

Detecting few-body quantum chaos: out-of-time ordered correlators at saturation

Dragan Marković^{a,b} and Mihailo Čubrović^b

^a*Department of Physics, University of Belgrade, Studentski Trg 12-16, 11000 Belgrade, Serbia*

^b*Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia*

E-mail: vokramnagard@gmail.com, cubrovic@ipb.ac.rs

ABSTRACT: We study numerically and analytically the time dependence and saturation of out-of-time ordered correlators (OTOC) in chaotic few-body quantum-mechanical systems: quantum Henon-Heiles system (weakly chaotic), BMN matrix quantum mechanics (strongly chaotic) and Gaussian random matrix ensembles. The growth pattern of quantum-mechanical OTOC is complex and nonuniversal, with no clear exponential regime at relevant timescales in any of the examples studied (which is not in contradiction to the exponential growth found in the literature for many-body systems, i.e. fields). On the other hand, the plateau (saturated) value of OTOC reached at long times decreases with temperature in a simple and universal way: $\exp(\text{const.}/T^2)$ for strong chaos (including random matrices) and $\exp(\text{const.}/T)$ for weak chaos. For small matrices and sufficiently complex operators, there is also another, high-temperature regime where the saturated OTOC grows with temperature. Therefore, the plateau OTOC value is a meaningful indicator of few-body quantum chaos. We also discuss some general consequences of our findings for the AdS/CFT duality.

KEYWORDS: AdS-CFT Correspondence, Field Theories in Lower Dimensions, Random Systems, Matrix Models

ARXIV EPRINT: [2202.09443](https://arxiv.org/abs/2202.09443)

Contents

1	Introduction	1
2	OTOC in quantum-mechanical systems	3
2.1	An upper bound on OTOC saturation	4
3	OTOC for random matrix ensembles	5
3.1	Estimate of the OTOC and its plateau	6
3.1.1	Kinematic OTOC	7
3.1.2	OTOC for dense and/or random operators	9
3.2	Numerical checks	9
4	OTOC for weakly and strongly chaotic Hamiltonians	11
4.1	Weak chaos: perturbation theory	12
4.2	Weak chaos: examples and numerics	15
4.3	Strong chaos: numerics and the return to random matrices	18
5	Discussion and conclusions	19
A	Detailed structure and calculation of OTOC for Gaussian orthogonal ensembles	21
A.1	The large matrix limit	24

1 Introduction

Recent years have seen a renewed interest in classical and quantum chaos in the context of high-energy physics, black holes and AdS/CFT, thanks to the relation of chaos to quantum information theory and the information problems of black holes. Sharp and reasonably rigorous results such as the celebrated MSS chaos bound [1] and its subsequent refinements [2, 3] establish a connection between chaos and the fundamental properties of gravity and black holes [4, 5]. Maximal chaos, with the Lyapunov exponent $\lambda = 2\pi T$ at temperature T , is reached for strongly coupled field theories in the large N limit, which have a classical gravity dual with a black hole. In [2] and other works it is explicitly shown how the Lyapunov exponent changes with finite N effects.

However, it has been pointed out many times, also in the pioneering MSS paper [1], that the multiple notions of quantum chaos in the literature mean different things. The out-of-time ordered correlation function (OTOC), given by the expectation value of the commutator of some operators A and B at times 0 and t :

$$C(t) \equiv \langle |[A(t), B(0)]|^2 \rangle, \quad (1.1)$$

is a natural quantity in quantum field theories, i.e. many-body systems, and defines the quantum Lyapunov exponent λ as the exponent of the time growth of OTOC. However, the classical limit of this exponent does not necessarily have much to do with the classical Lyapunov exponent λ_{class} , obtained by solving the variational equations [6–8]. The reason is the noncommutation of the three limits to be taken: the classical limit $\hbar \rightarrow 0$, the long-time limit $t \rightarrow \infty$, and the small initial variation limit $\delta x(0) \rightarrow 0$. The crucial insight of [7, 8] is that the mechanism of scrambling may be the chaotic dynamics, in which case λ_{class} is related though still not identical to the OTOC exponent (quantum Lyapunov exponent) λ , or it may originate in local instability (hyperbolicity), in which case even regular systems may have a nonzero λ exponent and likewise chaotic systems may have λ which is completely unrelated to the classical counterpart.

This mismatch between the classical and quantum Lyapunov exponent is just the tip of the iceberg. The problem is twofold: not only what is the relation between the quantum (OTOC) exponent and classical chaos, but also what is the relation between the quantum Lyapunov exponent λ and other indicators of quantum chaos such as, first and foremost, level statistics. The bread and butter of quantum chaos is the famous Dyson threefold way leading to the Wigner surmise, the level repulsion statistics determined solely by the time reversal properties of the Hamiltonian [10], which follows from the random matrix approximation of chaotic Hamiltonian operators [11]. It is no secret for several years already that the black hole quasinormal mode spectra follow the random matrix statistics [12], and the OTOC of a Gaussian unitary ensemble (GUE) has been computed analytically in terms of Bessel functions in [13, 14]; the outcome is close to the expected behavior of large- N field theories only at long timescales, longer than the scrambling time; at shorter timescales there are important differences. The authors of [13] have reached a deep conclusion in this respect: random matrices have no notion of locality as the correlation of any pair of eigenvalues is described by the same universal function. This is why the OTOC of a GUE system deviates from that of a local field theory at early times, when the perturbation in field theory has not had time to spread yet (i.e. when it is still localized). Therefore, the level repulsion does not imply the usual picture of the chaotic (exponential) OTOC behavior. However, we do not know yet how this correlates to the behavior of *few-body* or more precisely few-degrees-of-freedom quantum systems as opposed to the large- N field theories with a gravity dual. In few-body systems the notion of locality (and a classical gravity dual) does not exist anyway and the main problem found by [13] is irrelevant; at the same time, such systems are often very well described by random matrix statistics, i.e. Wigner-Dyson statistics [10]. In this paper we aim to understand the behavior of OTOC in such systems. Running a bit forward, we can say that the *growth* of OTOC is rather unremarkable: we find no universal trend, and little connection to level statistics. This confirms the results found for specific examples in [8, 15].

