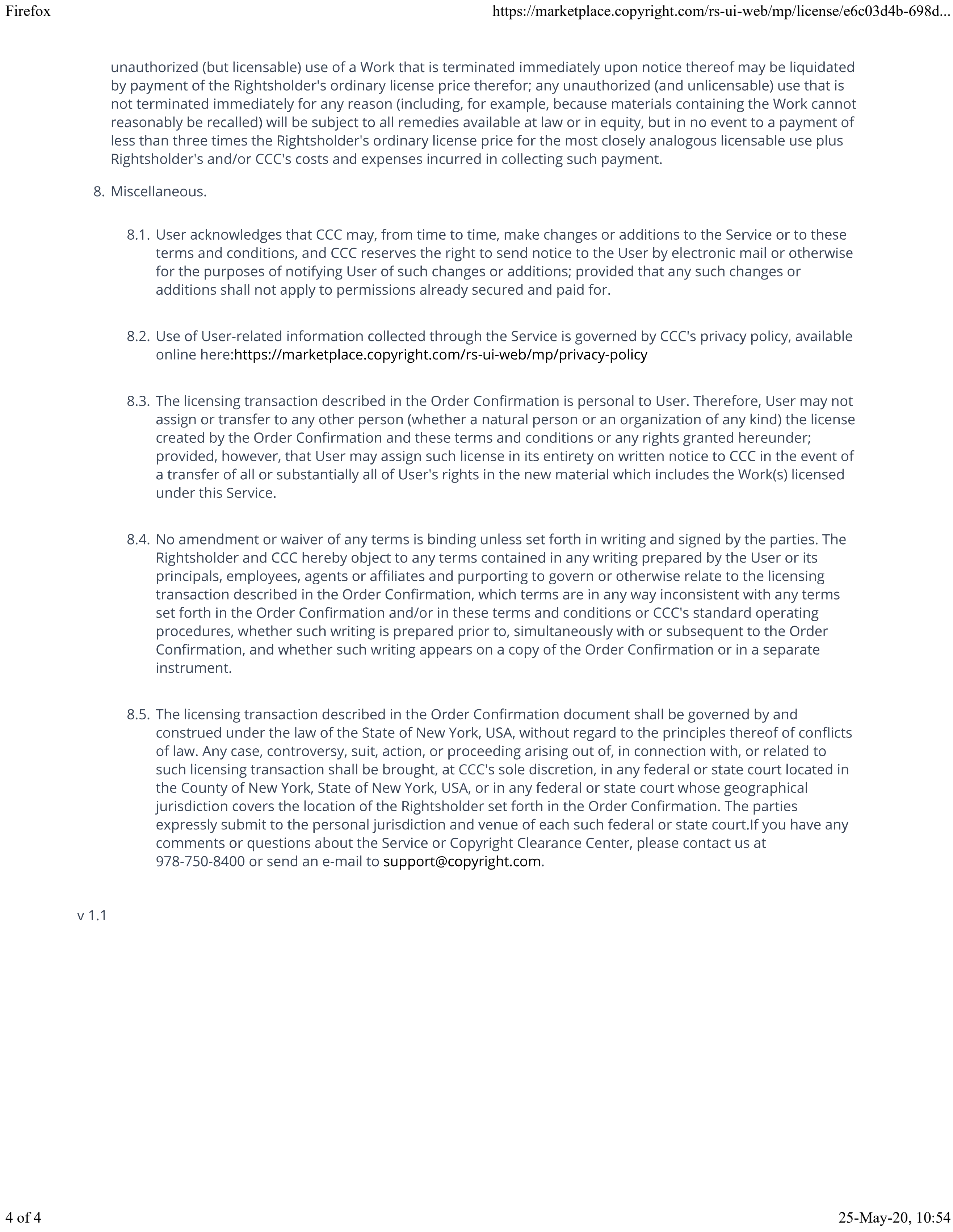
3.6. User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any



Firefox https://marketplace.copyright.com/rs-ui-web/mp/license/e6c03d4b-698d...

unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. Miscellaneous.

   8.1. User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

   8.2. Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here:https://marketplace.copyright.com/rs-ui-web/mp/privacy-policy

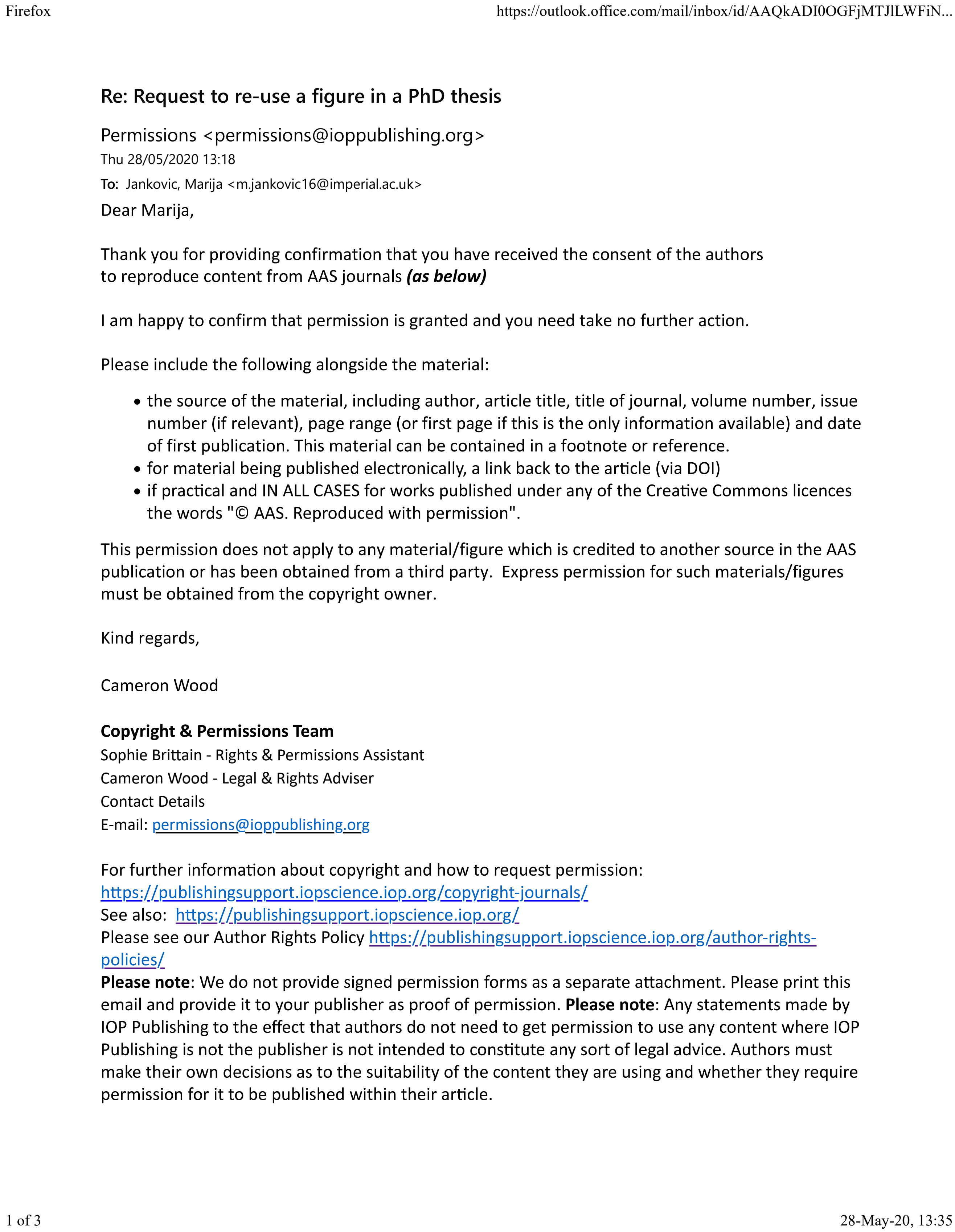
   8.3. The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

   8.4. No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

   8.5. The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to support@copyright.com.

v 1.1

4 of 4 25-May-20, 10:54



Firefox https://outlook.office.com/mail/inbox/id/AAQkADI0OGFjMTJlLWFiN...

Re: Request to re-use a figure in a PhD thesis

Permissions <permissions@ioppublishing.org>
Thu 28/05/2020 13:18
To: Jankovic, Marija <m.jankovic16@imperial.ac.uk>

Dear Marija,

Thank you for providing confirmation that you have received the consent of the authors to reproduce content from AAS journals ***(as below)***

I am happy to confirm that permission is granted and you need take no further action.

Please include the following alongside the material:

- the source of the material, including author, article title, title of journal, volume number, issue number (if relevant), page range (or first page if this is the only information available) and date of first publication. This material can be contained in a footnote or reference.
- for material being published electronically, a link back to the article (via DOI)
- if practical and IN ALL CASES for works published under any of the Creative Commons licences the words "© AAS. Reproduced with permission".

This permission does not apply to any material/figure which is credited to another source in the AAS publication or has been obtained from a third party. Express permission for such materials/figures must be obtained from the copyright owner.

Kind regards,

Cameron Wood

Copyright & Permissions Team
Sophie Brittain - Rights & Permissions Assistant
Cameron Wood - Legal & Rights Adviser
Contact Details
E-mail: permissions@ioppublishing.org

For further information about copyright and how to request permission:
https://publishingsupport.iopscience.iop.org/copyright-journals/
See also: https://publishingsupport.iopscience.iop.org/
Please see our Author Rights Policy https://publishingsupport.iopscience.iop.org/author-rights-policies/
Please note: We do not provide signed permission forms as a separate attachment. Please print this email and provide it to your publisher as proof of permission. **Please note**: Any statements made by IOP Publishing to the effect that authors do not need to get permission to use any content where IOP Publishing is not the publisher is not intended to constitute any sort of legal advice. Authors must make their own decisions as to the suitability of the content they are using and whether they require permission for it to be published within their article.

1 of 3 28-May-20, 13:35

Република Србија
Агенција за квалификације
Број: 612-01-03-3-281/2021
Датум: 21.04.2021. године
Мајке Јевросиме 51
Београд
мк

На основу члана 38. и члана 5. став 1. тачка 10. Закона о националном оквиру квалификација Републике Србије („Сл. гласник РС", бр. 27/18 и 6/20), члана 131. став 1. Закона о високом образовању („Сл. гласник РС", бр. 88/17, 27/18 – др. Закон, 73/18,67/19 и 6/20- др. закон), и члана 136. став 1. Закона о општем управном поступку („Сл. гласник РС", бр. 18/16 и 95/18 – Аутентично тумачење), решавајући по захтеву Марије Јанковић из Београда, Република Србија, за признавање високошколске исправе издате у Уједињеном Краљевству Велике Британије и Северне Ирске, ради запошљавања,

директор Агенције за квалификације доноси

РЕШЕЊЕ

1. Диплома издата 01.10.2020. године од стране Империал Колеџа Лондон (Imperial College London), Лондон, Уједињено Краљевство Велике Британије и Северне Ирске, на име Марија Јанковић, рођена 30.07.1991. године у Београду, о завршеним докторским академским студијама високог образовања, студијски програм: Физика, дисертација: „Акрециони дискови и стварање планета око младих звезда", звање/квалификација: Doctor of Philosophy/Доктор наука (на основу превода овлашћеног судског тумача за енглески језик), **признаје се** као диплома докторских академских студија трећег степена високог образовања (180 ЕСПБ), у оквиру образовно-научног поља: Природно-математичке науке, научна односно стручна област: Физичке науке, која одговара нивоу 8 НОКС-а, ради запошљавања.
2. Ово решење омогућава имаоцу општи приступ тржишту рада у Републици Србији, али га не ослобађа од испуњавања посебних услова за бављење професијама које су регулисане законом или другим прописом.
3. Звање/квалификација из тачке 1. диспозитива овог решења које је са оригиналне стране јавне исправе превео овлашћени судски тумач за енглески језик, не представља стручни, академски, научни односно уметнички назив који у складу са чланом 12. ставом 1. тачка 9. Закона о високом образовању, утврђује Национални савет за високо образовање.

Образложење

Агенцији за квалификације обратила се Марија Јанковић из Београда, Република Србија, захтевом од 05.04.2021. године, за признавање дипломе Империал Колеџ Лондон, Лондон, Уједињено Краљевство Велике Британије и Северне Ирске, докторске академске студије високог образовања, студијски програм: Физика, дисертација: „Акрециони дискови и стварање планета око младих звезда", звање/квалификација: Doctor of Philosophy/Доктор наука, ради запошљавања.

Уз захтев, подносилац захтева доставио је:

1) оверену копију дипломе издате 01.10.2020. године од стране Империал Колеџа Лондон, Лондон, Уједињено Краљевство Велике Британије и Северне Ирске, звање/квалификација Doctor of Philosophy/Доктор наука; студијски програм: Физика, звање/квалификација Doctor of Philosophy/Доктор наука;

2) оверени превод дипломе на српски језик;
3) оверену копију транскрипта испита;
4) примерак дисертације на изворном језику;
5) копију уверења о дипломирању са мастер студија;
6) апстракт докторског рада;
7) списак научних радова;
8) радну биографију;
9) пријавни формулар;
10) доказ о уплати накнаде за професионално признавање.

Одредбом члана 136. став 1. Закона о општем управном поступку прописано је да се решењем одлучује о праву, обавези или правном интересу странке.