The relation of OTOC, level statistics and the classical Lyapunov exponent was studied for few-body systems (quantum mechanics) in [8, 15–18] and the picture is inconclusive. One can have a nonzero growth exponent in integrable systems,¹ whereas fast scrambling with

¹This actually correlates with the classical variational equations in hyperbolic systems, which show exponential growth even in absence of chaos.

the exponent close to $2\pi T$ or at least growing linearly in T has not been found even in some clearly chaotic systems [8, 9]. Arguments for many-body systems such as spin chains even suggest that quantum-chaotic systems with Gaussian spectral statistics generically never show fast scrambling [19], but no claims of such generality have been tested or formulated for few-body quantum chaos.

Various indicators of chaos relevant also for small systems, and their relation to OTOC and scrambling were studied by [20–27] among others. In particular, in [26] some important insights can be found: even in small systems devoid of the locality notion, OTOC can be interpreted as a measure of delocalization of a state in phase space, and the oscillatory component of the OTOC dynamics has to do with the power spectrum of the system. This last insight provokes a more general question: can we learn something from the quasi-stationary regime of OTOC, where no systematic growth is present but only oscillations? In this paper we provide a partial answer from a detailed study of this saturated (asymptotic, plateau) OTOC regime: the magnitude of the OTOC average at the plateau has a simple temperature dependence, and apparently can differentiate between weak chaos (dominantly Poissonian level distribution with some admixture of the Wigner-Dyson statistics) and strong chaos (clear Wigner-Dyson level repulsion). We will demonstrate this on three representative systems: the quantum Henon-Heiles Hamiltonian, whose classical limit has mixed (regular/chaotic) phase space and thus we expect on average weak chaos, a simplified BMN matrix model (at small N) exhibiting strong chaos for most initial conditions, and Gaussian random matrices, the prototype of strong quantum chaos. The long-time limit of OTOC behaves in subtly different ways in each case.

Before we start, one caveat is in order (we will consider this issue in more detail later on): one might think that the saturated OTOC value is always trivially determined by the system size. We typically assume that the OTOC function C as defined in (1.1) behaves roughly as $C(t) \sim c/N^2 \times \exp(\lambda t)$ with c of order unity, so when $t \sim t_* \equiv \log N^2/\lambda$ the growth of $C(t)$ stops and OTOC approximately reaches unity (when appropriately normalized). But the twist is precisely that c is system-specific and in general poorly known. The leading N^2 behavior indeed determines the OTOC values for N large, but when N and c are comparable within an order of magnitude the effects of fluctuations and finite N corrections are significant. This is at the root of our observations in this work.

The plan of the paper is as follows. In section 2 we recapitulate and generalize some results on computing OTOC in quantum mechanics, and show how OTOC sensitively depends on both the Hamiltonian and the operators A, B from the definition (1.1). In section 3 we apply the general formalism to random matrix ensembles and show that the OTOC growth is a complicated and nonuniversal function but that its asymptotic value behaves in a rather simple way. Section 4 discusses the behavior of OTOC for deterministic quantum-chaotic Hamiltonians. Section 5 sums up the conclusions.

2 OTOC in quantum-mechanical systems

Consider a four-point time-disordered correlation function for a quantum-mechanical system in $0 + 1$ dimensions at temperature $T = 1/\beta$. Starting from the usual definition (1.1) as the

squared module of the commutator of the two operators A and B , we can write it out as

$$C(t) = \frac{1}{Z} \langle |[A(t), B(0)]|^2 \rangle = \frac{1}{Z} \sum_n e^{-\beta E_n} \langle n | [A(t), B(0)]|^2 | n \rangle, \quad (2.1)$$

where the averaging is both thermal and quantum mechanical: $\langle \dots \rangle = \text{tr} e^{-\beta H} \langle \text{vac} | \dots | \text{vac} \rangle$. We can pick a basis of states and express the above defining expression in terms of matrix elements of the operators (this closely follows the derivation in [8, 16]):

$$C(t) = \frac{1}{Z} \sum_{nm} e^{-\beta E_n} \langle n | [A(t), B(0)] | m \rangle \langle m | [A(t), B(0)] | n \rangle = \frac{1}{Z} \sum_{nm} e^{-\beta E_n} |c_{mn}(t)|^2, \quad (2.2)$$

where we have inserted the completeness relation $1 = \sum_m |m\rangle \langle m|$. For a single element $c_{mn}(t)$ one gets:

$$\begin{aligned} c_{mn}(t) &= \langle n | [e^{iHt} A e^{-iHt}, B] | m \rangle = \\ &= \sum_k \left(\langle n | e^{iHt} A e^{-iHt} | k \rangle \langle k | B | m \rangle - \langle n | B | k \rangle \langle k | e^{iHt} A e^{-iHt} | m \rangle \right) = \\ &= \sum_k \left(\langle n | e^{iE_n t} A e^{-iE_k t} | k \rangle \langle k | B | m \rangle - \langle n | B | k \rangle \langle k | e^{iE_k t} A e^{-iE_m t} | m \rangle \right) = \\ &= \sum_k \left(a_{nk} b_{km} e^{-iE_{kn} t} - b_{nk} a_{km} e^{-iE_{mk} t} \right), \end{aligned} \quad (2.3)$$

where in the second line we have again inserted a completeness relation and in the third line we have used the fact that we work in the energy eigenbasis. The outcome is expressed in terms of the matrix elements a_{mn}, b_{mn} of the operators in the energy basis. In practice, it may or may not be possible to compute these analytically. Specifically, for $A = x, B = p$, we get the analogue of the classical Lyapunov exponent. From now on we call this the kinematic OTOC as it is directly related to the classical trajectory. Let us now see what general bounds can be put on (2.3) from the properties of quantum-mechanical Hamiltonians.

2.1 An upper bound on OTOC saturation

We begin with a very general and very formal result, which immediately makes it clear that in a generic quantum-mechanical system (integrable or nonintegrable) OTOC can be bounded from above by a quantity which solely depends on the energy spectrum of the Hamiltonian and the choice of the operators A and B . This upper bound remains valid no matter what is the time dependence of OTOC, even if it does not have a nonzero growth exponent at all (which is quite generic in quantum mechanics). Starting from the basic equations (2.2)–(2.3), let us denote $C_{nmk} = a_{nk} b_{km}$ and $D_{nmk} = -b_{nk} a_{km}$, and estimate a single coefficient $c_{mn}(t)$ in the sum. We clearly have

$$\begin{aligned} |c_{mn}(t)| &= \left| \sum_k C_{nmk} e^{-iE_{kn} t} + D_{nmk} e^{-iE_{mk} t} \right| \leq \sum_k |C_{nmk} e^{-iE_{kn} t} + D_{nmk} e^{-iE_{mk} t}| \Rightarrow \\ |c_{mn}(t)|^2 &\leq \left(\sum_k |C_{nmk} e^{-iE_{kn} t} + D_{nmk} e^{-iE_{mk} t}| \right)^2 \leq \sum_k |C_{nmk} e^{-iE_{kn} t} + D_{nmk} e^{-iE_{mk} t}|^2 \leq \\ &\leq \sum_k \left(|C_{nmk}|^2 + |D_{nmk}|^2 + 2|C_{nmk}| |D_{nmk}| \cos(E_{mk} - E_{kn}) t \right), \end{aligned} \quad (2.4)$$

where N is the matrix size. In the second and third line we have used the inequality between the arithmetic and harmonic mean. Now we can bound the value of $C(t)$:

$$0 \leq C(t) \leq \frac{1}{Z} \sum_{nmk} e^{-\beta E_n} \left(|C_{nmk}|^2 + |D_{nmk}|^2 + 2|C_{nmk}||D_{nmk}| \cos(E_{mk} - E_{kn})t \right) \quad (2.5)$$

This means that $C(t)$ is bounded at all times by an oscillatory function of time, whose frequencies are linear combinations of three eigenenergies ($E_{mk} - E_{kn} = E_m + E_n - 2E_k$). Such a combination is generically always nonzero for a chaotic system except when the energies coincide, e.g. $E_m = E_n = E_k$ (according to the non-resonance condition). Therefore, since OTOC is typically a non-decreasing function of time, the behavior of $C(t)$ for t large is roughly its maximum value and is likely close to the right-hand side in (2.5). This suggests that the OTOC dynamics after saturation likely consists of a very complex oscillatory pattern (with $\sim N^3$ frequencies if the Hilbert space has dimension N) superimposed on a plateau. The numerics will indeed confirm such behavior.

Another estimate, which is time-independent and relevant for our main result — the magnitude of the saturation (plateau) OTOC value, is obtained from the triangle and mean inequalities:

$$\begin{aligned} |c_{mn}(t)|^2 &\leq \left| \sum_k C_{nmk} e^{-iE_{kn}t} \right|^2 + \left| \sum_k D_{nmk} e^{-iE_{mk}t} \right|^2 \leq \sum_k |C_{nmk}|^2 + \sum_k |D_{nmk}|^2 \Rightarrow \\ C(t) &\leq \frac{1}{Z} \sum_{nm} e^{-\beta E_n} \left(|(A \cdot B)_{nm}|^2 + |(B \cdot A)_{nm}|^2 \right) \leq \frac{2}{Z} \sum_{nm} e^{-\beta E_n} |(A \cdot B)_{nm}|^2, \end{aligned} \quad (2.6)$$

where we have used the obvious relations $\sum_k C_{nmk} = (A \cdot B)_{nm}$ and $\sum_k D_{nmk} = (B \cdot A)_{nm} = (A \cdot B)_{mn}^* = (A \cdot B)_{nm}$, assuming also the hermiticity of the operators. For some models (e.g. random matrices, Henon-Heiles), this sum can be estimated in a controlled way and provides an approximation for the plateau of OTOC. These estimates are obviously very simple and very weak (in the mathematical sense) but provide us with a framework into which we can insert specific A , B and H (the Hamiltonian with energies E_n) and perform back-of-the-envelope calculations which explain the numerical findings.