Одредбом члана 38. став 1. Закона о националном оквиру квалификација Републике Србије прописано је да захтев за професионално признавање заинтересовано лице подноси Агенцији. Ставом 2. наведеног члана прописано је да професионално признавање врши ENIC/NARIC центар, као организациони део Агенције, по претходно извршеном вредновању страног студијског програма, у складу са овим и законом који уређује високо образовање. Ставом 3. наведеног члана прописано је да вредновање страног студијског програма из става 2. овог члана, уколико међународним уговором није предвиђено другачије, врши се на основу врсте и нивоа постигнутих компетенција стечених завршетком студијског програма, узимајући у обзир систем образовања, односно систем квалификација у земљи у којој је високошколска исправа стечена, услова уписа, права која проистичу из стране високошколске исправе у земљи у којој је стечена и других релевантних чињеница, без разматрања формалних обележја и структуре студијског програма, у складу са принципима Конвенције о признавању квалификација из области високог образовања у европском региону ("Службени лист СЦГ - Међународни уговори", број 7/03), као што је уређено и одредбом члана 131. став 1. Закона о високом образовању. Ставом 4. наведеног члана прописано је да решење о професионалном признавању посебно садржи: назив, врсту, степен и трајање (обим) студијског програма, односно квалификације, који је наведен у страној високошколској исправи - на изворном језику и у преводу на српски језик и научну, уметничку, односно стручну област у оквиру које је остварен студијски програм, односно врсту и ниво квалификације у Републици Србији и ниво НОКС-а којем квалификација одговара. Ставом 5. наведеног члана прописано је да директор Агенције доноси решење о професионалном признавању у року од 60 дана од дана пријема уредног захтева. Ставом 6. наведеног члана прописано је да решење из става 4. овог члана не ослобађа имаоца од испуњавања посебних услова за обављање одређене професије прописане посебним законом. Ставом 7. наведеног члана прописано је да је решење о професионалном признавању коначно. Ставом 8. наведеног члана прописано је да изузетно од става 3. овог члана, уколико је високошколска исправа стечена на једном од првих 500 универзитета рангираних на једној од последње објављених међународних листа рангирања универзитета у свету Shanghai ranking consultancy (Шангајска листа), US News and World Report Ranking (листа рејтинга US News and World Report) или The Times Higher Education World University Rankings (Тајмсова листа рејтинга светских универзитета) решење о професионалном признавању доноси се без спровођења поступка вредновања страног студијског програма из става 2. овог члана у року од осам дана од дана пријема уредног захтева. Ставом 9. наведеног члана прописано је да се, уколико није другачије прописано, на поступак професионалног признавања примењује закон којим се уређује општи управни поступак. Ставом 10. наведеног члана прописано је да решење о професионалном признавању има значај јавне исправе. Ставом 11. наведеног члана прописано је да ближе услове у погледу начина спровођења поступка професионалног признавања прописује министар надлежан за послове образовања.

Одредбом члана 5. став 1. тачка 10. Закона о националном оквиру квалификација Републике Србије, прописано је да се осми ниво стиче завршавањем докторских студија обима 180 ЕСПБ бодова (уз претходно завршене интегрисане академске, односно мастер академске студије).

Одлучујући о захтеву подносиоца, извршено је вредновање страног студијског програма на основу врсте и нивоа постигнутих компетенција стечених завршетком студијског програма, узимајући у обзир систем образовања у земљи у којој је високошколска исправа стечена, услове уписа, права која проистичу из стране високошколске исправе у земљи у којој је стечена и друге релевантне чињенице,без разматрања формалних обележја и структуре студијског програма, одлучено је да се диплома Империал Колеџа Лондон, Лондон, Уједињено Краљевство Велике Британије и Северне Ирске, може признати као диплома докторских академских студија трећег степена високог образовања (180 ЕСПБ), која одговара нивоу 8 НОКС-а.

Са напред наведених разлога директор Агенције је нашао да су у конкретном случају испуњени претходно наведени сви законом прописани услови да се призна диплома Империал Колеџа Лондон, Лондон, Уједињено Краљевство Велике Британије и Северне Ирске, као диплома докторских академских студија трећег степена високог образовања (180 ЕСПБ), у оквиру образовно-научног поља: Природно-математичке науке, научна односно стручна област: Физичке науке, која одговара нивоу 8 НОКС-а, ради запошљавања.

Накнада за решење по захтеву се наплаћује на основу члана 2. став 3. Правилника о висини накнаде за трошкове поступка признавања страних школских исправа и признавање страних високошколских исправа у сврху запошљавања и о висини накнада за трошкове поступка давања одобрења другој организацији за стицање статуса јавно признатог организатора активности образовања одраслих ("Службени гласник РС", бр. 1/2020) плаћена је и поништена.

Сходно претходно наведеном, донета је одлука као у диспозитиву решења.

Упутство о правном средству: Ово решење је коначно у управном поступку и против истог може се покренути управни спор. Тужба се подноси Управном суду у року од 30 дана од дана пријема овог решења.

Решење доставити:
- Марија Јанковић;
- Архиви.

ДИРЕКТОР

проф. др Часлав Митровић

From Stars to Planets II - Connecting our understanding of star and planet formation

Monday June 17th - Friday June 21st 2019, Chalmers University of Technology, Gothenburg, Sweden
http://cosmicorigins.space/fstpii

Sponsored by: Chalmers Initiative on Cosmic Origins (CICO) & Vetenskapsrådet (Swedish Research Council)

SOC: Jonathan C. Tan (Chair, Chalmers/UVa), Ilse Cleeves (UVa), Maria Drozdovskaya (Bern), Eric Herbst (UVa), Jouni Kainulainen (Chalmers), Zhi-Yun Li (UVa), Yamila Miguel (Leiden), Darin Ragozinne (BYU), John Tobin (NRAO), Jonathan Williams (UH), Andrew Youdin (UA)
LOC: Jonathan C. Tan (chair), Thomas Bisbas, Yu Cheng, Antea Cooper, Juan Farias, Madeline Gyllenhoff, Chia-Jung Hsu, Dylan Jones, Boy Lankhaar, Mengyao Liu, Andri Spilker, Liam Walters

Science Program

Key: I - invited talk (22+3min); IS - invited short talk (17+3); C - contributed talk (12+3); CS - contributed short talk (8+2); P - poster (<2)

Sun. 16th June 2019

17.00-20.00 Welcome reception at Vasaparken (find us in middle of park; if weather is bad, go to Vasa kyrkogata 1, south of park, next to Levantine)

Mon. 17th June 2019

8.30			**Registration & poster set-up (Chalmers Conference Center) - posters are displayed until Thursday 2pm**		
8.45	1		**Welcome & Science Overview: Open Questions from the SOC**		

A Large-Scale Interstellar Medium, Molecular Clouds and Star Formation

9.00	2	IS	Kainulainen, Jouni	Chalmers	The PROMISE Survey: the Dark Clouds of the Galactic Plane
9.20	3	C	Wienen, Marion	Exeter	A New Perspective On Star Formation and Spiral Structure in Our Home Galaxy
9.35	4	CS	Cosentino, Giuliana	UCL/ESO	Widespread SiO emission in IRDCs as a probe of cloud-cloud collisions and other shock interactions
9.45	5	CS	Petkova, Maya	Heidelberg	A dominant role for galactic dynamics in turbulence driving, cloud evolution, & star formation in the CMZ
9.55			**Welcome from Chalmers University President, Stefan Bengtsson**		
10.00	6	C	Onishi, Toshikazu	Osaka Pre. U.	High-mass star formation under various environments probed in the Magellanic Clouds
10.15	7	P	Guszejnov, David	Texas	Evolution of giant molecular clouds across cosmic time
	8	P	Gezer, Ilknur	Copernicus AC	Star Formation in the Outer Galaxy
	9	P	Siodmiak, Natasza	Copernicus AC	IRAS 22147+5948 – a cradle of star birth
	10	P	Spilker, Andri	Chalmers	Interpreting Scaling Relations of Star Formation
	11	P	Law, Chi Yan	CUHK/Chalmers	Links between magnetic field and filamentary clouds
	12	P	Tanaka, Kei	Osaka/NAOJ	Synthetic Observations of Molecular Cloud Formation
	13	P	Bisbas, Thomas	Athens/AUT/Köln	How do the observables of molecular clouds depend on the ISM environmental parameters?
	14	P	Gahm, Gösta	Stockholm	Expanding star forming shells around young stellar clusters
	15	P	Rezaei Kh., Sara	Chalmers/MPIA	3D structure of the Orion-Eridanus superbubble with Gaia DR2

10.30 Coffee break & posters

B Exoplanet Properties and Evolution

11.00	16	I	Howard, Andrew	Caltech	Observed Properties of Close-in, Small Planets
11.25	17	I	Ragozzine, Darin	BYU	The underlying distribution of exoplanetary system architectures
11.50	18	C	Janson, Markus	Stockholm	Planet formation around stars of different masses
12.05	19	C	Persson, Carina	Chalmers	Super-Earths: rocky extrasolar planets
12.20	20	CS	Brandeker, Alexis	Stockholm	Is it raining lava in the evening on the super Earth 55 Cancri e?

12.30 Lunch

14.00	21	C	Mustill, Alexander	Lund	How do distant planets affect the dynamics of those we see close to the star?
14.15	22	C	Cai, Maxwell Xu	Leiden	Diverse stellar environments result in diverse exoplanet architectures
14.30	23	C	Veras, Dimitri	Warwick	Autopsies: From Dying Stars to Dying Planets
14.45	24	P	Müller, Simon	Zürich	The Evolution of Giant Planets with Composition Gradients
	25	P	Wallace, Alexander	ANU	Expected Exoplanet Yield from High-Contrast Imaging and Interferometry Surveys
	26	P	Georgieva, Iskra	Chalmers	Methods for Exoplanet Discovery
	27	P	Fridlund, Malcolm	Chalmers	The KESPRINT collaboration
	28	P	McIntyre, Sarah	ANU	Planetary Magnetism as a Parameter in Exoplanet Habitability
	29	P	Li, Daohai	Lund	Flyby encounters between two solar systems
	30	P	Tan, Jonathan	Chalmers/UVA	A Swedish Contribution to ARIEL

C Star Cluster Formation and Evolution

15.00	31	I	Offner, Stella	Texas	The Role of Stellar Feedback in Star Cluster Formation

15.30 Coffee break & posters

16.00	32	C	Inutsuka, Shu-ichiro	Nagoya	Filamentary Molecular Clouds and the Angular Momenta of Cloud Cores
16.15	33	CS	Kuznetsova, Aleksandra	Michigan	Where do Protostellar Cores Get their Angular Momenta?
16.25	34	CS	Treviño Morales, Sandra	Chalmers	Dynamical signs of a hub-filament system
16.35	35	CS	Suri, Sümeyye	MPIA	Are filaments in Orion PDRs?
16.45	36	CS	Andersen, Morten	Gemini	The formation of massive star clusters
16.55	37	CS	Farias, Juan	Chalmers	The Dynamics of Star Cluster Formation
17.05	38	C	Reiter, Megan	UKATC	Cluster dynamics in the typical birthplaces of stars and planets
17.20	39	CS	van Terwisga, Sierk	Leiden	Disk masses in the Orion Molecular Cloud-2: Distinguishing time and environment
17.30	40	C	Hands, Thomas	Zurich	Planetesimal discs in open clusters: implications for 1I/'Oumuamua, Kuiper belt, Oort cloud & more
17.45	41	P	Pelkonen, Veli-Matti	ICCUB	From CMF to IMF: can we predict the mass of the forming star?
	42	P	Parker, Richard	Sheffield	Enlarging habitable zones around binary stars in hostile environments
	43	P	Yi, Hee-Weon	Kyung Hee Uni.	Physical and chemical properties of cores in different environments; the Orion A, B and λ Orionis clouds
	44	P	Walters, Liam	Virginia	Dense gas tracers in the most massive Infrared Dark Cloud
	45	P	Gyllenhoff, Madeline	Virginia	Star Cluster Formation from Ellipsoidal Clouds
	46	P	Cheng, Yu	Virginia	A cluster of protostellar disks in a massive star formation region
	47	P	Zinnecker, Hans	SOFIA	Star formation studies with SOFIA

6-8pm - Poster reception

Tue. 18th June 2019

D Formation and Evolution of the Solar System

8.30	48	IS	Miguel, Yamila	Leiden	Revealing giant planets interiors with Cassini and Juno missions
8.50	49	C	Haugboelle, Troels	Copenhagen	Probing the protosolar disk using dust filtering of the gas giants in the early Solar System
9.05	50	C	Haghighipour, Nader	Hawaii	Accurate Modeling of Terrestrial Planet Formation and Post-Formation Instability of the Solar System
9.20	51	CS	Lichtenberg, Tim	Oxford	A water budget dichotomy of rocky protoplanets from 26Al-heating
9.30	52	C	Li, Rixin	Arizona	Demographics of Planetesimals Formed by the Streaming Instability
9.45	53	I	Nesvorny, David	SWRI	Trans-Neptunian Binaries Provide Evidence for Planetesimal Formation by the Streaming Instability

10.10 Coffee break

10.30	54	C	Pfalzner, Susanne	FZ Jülich	Outer Solar System possibly Shaped by Stellar Flyby
10.45	55	IS	Drozdovskaya, Maria	Bern	Ingredients for Solar-like Systems: IRAS 16293-2422 B versus 67P/Churyumov-Gerasimenko

E Astrochemistry of Star and Planet Forming Regions and Exoplanets

11.05	56	I	Sakai, Nami	RIKEN	Disk Formation and Evolution toward Protoplanetary Disks

11.30 Leave for excursion to Brännö and the southern archipeligo, stopping first at hotels, if needed.
We will take tram 11 to Saltholmen (travel card provided, aim to be there by 12.45) and then the 13.05 ferry to Brännö.
~13.30 Lunch at Brännö Varv, then option for a ~5 km walk in Galterö nature reserve and/or swimming in the Kattegat
Ferries return from Brännö about once per hour.

19.30 Inaugural Astronomy on Tap Gothenburg event (Haket Pub), featuring talks by Yamila Miguel and Christian Eistrup

Wed. 19th June 2019

		E	**Astrochemistry of Star and Planet Forming Regions and Exoplanets (cont.)**		
9.00	57	C	Padovani, Marco	Arcetri	Physics and chemistry of cosmic rays in star-forming regions
9.15	58	C	Rivilla, Victor	Arcetri	Abundant Z-cyanomethanimine in the interstellar medium: paving the way to the synthesis of adenine
9.30	59	I	van Dishoeck, Ewine	Leiden	Astrochemistry on solar-system scales: from young to mature disks
9.55	60	I	Guzmán, Viviana	PUC	Complex molecules in protoplanetary disks
10.20	61	P	Colzi, Laura	Arcetri	Enhanced nitrogen fractionation at core scales: the high-mass star-forming region IRAS 05358+3543
	62	P	Grassi, Tommaso	LMU	The impact of cosmic rays on chemistry and microphysics in hydrodynamical simulations
	63	P	Punanova, Anna	Ural Federal U.	Methanol vs visual extinction in cold cores: test the model predictions
	64	P	Hsu, Chia-Jung	Chalmers	Simulating deuterium fractionation in massive pre-stellar cores
	65	P	Kulterer, Beatrice Marie	Bern	Modelling of deuterated methanol formation around Class 0 objects
	66	P	Calcutt, Hannah	Chalmers	From low-mass to high-mass: Chemical variability across the Galaxy
	67	P	Vasyunin, Anton	Ural Federal U.	Modeling the complex organic molecules formation in IRAS16293-2422
	68	P	Van Gelder, Martijn	Leiden	Directly linking gas and ice abundances in low-mass protostars
	69	P	Jones, Dylan	Virginia	Chemical Evolution During Massive Star Formation
	70	P	Angarita Arenas, Yenifer	Leeds	Pattern Finding in Samples of Spectroscopic Data
	71	P	Cooper, Antea	Virginia	Quantum Chemical Predictions of Prebiotic Polymer Formation and Purine Synthesis