3 OTOC for random matrix ensembles

Random matrix theory [10, 11] provides a highly detailed and rigorous (within its starting assumptions) stochastic effective description of the few-body quantum chaos, and allows an analytic calculation of OTOC along the lines of (2.3). Let us focus on Gaussian orthogonal ensembles of size $N \times N$, appropriate when there is full time reversal invariance. It is known [10] that the joint distribution all the elements of all eigenvectors is obtained simply from the statistical independence of the eigenvectors from each other and of the elements in each eigenvector (and the orthogonality of the eigenvectors):

$$P(\{c\}) = \left(\prod_{n=1}^N \delta \left(\sum_i (c_i^n)^2 - 1 \right) \right) \left(\prod_{n < m} \delta \left(\sum_i c_i^n c_i^m \right) \right), \quad (3.1)$$

where $i = 1 \dots N$ is the component of the eigenvector and $1 \leq n, m \leq N$ count the eigenvectors themselves, i.e. the energy levels; so the n -th eigenvector $|n\rangle$ is represented by the column vector $\psi^{(n)}$ with the elements (c_1^n, \dots, c_N^n) . Special cases like the probability distribution for the p -tuple of elements of a single eigenvector are obtained from (3.1) by integrating out all the other elements [10]. We will also need the probability distribution of the energy levels $\{E\} = E_1, E_2, \dots, E_N$, the celebrated Wigner-Dyson distribution function [10]:

$$\mathcal{P}(\{E\}) = \text{const.} \times \prod_{n < m} |E_n - E_m|^b \exp\left(-\sum_k \frac{E_k^2}{\sigma^2}\right), \quad (3.2)$$

where σ is the standard deviation, fixing the unit of energy, and $b = 1, 2$ or 4 for orthogonal, unitary and symplectic ensembles respectively. Most of our work is independent of the symmetry class, however our default class will be the Gaussian orthogonal ensemble (GOE) with $b = 1$ when not specified otherwise.

3.1 Estimate of the OTOC and its plateau

The idea is to use the results recapitulated in the previous section to find the ensemble expectation value of OTOC from the “master formulas” (2.2)–(2.3). Representing the eigenvectors and the operators as matrices in some (arbitrary) basis we can obviously write out

$$a_{nk} = \sum_{ij} \psi_i^{(n)} \psi_j^{(k)} A_{ij} \Rightarrow \langle a_{nk} \rangle = \int d^N \psi^{(n)} \int d^N \psi^{(k)} P(\psi^{(n)}, \psi^{(k)}) \psi_i^{(n)} \psi_j^{(k)} A_{ij}, \quad (3.3)$$

and similarly for b_{nk} . Inserting the above expression for the matrix elements into (2.3), multiplying $c_{mn}(t)$ by its complex conjugate taking into account the reality of the eigenvectors and relabelling the indices in the sums where convenient we find (denoting the average over the random matrix ensemble by $\langle C(t) \rangle$):

$$\begin{aligned} \langle C(t) \rangle = & \int d^{N^2} \{c\} \int d^N \{E\} \mathcal{P}(\{E\}) P(\{c\}) \sum_{n,m} \sum_{k,k'} \sum_{i_1,2} \sum_{j_1,2} \sum_{i'_1,2} \sum_{j'_1,2} \sum c_{j_1}^k c_{i_2}^k c_{j_1}^{k'} c_{i_2}^{k'} c_{i_1}^n c_{i_1}^n c_{j_2}^m c_{j_2}^m e^{-\beta E_n} \times \\ & \times \left(A_{i_1 i_2} A_{i'_1 i'_2} B_{j_1 j_2} B_{j'_1 j'_2} e^{i(E_{k'} - E_k)t} + A_{i_2 j_2} A_{i'_2 j'_2} B_{i_1 j_1} B_{i'_1 j'_1} e^{i(E_k - E_{k'})t} - \right. \\ & \left. - A_{i_2 j_2} A_{i'_1 i'_2} B_{i_1 j_1} B_{j'_1 j'_2} e^{i(E_k + E_{k'} - E_m - E_n)t} - A_{i_1 i_2} A_{i'_2 j'_2} B_{j_1 j_2} B_{i'_1 j'_1} e^{i(E_m + E_n - E_k - E_{k'})t} \right), \end{aligned} \quad (3.4)$$

where $\{c\}$ determines the whole set of N^2 random elements $c_j^{(n)}$ with $j, n = 1 \dots N$ and likewise $\{E_n\}$ is the whole set of eigenenergies. All the sums run from 1 to N . The integral over the eigenvector elements $\{c\}$ in (3.4) produces only an overall constant as these coefficients do not couple to the other quantities (in fact the integral $d^{N^2} \{c\}$ is a textbook Jeans integral, but we do not need its value as it only produces an N -dependent, T -independent constant). The remaining integral, over the eigenenergies, is again a sum of products of Jeans-type integrals but with an additional linear term $-\beta E$ in the exponent. Notice that the imaginary (sine) terms in (3.4) cancel out when the sum is performed; this

is a consequence of the module squared in $|c_{mn}|^2$, i.e. of the reality of OTOC. Now we see that (3.4) becomes a sum where each term is a product of factors of the form