10.40 Coffee break & posters

11.10	72	I	Walsh, Catherine	Leeds	Probing the composition of the planet-building reservoir in disks: the hunt for gas-phase methanol
11.35	73	CS	Ilee, John	Leeds	The interplay between dynamics and chemistry during planet formation
11.45	74	CS	Krijt, Sebastiaan	Arizona	Coupled physical and chemical evolution of CO in protoplanetary disks
11.55	75	CS	Eistrup, Christian	Virginia	From disk to planet atmosphere: the chemical route
12.05	76	CS	Cridland, Alex	Leiden	Connecting planet formation & astrochemistry: refractory carbon depletion & Hot Jupiter atmospheric C/
12.15	77	C	Rahm, Martin	Chalmers	The Azotosome, a Challenge for Computational Molecular Astrobiology

12.30 Lunch

F Pre-stellar and Protostellar Cores, Outflows, and Young Stellar Objects

14.00	78	I	Yang, Yao-Lun	Texas	The Dynamical and Chemical Evolution from Cores to Disks
14.25	79	CS	Küffmeier, Michael	Heidelberg	Modeling embedded Star Formation
14.35	80	CS	Liu, Mengyao	Virginia	The SOFIA Massive (SOMA) Star Formation Survey - Tests of Massive Star Formation Theories
14.45	81	P	Cao, Yue	Nanjing	Striking resemblance between the mass spectrum of sub-pc clumps and the initial mass function
	82	P	Gieser, Caroline	MPIA	Physical and chemical complexity in high-mass star-forming regions
	83	P	Johnston, Katharine	Leeds	A high-resolution picture of spiral arms and stability within the AFGL 4176 disk
	84	P	Koumpia, Evgenia	Leeds	Spatially resolving the MYSO binaries PDS 27 and PDS 37 with VLTI/PIONIER
	85	P	Olech, Mateusz	Copernicus AC	Studies of high-mass star formation with periodic methanol masers
	86	P	Lankhaar, Boy	Chalmers	Characterizing maser polarization
	87	P	Kang, In	Kyung Hee U.	Probing the inner disk of a massive young stellar object with the high-resolution spectrometer, IGRINS
	88	P	Staff, Jan	U. Virgin Islands	Disk wind feedback from high-mass protostars
	89	P	Tokuda, Kazuki	Osaka PU	ALMA studies of the substructure of protostellar/prestellar cores in Taurus
	90	P	Smullen, Rachel	Arizona	The highly variable time evolution of cores
	91	P	Young, Alison	Exeter	Synthetic molecular line observations of the first hydrostatic core
	92	P	Postel, Andreas	Geneva	Analysis and modeling of FU Ori Objects observed with Herschel and Spitzer
	93	P	Djupvik, Amanda	NOT	Time-variable velocities of knots in protostellar jets
	94	P	Habel, Nolan	Toledo	Surveying Protostars within 500PC with HST
	95	P	Skliarevskii, Aleksandr	SFU	Global evolution of a gravitoviscous protoplanetary disk. The importance of the inner sub-au region
	96	P	Waterfall, Charlotte	Manchester	Modelling the radio and X-ray emission from T-Tauri flares
	97	P	Tychoniec, Łukasz	Leiden	Chemistry and kinematics of molecular jets from the protostars - ALMA perspective
	98	P	Zhang, Yichen	RIKEN	Rotation in the NGC 1333 IRAS 4C Outflow
	99	P	Mardones, Diego	U. Chile / MPE	Multiple Outflow Launching Cavities in HH46/47

15.30 Coffee break & posters

16.00	100	CS	Zhang, Yichen	RIKEN	Massive star formation in 0.03-resolution view of ALMA
16.10	101	CS	Frost, Abigail	Leeds	Unveiling a potential transition disk around a massive young stellar object (MYSO)
16.20	102	CS	Kreplin, Alexander	Exeter	Hydrogen-line emission in YSOs: Resolving the origin of Brγ-line emission with NIR interferometry
16.30	103	CS	Tabone, Benoit	Leiden	Unveiling the accretion-ejection connection in protostars with ALMA
16.40	104	CS	Bjerkeli, Per	Chalmers	Resolving star and planet formation with ALMA
16.50	105	CS	Cruz-Sáenz de Miera, Fernando	Konkoly	L1551 IRS 5: ALMA observations of a binary system with a FUor-like star and a circumbinary ring
17.00	106	IS	Li, Zhi-Yun	Virginia	Formation of rings and gaps in wind-launching non-ideal MHD disks
17.20	107	C	Chabrier, Gilles	CRAL, ENS-Lyon	Exploring new fundamental physical mechanisms in star and planet formation and evolution
17.35	108	C	Jayawardhana, Ray	Cornell	Watching Brown Dwarfs Go ’Round and ’Round: Angular Momentum Evolution of Sub-Stellar Objects

19.00- 22.30 Conference dinner at Wijkanders
22.30- … Afterparty at Yaki-Da (free entry before 23.00)

Thursday 20th June 2019

G Protostellar and Protoplanetary Disks

9.00	109	IS	Tobin, John	NRAO	The Protostellar Disk Population in the Orion Molecular Clouds
9.20	110	IS	Williams, Jonathan	Hawaii	Disk Dust Mass Distributions across Protostellar Evolutionary Classes
9.40	111	I	Testi, Leonardo	ESO	Demographical properties of disks from solar mass stars to the BD regime
10.05	112	CS	van 't Hoff, Merel	Leiden	Unveiling the temperature structure of embedded disks
10.15	113	CS	Mori, Shoji	Tokyo	Inefficient Magnetic Accretion Heating in Protoplanetary Disks
10.25	114	P	Savvidou, Sofia	MPIA	Influence of grain growth on the thermal structures of protoplanetary discs
	115	P	Concha-Ramírez, Francisca	Leiden	External photoevaporation of circumstellar disks constrains the timescale for planet formation
	116	P	Cherkos, Alemayehu	Addis Ababa	Effect of viscosity on propagation of MHD waves in astrophysical plasma
	117	P	Nauman, Farrukh	Chalmers	Machine learning applied to disk dynamos
	118	P	Vlemmings, Wouter	Chalmers	A limit to the magnetic field around TW Hya from ALMA CN polarisation observations
	119	P	Trapman, Leon	Leiden	Gas versus dust sizes of protoplanetary disks: Effects of dust evolution
	120	P	Elbakyan, Vardan	SFU	Spatial distribution and growth of dust in a clumpy gravitoviscous protoplanetary disk
	121	P	Akimkin, Vitaly	IoA, RAS	Early evolution of self-gravitating circumstellar disks with a dust component
	122	P	Price, Daniel	Monash	Protoplanetary discs: The messy leftovers of star formation

10.40 Coffee break

11.10	123	CS	Klarmann, Lucia	MPIA	Characterising dust and gas at the inner rim of protoplanetary disks with GRAVITY
11.20	124	CS	Doppmann, Gregory	Keck Obs.	Diagnosing Protoplanetary Conditions with High Resolution IR Spectroscopy
11.30	125	CS	Panic, Olja	Leeds	Giant planet formation sweet spot - Discs around Herbig Ae stars
11.40	126	CS	Hogerheijde, Michiel	Leiden	Locating cold water vapour in the disk of HD100546
11.50	127	P	Notsu, Shota	Leiden	Water lines & multiple ring & gap structures of the disk around HD 163296 observed by ALMA
	128	P	Ros, Katrin	Lund	Icy pebble growth by nucleation and deposition at protoplanetary disc ice lines
	129	P	Miley, James	Leeds	ALMA's view of the mid-plane in Herbig Discs
	130	P	Booth, Alice	Leeds	The First Detection of 13C17O in a Protoplanetary Disk: A Robust Tracer of Disk Mass
	131	P	Vanon, Riccardo	Sheffield	Disc gravito-turbulence regulated by zonal flow instabilities
	132	P	Kadam, Kundan	Konkoly	Dynamical Gaseous Rings in the Dead Zone of Protoplanetary Disks
	133	P	Tominaga, Ryosuke	Nagoya	Dust diffusion & a new secular instability in disks: a way to create ring structures & planetesimals
	134	P	Hu, Xiao	UNLV	Non-ideal MHD simulation of HL Tau disk: formation of rings
	135	P	Bosman, Arthur	Leiden	Probing planet formation and disk substructures in the inner disk
	136	P	Ziampras, Alexandros	Tübingen	Location and shape of water iceline in protoplanetary disks
	137	P	Musso Barcucci, Arianna	MPIA	Detection of Hα Emission from a Brown Dwarf Companion with SPHERE/ZIMPOL
	138	P	Calcino, Josh	Queensland	Signatures of an eccentric disc: Dust and gas in IRS48
	139	P	Jordan, Lucas	Tübingen	Protoplanetary disks in close binaries
	140	P	Rometsch, Thomas	Tübingen	Forming transition disks with planets
	141	P	Chen, Lei	Konkoly	Variable warm dust around the Herbig Ae star HD 169142: is it of secondary origin?
	142	P	Woelfer, Lisa	MPE	Radiation-Hydrodynamical Models of X-ray Photoevaporation in Carbon Depleted Circumstellar Discs
	143	P	Franz, Raphael	LMU	Dust Entrainment in Photoevaporative XEUV Winds

H Planet Formation

144	P	Lyra, Wladimir	Cal State	Initial conditions for planet formation: Turbulence driven by hydrodynamical instabilities in disks
145	P	Schaefer, Urs	Hamburg	On the Coexistence of Streaming Instability and Vertical Shear Instability in Protoplanetary Disks
146	P	Carpenter, Vincent	MPIA	Simulations of the Onset of Collective Motion of Sedimenting Particles
147	P	Eriksson, Linn	Lund	Pebble drift and planetesimal formation in protoplanetary disks with embedded planets
148	P	Schulik, Matthäus	Lund	Global 3D radiation-hydro simulations of gas accretion: opacity dependent growth of Saturn-mass plane
149	P	Moldenhauer, Tobias	Tübingen	Planetary Proto-Atmospheres
150	P	Hasegawa, Yasuhiro	JPL/Caltech	The occurrence rate of close-in giant exoplanets as a link between formation of stars and planets
151	P	Tan, Jonathan	Chalmers/UVA	Tests of Inside-Out Planet Formation
152	P	Penzlin, Anna	Tübingen	Dynamics of circumbinary discs and planet migration
153	[P]	Nordlund, Åke	Copenhagen	Task based computing with the DISPATCH code framework

12.30 Lunch

14.00	154	I	Johansen, Anders	Lund	From dust to planetesimals and beyond
14.25	155	CS	Liu, Beibei	Lund	Growth after the streaming instability: From planetesimal accretion to pebble accretion
14.35	156	CS	Ramsey, Jon	Virginia	Simulating pebble accretion with DISPATCH, a next-generation tool for computational astrophysics
14.45	157	I	Nelson, Richard	Queen Mary	Planet migration in laminar and inviscid protoplanetary discs
15.10	158	I	Zhu, Zhaohuan	UNLV	Using Disk Substructures to Reveal Young Planet Population
15.35	159	CS	Nayakshin, Sergei	Leicester	ALMA dusty annular structures: significant implications for planet and star formation theories

15.45 Coffee break

16.10	160	I	Mulders, Gijs	U. Chicago	Exoplanets around stars of different masses as empirical constraints for planet formation models
16.35	161	I	Mordasini, Christoph	Bern	A new generation of planetary population syntheses
17.00	162	CS	Bitsch, Bertram	MPIA	Formation and composition of super-Earths
17.10	163	CS	Hu, Xiao	UNLV	Accreting Protoplanetary Disks and In Situ Inside-Out Planet Formation
17.20	164	CS	Jankovic, Marija	Imperial	In situ formation of close-in super-Earths due to the MRI
17.30	165	CS	Marleau, Gabriel-Dominiqu	Tübingen	The accretion shock in planet and star formation
17.40	166		Conference Summary / Discussion		

18.30 End of Science Sessions

Friday 21st June 2019

Midsummer activities and farewell party

11.00 Bus leaves Elite Hotel to go to Onsala Space Observatory for a picnic lunch and hike by the sea and telescopes (bring your own food)

~16.00 Bus returns to Gothenburg first to Park Slottsskogen (drop off at Natural History Museum), then continuing back to Elite Hotel.

~16.00 to ~20.00 Enjoy the outdoor party and BBQ at Park Slottskogen (also close to Botanical Gardens)

~21.00 Farewell party near Vasa Park

Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc

Marija Jankovic – Imperial College London

At the orbital distances of the abundant close-in super-Earths, the inner regions of protoplanetary discs are viscously accreting due to the magneto-rotational instability (MRI). The MRI-induced turbulence limits the size of dust particles in the inner disc, as the particles fragment due to relative turbulent velocities. We show that this leads to enhancement of the dust-to-gas ratio, i.e., to the first stage of planet formation in the inner disc. We further explore how the high temperatures and low gas surface densities of the MRI-accreting inner disc affect the last stages of planet formation, the accretion of planet atmospheres onto solid cores. We calculate how much gas solid cores accrete from the disc, how much of their atmospheres they subsequently lose to photoevaporation, and the resulting planet mass-radius relationship. Finally, we compare the results to the observed exoplanet masses and radii.

Tracing volatile elements from disks to planets

Oliver Shorttle – University of Cambridge – *Invited speaker*

The elements carbon, water, and sulfur, are some of the most abundant in the universe, are major constituents of planetary atmospheres, and are essential constituents of all known life. Yet, their arrival on planets and distribution therein remains uncertain. In large part this is due to their volatile nature: each element may form species that readily enter the gas phase at the temperatures and pressures of planetary differentiation. Here, we combine observations from Earth with observations from protoplanetary disks. Combining geochemical and astrophysical observations in this way gives us new insight into how these elements were partitioned between gas and dust during accretion, and where they are ultimately distributed in planets.

The innermost region of protoplanetary discs and its connection to the origin of planets.