$$p_i(E_i)e^{-E_i^2/4\sigma^2-\beta E_i} \cos(sE_it), \quad s \in \{0, 1\},$$

where p_i is some polynomial and s may be zero or unity, i.e. some terms have this factor and some do not. Every such term is a Jeans-type integral. The number of terms in $\mathcal{P}(\{E\})$ equals the number of partitions of $N(N-1)b/2$, and the sums over the coefficients $\{c\}$ bring altogether N^{12} terms. When everything is said and done (for details see appendix A), the final outcome, ignoring the multiplicative constant factors, reads:

$$\begin{aligned} \langle C(t) \rangle = & \prod_{a=1}^4 \sum_{\substack{\sum_j \alpha_j^a = N(N-1)b/2 \\ \alpha_1^a, \dots, \alpha_N^a}} \left[{}_1F_1 \left(\frac{1 + \alpha^a}{2}, \frac{1}{2}, \frac{\sigma^2}{4} (\beta - it)^2 \right) \right. \\ & + (\beta - it) {}_1F_1 \left(\frac{2 + \alpha^a}{2}, \frac{3}{2}, \frac{\sigma^2}{4} (\beta - it)^2 \right) + {}_1F_1 \left(\frac{1 + \alpha^a}{2}, \frac{1}{2}, \frac{\sigma^2}{4} (\beta + it)^2 \right) \\ & \left. + (\beta + it) {}_1F_1 \left(\frac{2 + \alpha^a}{2}, \frac{3}{2}, \frac{\sigma^2}{4} (\beta + it)^2 \right) \right], \end{aligned} \quad (3.5)$$

where ${}_1F_1$ is the confluent hypergeometric function. The sum runs over all partitions of $N(N-1)b/2$, and the product has four terms as each factor $|c_{mn}|^2$ has four matrix elements of A and B .

3.1.1 Kinematic OTOC

In order to move further we need to specify at least to some extent the operators A and B . We will consider (1) the kinematic OTOC, with $A = x$, $B = p$ (2) generic sparse operators, with $O(N)$ nonzero elements in the matrices a_{mn} and b_{mn} , and (3) dense operators A and B , with $O(N^2)$ nonzero elements, in particular the case when the operators A , B are themselves represented by Gaussian random matrices. Let us estimate OTOC for each case.

For the kinematic OTOC, $A_{ij} = x_i \delta_{ij}$ is diagonal and in the large- N limit B can be approximated as $B_{ij} \sim \delta_{ij}/x_i$. The Kronecker deltas reduce the number of terms in the sums over $\{c\}$ to N^4 , the number of partitions $\sum_j \alpha_j = n$ can be approximated as $p(n) \sim \exp(\pi\sqrt{2n/3})/\sqrt{n}$, and the general expression (3.4) becomes²

$$\langle C(t) \rangle \sim e^{\pi\sqrt{\frac{1}{3}}N} N^3 e^{\frac{\sigma^2\beta^2}{4}} \left(W_0(\sigma\beta) + Q_1 \left(\cos \frac{\sigma^2\beta t}{2} \right) W_1(\sigma\beta) + Q_2 \left(\sin \frac{\sigma^2\beta t}{2} \right) W_2(\sigma\beta) \right), \quad (3.6)$$

where $W_{0,1,2}$ are polynomials in $\sigma\beta$ of degree $N(N-1)b/2 \sim N^2b/2$, Q_1 is an even polynomial (with only even powers) of the same degree, and Q_2 is an odd polynomial of the same degree. Each coefficient in the polynomials $W_{0,1,2}$ comes from $\sim N^2$ terms (appendix A), therefore the size of the coefficients scales approximately with N^2 . Eq. (3.6)

²One might be surprised by the unusual dependence on N . This happens because we have not normalized $C(t)$ by the product $\langle AA \rangle \langle AA \rangle$ as it is usually done. With appropriate normalization, $C(t)$ would of course be of order unity.

is a very complicated oscillating function as many terms are involved. But if we are only interested in the average value of $C(t)$ at long times, we may simply ignore the oscillations (which in the first approximation average out to some value of order unity) and write the estimate for the long-term, saturated or plateau OTOC value that we denote by C_∞ :

$$C_\infty \sim \langle C(t \rightarrow \infty) \rangle \sim e^{\pi\sqrt{\frac{b}{3}}N} N^3 e^{\frac{\sigma^2\beta^2}{4}} W_0(\sigma\beta) \quad (3.7)$$

We deliberately do not write $\lim_{t \rightarrow \infty}$ in the above definition as the limit in the strict sense does not exist because of the oscillatory functions, and in addition our derivation is obviously nothing but a crude estimate. A similarly rough estimate of the temperature dependence of C_∞ can be obtained in the following way. For sufficiently large $\sigma\beta$, roughly $\sigma\beta/N^2 > 1$, the polynomial W_0 is dominated by the highest-degree term and we have, from (3.7):

$$C_\infty \sim (\sigma\beta)^{\frac{N^2b}{2}} e^{\frac{\sigma^2\beta^2}{4}} + \dots \sim e^{\frac{\sigma^2}{4T^2}} + \dots, \quad (3.8)$$

where in the second step we have assumed $\beta \gg 1$ so that the power-law prefactor $\beta^{N^2b/2}$ becomes negligible compared to the exponential. We deliberately emphasize that there are other terms in the expansion (...), including also a constant term (from the zeroth-order term in W_0). This is important as it tells us that the scaling is in general of the form $C_\infty \approx \text{const.} + \exp(\sigma^2/4T^2)$, i.e. the temperature dependence is superimposed to a constant. This is also expected as the (appropriately normalized) saturated value C_∞ should always be of order unity, and the temperature dependence will only account for the relatively small differences between the plateau values of $C(t)$, as we will see later in figures 1 and 2.

On the other hand, for sufficiently small $\sigma\beta$, the polynomial W_0 can be estimated as a geometric sum of monomials in $-\sigma\beta N^2$ (remember the terms in W_0 have alternating signs):

$$C_\infty \sim \frac{e^{\frac{\sigma^2\beta^2}{4}}}{1 + \sigma\beta N^2} \sim 1 - \sigma\beta N^2 + O(\beta^2). \quad (3.9)$$

We have now reached an important point: the plateau OTOC falls off exponentially with $1/T^2$ at low temperatures³ and grows as a function of $1/T$ at high temperatures (we are not sure which function, as there are higher order terms in addition to the one written in (3.9), and there is no clearly dominant term like the exponential at large β), with the crossover temperature:⁴

$$T_c \sim \sigma N^2. \quad (3.10)$$

If we consider a pair of arbitrary sparse operators A and B , the whole above reasoning remains in place, except that the products of matrix elements such as $A_{i_1 i_2} A_{i'_1 i'_2} B_{j_1 j_2} B_{j'_1 j'_2}$ remain as arbitrary constants. Therefore we get the same qualitative behavior with two regimes and a crossover between them. The crossover temperature is very high for typical $N \gg 1$ (otherwise the random matrix formalism makes little sense) and finite σ (again, $\sigma \rightarrow 0$ makes little sense). In particular, in the $N \rightarrow \infty$ limit the crossover temperature becomes infinite and the only regime is the exponential decay.

³Actually, the falloff rate equals $\text{const.}/T^2$ with some system-specific constant, but for brevity we will denote it schematically as the $1/T^2$ regime throughout the paper.

⁴The crossover temperature is determined simply as $\beta_c \sigma N^2 = 1$, i.e. whether the terms in W_0 grow or decay at higher and higher order.

3.1.2 OTOC for dense and/or random operators

Now consider the case when the matrix elements in (3.5) are generically all nonzero (and for now nonrandom, i.e. we fix the operators and do not average over them). The large- t limit yields the expression

$$\langle C(t) \rangle \sim e^{\pi\sqrt{\frac{b}{3}}N} N^{11} e^{\sigma^2(\beta^2-t^2)} [q_0(\sigma t) w_0(\sigma\beta) + q_1(\sigma t) w_1(\sigma\beta + it, \sigma\beta - it)], \quad (3.11)$$

where $q_{0,1}$ and w_0 are polynomials of degree $N(N-1)b/2 \sim N^2b/2$, and w_1 is the polynomial of the same total degree of two variables, $\sigma\beta + it$ and $\sigma\beta - it$. The coefficients of $w_{0,1}$ are proportional to products of matrix elements $A_{i_1j_1} B_{k_1l_1} \dots A_{i_8j_8} B_{k_8l_8}$, which are roughly proportional to $|A|^8|B|^8$. The long-time limit yields

$$C_\infty \sim \frac{e^{\pi\sqrt{\frac{b}{3}}N}}{N^5} (|A||B|)^8 e^{\frac{\sigma^2\beta^2}{4}} w_0(\sigma\beta)w_1(\sigma\beta, \sigma\beta), \quad (3.12)$$

but now a typical coefficient of the polynomials $w_{0,1}$ behaves as $N^2(|A||B|)^8$. Therefore, the scaling in the low-temperature regime remains the same as (3.8): $C_\infty \sim \exp(\sigma^2/T^2)$. But the high-temperature regime yields

$$C_\infty \sim 1 - \sigma\beta N^2 (|A||B|)^8 + O(\beta^2), \quad (3.13)$$

therefore the crossover now happens at

$$T_c \sim \sigma N^2 (|A||B|)^8 \quad (3.14)$$

and therefore may be lower than the very high value (3.10), depending on the norm of the operators A and B .

Finally, if the operators A and B are both random Hermitian matrices (for concreteness, from the Gaussian unitary ensemble with the distribution function Π and the standard deviation ξ), we need to average also over the distribution functions for A and B and work with the double average $\langle\langle C(t) \rangle\rangle$:

$$\langle\langle C(t) \rangle\rangle \equiv \int d^N\{a\} \int d^N\{b\} \Pi(\{a\}) \Pi(\{b\}) \langle C(t) \rangle \sim \xi^{N^2} \langle C(t) \rangle. \quad (3.15)$$

This estimate is very crude, based simply on the fact that the distribution functions Π have $\sim N^2/2$ pairs of the form $(a_i - a_j)^2$. The important point is that the scaling from (3.13) that behaves essentially as $\sim \xi^{16}$ now becomes $\sim \xi^{N^2}$, therefore the crossover temperature is significantly reduced compared to (3.14) and behaves as $T_c \sim \sigma N^2 \lambda^{N^2}$. So for random operators the crossover may happen at temperatures that are not very high and thus can be clearly visible in the numerics and experiment.