15–19 OCTOBER 2018

ESO HQ, GARCHING BEI MÜNCHEN, GERMANY

TAKE A
CLOSER
LOOK

Program & Abstracts

Take a close look

The quest for detecting exoplanets (e.g., via Kepler and HARPS RV surveys) has revealed the existence of a large population of systems comprising one to several planets very close to the central star, i.e. at distances of 0.1-1 au, even around TTauri (age<5 Myr) stars. These are usually slightly bigger than the Earth and up to Neptune sizes, with rare Jupiter analogues. This finding differs to what we observe in our own Solar System, and raises the question of how such planets form. From a theoretical point of view, it is still hard to show that these planets formed in-situ, but it is similarly complicated to explain this large population of close-in planets as a result of migration through the disk. Additional evidence of the importance of this region comes from our own Solar System, where studies have established that material routinely observed in meteorites (e.g., Ca-Al-rich inclusions, CAI) must have formed very close to the central star, or in a very hot region of the disc.

To advance our understanding of planet formation and migration, it is crucial to study the conditions within the inner regions of their progenitor protoplanetary discs. The innermost part of the disc is where most of the star-disc interaction processes take place. The magnetic field topology of the central star truncates the disc at a few stellar radii and drives accretion of material onto the central star, as well as the ejection of fast-collimated jets and slow winds. Recent studies indicate that this star-disc interaction evolves quickly at the same time that giant planet formation ceases. Also, this region is known to undergo rapid evolution, for example, short or long lasting dimming events (e.g., AA-Tau, RW Aur, dippers). This rapid evolution is, in itself, likely to impact the formation of planets. Finally, a fraction of discs known as transition discs, show a deficit of dust in the inner few au of the disc, which could be related to the mechanism driving disc evolution in this planet-forming region.

Studies of this key inner disc region require innovative techniques and a wide range of instrumentations, since radio interferometers cannot resolve spatial scales smaller than 10 au in most discs. Observations with instruments on the ESO/VLT and VLTI and other facilities provide us with unprecedented detail and motivate this workshop. Specifically, this workshop aims at discussing the present-day knowledge of the morphology and composition of the innermost regions of the disc, of the star-disc interaction processes, and of the theories to describe the evolution of the innermost regions of discs and of the formation of close-in planets.

SOC

Hussain, G. (ESO) (chair)
Benisty, M. (IPAG Grenoble)
Birnstiel, T. (LMU Munich)
Bouvier, J. (IPAG Grenoble)
Calvet, N. (U. Michigan)
Ciesla, F. (U. Chicago)
Ercolano, B. (LMU Munich)
Lopez, B. (OCA Nice)
Natta, A. (DIAS Dublin)
Manara, C.F. (ESO) (organizer)
Schneider, P.C. (Hamburg) (organizer)

LOC

Stella Chasiotis-Klingner
Carlo F. Manara
Gaitee Hussain
Paolo Cazzoletti
Giuliana Cosentino
Stefano Facchini
Anna Miotello
Tereza Jeăábková
Enrique Sanchis Melchor
Tomáš Tax
Maria Giulia Ubeira Gabellini

Workshop venue

The workshop will be held in the ESO Auditorium, located at:

Karl-Schwarzschild-Str. 2, Garching b. München.

Contact phone and fax number

Contact phone number: +49 89 320 060 (ESO Reception Desk)
Contact fax numbers: +49 89 320 23 62

Registration

Registration will take place on Monday, October 15th in front of the new auditorium (Eridanus) from 12:30.

Contact

Please use the following email-address to contact us: `tcl2018@eso.org`

Webpage

`https://www.eso.org/sci/meetings/2018/tcl2018.html`

What's in the booklet

1. Program Overview p. 4
2. Detailed Program p. 5
3. List of Posters p. 10
4. Talk abstracts p. 15
5. Poster abstracts p. 44

Program Overview

Monday, Oct 15

12:30	Registration
13:30	Welcome
13:50	**Session 1 – Close-in exo-planets observations and origin**
18:00	Reception at Supernova
18:15	Planetarium show & visit of exhibition
& 18:45	(two groups of ∼70 people)
19:15	Food served in the Supernova building

Tuesday, Oct 16

09:00	**Session 2 – Morphology of the inner disk**
12:30	Lunch
16:50	Poster viewing with Beer and Brezn

Wednesday, Oct 17

09:00	Session 2 continues
09:40	**Session 3 – Evolution of protoplanetary disks**
12:30	Lunch
17:00	Poster viewing with Beer and Brezn

Thursday, Oct 18

09:00	**Session 4 – Angular momentum transport and connection of outer and inner disk**
12:30	Lunch
14:00	**Session 5 – Star-disk interaction**
16:30	Breakout Sessions
	AUDITORIUM ERIDANUS RW Aur: origin of the dimming events
19:00	Social dinner: Gasthof Neuwirt in Garching

Friday, Oct 19

09:00	Session 5 continues
11:00	**Final Discussion and Concluding Remarks**
12:30	End of program

Monday, Oct 15

12:30	**Registration**		
13:30	**Introduction**		

Session 1 – Close-in exo-planets observations and origin

13:50	Haywood, Raphaëlle	Invited review	Current status of the observational search for close-in planets
14:30	Ormel, Chris	Invited review	Theories of formation of close-in exoplanets
15:10	Kreplin, Alexander		VLTI+ALMA imaging of potential planet-formation processes in the pre-transitional disc of V1247 Ori
15:30	Mendigutía, Ignacio		Detecting signatures of planet formation in disks around young stars
15:50	**Break**		
16:20	Keppler, Miriam		Discovery of a planetary mass companion in the gap of the transition disk PDS 70
16:40	Jankovic, Marija		Formation of close-in super-Earths: dust enrichment of the inner disk due to the MRI
17:00	Ogihara, Masahiro		Formation of close-in super-Earths from embryos with suppressed type I migration
17:20	Drazkowska, Joanna		Planetesimal formation in the inner region of the protoplanetary disk
17:40	Tan, Jonathan		Inside-Out Planet Formation

18:00	**Reception at Supernova**
18:15 & 18:45	Planetarium show and visit of the exhibition (two groups of ∼70 people)
19:15	**Food served in the Supernova building**

Tuesday, Oct 15

Session 2 – Morphology of the inner disk

09:00	Facchini, Stefano	Invited review	The inner regions of protoplanetary disks: a dynamical perspective
09:40	Kraus, Stefan	Invited review	Spatially resolved observations of inner disc structure
10:20	Banzatti, Andrea	Invited talk	Observing the evolution of exoplanet-forming disks at 0.01-10 au with gas, dust, and wind tracers
10:50	**Break**		
11:30	Ansdell, Megan		Misaligned inner disk demographics using "dipper" stars
11:50	Pinilla, Paola		Variable Outer Disk Shadowing Around the Dipper Star J160421.7-213028
12:10	Nealon, Rebecca		Warping protoplanetary discs by planets on inclined orbits
12:30	**Lunch**		
14:00	Brown, Alexander		Dust and Gas Composition at the Inner Edge of the Transitional Disk Surrounding 7Myr old T Cha
14:20	Arulanantham, Nicole		Mapping the Inner Disk Gas around Young Stars in the Lupus Complex
14:40	France, Kevin	Invited talk	UV Characterization of Inner Disks: from Hubble to LUVOIR
15:10	Perraut, Karine		The inner dust rim of protoplanetary disks as probed by GRAVITY
15:30	**Break**		
16:00	Matter, Alexis	Invited talk	Observing and characterizing the planet-forming region of protoplanetary disks with MATISSE
16:30	Davies, Claire		Grain growth-induced curvature of the inner rim of protoplanetary disks
16:50	**Beer and Brezn**		**Poster Viewing**
19:00	**End of day**		