3.2 Numerical checks

Now we demonstrate numerically that the crude estimates from the previous subsection indeed make sense and describe the characteristic behavior of OTOC. Our chief goal is to understand the behavior of C_∞ , however it is instructive to start from the time dependence

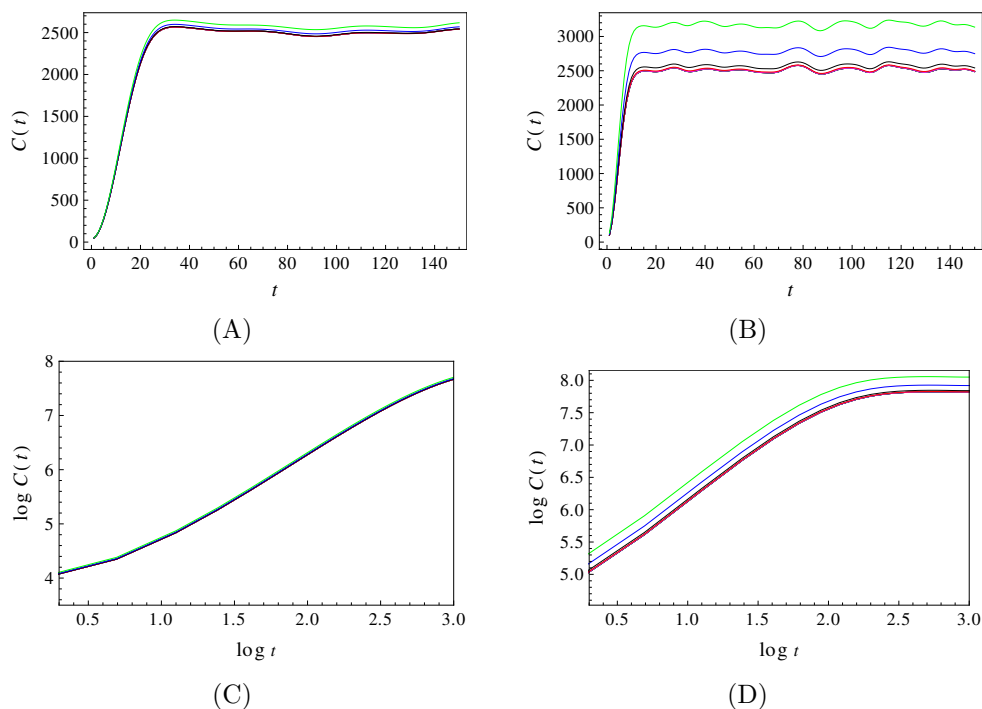


Figure 1. Numerically computed and averaged kinematic OTOC $C(t)$ for an ensemble of $l = 1000$ Gaussian orthogonal matrices of size $N = 20$ for the deviation $\sigma = 0.02$ (A, C) and $\sigma = 0.05$ (B, D), at temperatures 0.67 (black), 1.00 (blue), 1.25 (green), 2.50 (magenta), 5 (red). The plots (A, B) show the linear scale and the plots (C, D) the log-log scale. Crucially, the growth regime is not exponential and is actually closer to a power law. The growth ends on a plateau with superimposed oscillations. The plateaus differ slightly for different temperatures — the main effect we look at in this paper. Times is in units $1/\sigma$ in all plots.

of the kinematic OTOC (figure 1). We find the expected pattern of early growth followed by a plateau, however the growth is closer to a power law than to an exponential; this follows from the polynomial terms in (3.6), although the power law is not perfect either, as we see in the panels (A, C). This is in line with the prediction of [13], where the authors find

$$\langle C(t) \rangle \approx J_1^4(2t)/t^4 + t(t/2 - 1), \quad (3.16)$$

for a slightly different ensemble of random matrices (J_1 is the Bessel function of the first kind). This function is also neither an exponential nor a power law but at early times it is best approximated by a power law at leading order (at long times it falls off exponentially but the saturation is reached already prior to that epoch). In figure 2 we focus on the plateau behavior. It has the form of a constant function with superimposed aperiodic oscillations, and the differences of the plateau values are the subject of our theoretical predictions. These are relatively small and become important only when N is finite and not very large. In figure 2 we plot again the time dependence of the kinematic OTOC but now we focus on long timescales, to confirm that the plateau is indeed stable, and to show the very complex oscillation pattern superimposed on the plateau.

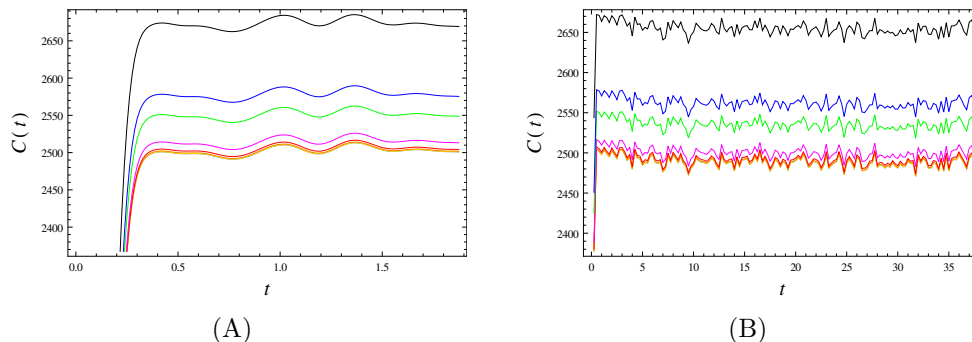


Figure 2. Numerically computed and averaged kinematic OTOC $C(t)$ for an ensemble of $l = 1000$ Gaussian orthogonal matrices of size $N = 60$ for the deviation $\sigma = 0.1$, at temperatures 0.67 (black), 1.00 (blue), 1.25 (green), 2.50 (magenta), 5 (red), 10 (orange), 20 (yellow) and 100 gray. In (B) we plot the same as in (A) but over a longer timescale, showing that the plateau remains stable for long times, i.e. represents true asymptotic behavior.