Wednesday, Oct 17

09:00	Garufi, Antonio		The outer disk knows about the inner disk
09:20	Ricci, Luca		Imaging the inner regions of protoplanetary disks with sub-millimeter and millimeter interferometry

Session 3 – Evolution of protoplanetary disks

09:40	Rosotti, Giovanni	Invited review	Disc evolution processes: how they affect the inner disc
10:20	McClure, Melissa	Invited review	Eye to the telescope: observing the inner 1 AU of protoplanetary disks
11:00	**Break**		
11:30	Parker, Raeesa		Understanding the formation and distribution of cm-sized dust grains in DG Tau A
11:50	Bitsch, Bertram		Origin of super-Earths planets: influence of pebble accretion, migration and instabilities
12:10	McNally, Colin		How planets move in the inner disc
12:30	**Lunch**		
14:00	Bosman, Arthur		Inner disk chemistry and the effects of drifting icy grains: the case of CO2
14:20	Booth, Richard		The inner disc tracers of the physics and chemistry of disc evolution
14:40	Chen, Christine	Invited talk	Observing the Inner Regions of Protoplanetary Disks with JWST
15:10	Simon, Jake		What Drives Accretion in Protoplanetary Disks?
15:30	**Break**		
16:00	Ballabio, Giulia		Empirical diagnostics of protoplanetary disc winds
16:20	Picogna, Giovanni		The dispersal of planet-forming discs. A new generation of X-ray photoevaporation models.
16:40	Monsch, Kristina		The imprint of X-ray photoevaporation on the orbital distribution of giant planets
17:00	**Beer and Brezn**		**Poster viewing**
19:00	**End of the day**		

Thursday, Oct 18

Session 4 – Angular momentum transport and connection of outer and inner disk

09:00	Sheikhnezami, Somayeh	Invited review	Disk-outflow interaction: exchange of mass and angular momentum
09:40	García López, Rebeca	Invited review	Observations of outflows and angular momentum transport from the inner to the outer disk
10:20	Dougados, Catherine		What can jets and outflows tell us about the central astronomical units ?
10:40	**Break**		
11:10	McGinnis, Pauline		Spectral signatures of jets and winds in the young, open cluster NGC 2264
11:30	Nisini, Brunella		Connection between jets, winds and accretion in T Tauri stars
11:50	Fischer, Will		Complementary Insights from Helium 10830 in the Age of Interferometry
12:10	Nolan, Chris		Determining the launching regions of centrifugally-driven disc winds from non-ideal MHD simulations
12:30	**Lunch**		

Session 5 – Star-disk interaction

14:00	Gregory, Scott	Invited review	The impact of magnetic fields on the star-disk interaction and planet formation
14:40	Venuti, Laura	Invited review	Observations of star-disk interaction and link to disk evolution at the epoch of planet formation
15:20	Takasao, Shinsuke		Fast accretion onto a weakly magnetized star
15:40	Alencar, Silvia		The inner disk structure of the classical T Tauri star LkCa 15
16:00	**Break**		
16:30	**Breakout Sessions**		
	AUDITORIUM ERIDANUS – RW Aur: origin of the dimming events		
	PAVO, TUCANA, DORADO, VELA – Rooms to be requested by participants for oher sessions		
18:00	**End of the day**		
19:00	Social Dinner		Gasthof Neuwirt Garching

Friday, Oct 19

09:00	Hill, Colin		Magnetic Fields and Planets of Weak-line T Tauri Stars
09:20	Beccari, Giacomo		Accretion properties of low-mass stars in the Large Magellanic Cloud: the case of LH95
09:40	Aly, Hossam		Dusty warps: Can Warps in Protoplanetary Discs Form Dust Traps?
10:00	Ramsay, Suzanne	Invited talk	Looking closely at protoplanetary discs with the European ELT and its instruments
10:30	**Break**		
11:00	**Final Discussion and Concluding Remarks**		
12:30	**End of Program**		

Breakout session 1 – RW Aur: Origin of the dimming events

16:30	Facchini, Stefano	Brief overview on the RW Aur dimming events
16:40	Lamzin, Sergei	Analysis of photometric and polarimetric observations of RW Aur A: arguments in favor of dusty wind
16:45	Safonov, Boris	Resolving dusty wind of RW Aur A
16:50	Koutoulaki, Maria	Shedding light on the dimming events of the CTTS RW Aur A.
16:55	Günther, H. Moritz	X-ray news from RW Auriga: Optical dimming with iron rich plasma and an exceptional column density
17:00	Gárate, Matias	The dimmings of RW Aur. Is the accretion of dust preceding an outburst?
17:05	Discussion on RW Aur lead by S. Facchini	

014-9658-4) was fundamentally flawed by the authors' misunderstanding of chemical equilibria (Jackson JB, 2017, J Mol Biol, doi: 10.1007/s00239-017-9805-9).

We should not be seduced by the apparent simplicity of the natural pH gradient hypothesis into thinking that the origin of other forms in the universe (on exoplanets) might have easily and commonly involved natural pH gradients. Modern molecular machines capable of utilising Δ pH in living organisms on earth probably only emerged after the evolutionary arrival of proteins- that is, rather late in the development of life.

Marija Jankovic

Imperial College London

A057

Dust enhancement in the inner disk due to the MRI

The formation pathway of the abundant close-in super-Earths and mini-Neptunes is still unclear. Do they form *in situ* or migrate inwards as embryos? If the former, then the key step to planet formation is dust enrichment of the inner protoplanetary disk. Dust is subject to growth by coagulation, and radial drift and fragmentation due to gas drag and turbulence. In the outer disk the dust particles are expected to quickly grow and drift inwards. The inner disk structure, governed by viscous accretion due to the magneto-rotational instability (MRI), features a local gas pressure maximum that could potentially trap the drifting dust particles.

We examine the evolution of dust radial distribution and dust particle size in a steady-state gaseous disk, with the inner disk gas structure self-consistently determined from MRI criteria (Mohanty et al. 2017).

We find that in the inner disk turbulent fragmentation limits the size of dust particles, which thus remain too small to feel significant gas drag and therefore do *not* accumulate at the local gas pressure maximum. However, as a decrease in particle size also implies less radial drift, this ultimately leads to *dust enhancement throughout the inner disk*, interior to the pressure maximum.

Luke Jonathan Johnson, Charlotte M. Norris*

Imperial College London

A054

Simulations of Stellar Variability

The radiative output of stars is not constant, with brightness fluctuations observed as they rotate due to magnetic surface features such as cool starspots and hot active regions. To facilitate analysis of this phenomenon, we present ACTReSS, a software tool for calculating the incident flux from a model active stellar surface as it varies throughout a rotation. The model uses a non-linear limb-darkening law with coefficients derived from MURaM 3D magneto-convection simulations for the quiet Sun and bright active regions. This allows us to investigate the dependence of the flux variation on spectral type and wavelength.

*in collaboration with Yvonne C. Unruh (Imperial College); Sami K. Solanki, Natalie A. Krivova (MPS Göttingen)