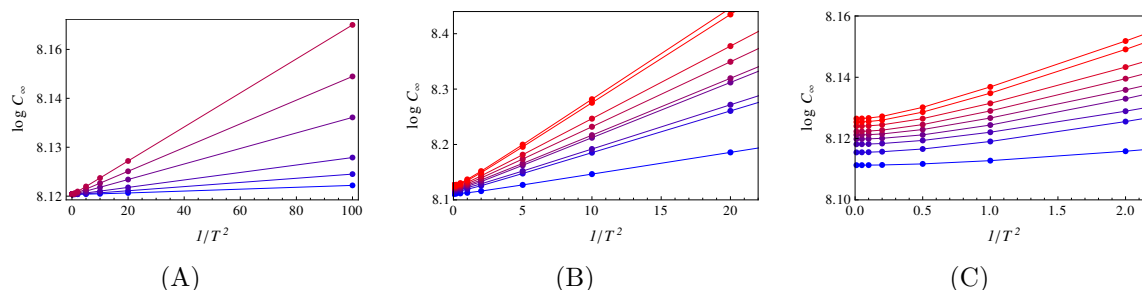


Figure 3. (A) The logarithm of the amplitude of the plateau C_∞ of the kinematic OTOC for the deviations $\sigma = 0.05, 0.1, 0.2, 0.4, 0.6, 0.8$ (blue to red) as a function of temperature for β values. The linear dependence is nearly perfect, in accordance with the predicted scaling $\log C_\infty \propto 1/T^2$. The matrix size is $N = 20$. (B) Same as (A) but for the deviation $\sigma = 0.4$ and varying matrix size $N = 10, 20, 40, 60, 80, 100$ and 120 (blue to red). In (C) we bring the zoom-in of the plot (B) for high temperatures. Apart from a slight deviation near $\beta = 0$, the behavior for larger matrices is still fully consistent with the analytical prediction. The solid lines are just to guide the eye.

Figure 3 confirms our main prediction for the low-temperature regime (again for the kinematic OTOC) — clear linear scaling of $\log C_\infty$ with $1/T^2$ in a broad range of temperatures. At small inverse temperatures there is some deviation from the linear scaling law but this we also expect. Looking now at the OTOC for a pair of random Hermitian operators in figure 4, we detect also the other regime at small enough temperatures — $\log C_\infty$ decays with the inverse temperature. This regime is likely present also in figure 3, but only at extremely high temperatures (which we have not computed in that figure).

4 OTOC for weakly and strongly chaotic Hamiltonians

For quasi-integrable few-degrees-of-freedom Hamiltonians one would naively expect that OTOC closely resembles the Lyapunov exponent, at least for high quantum numbers, approaching the classical regime. As we have already commented in the Introduction, it

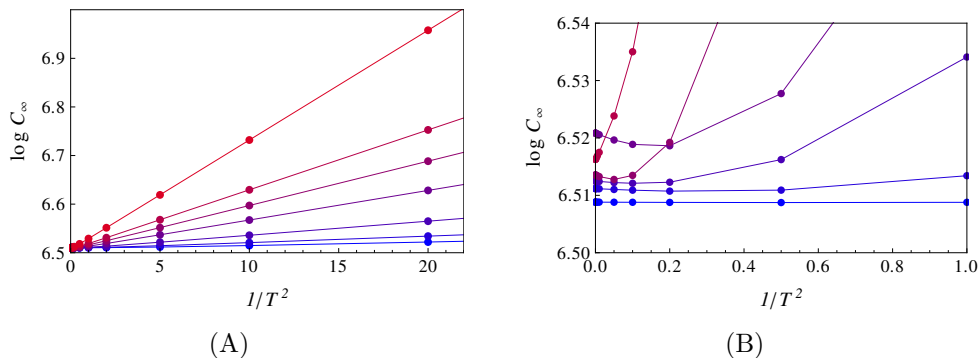


Figure 4. The saturated OTOC $C_\infty(T)$ of a pair of dense random operators A and B for the Gaussian orthogonal random Hamiltonian with $\sigma = 0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8$ (blue to red). In (A), looking at the full range of C_∞ values, it is obvious that the dominant regime is still the $\exp(c/T^2)$ scaling. However, focusing on the low- σ ensembles (B), we notice the high-temperature growing behavior of the OTOC plateau which is absent for sparse operators.

is known that this is not true in general [6, 8, 9, 16, 17] and that both chaotic systems with zero quantum Lyapunov exponent and regular systems with a nonzero exponent exist. We will now try to find some common denominator of OTOC dynamics in (deterministic) quantum-mechanical systems. It will quickly become clear from our general analysis of the master formula (2.3) that the function $C(t)$ is as complicated as for random matrices (indeed, even more so). But we will again construct an upper bound for the saturated OTOC value and arrive at a rough scaling estimate.

4.1 Weak chaos: perturbation theory

As an example of a quasi-integrable system (of the form $H = H_0 + \epsilon V$ where H_0 is integrable and the perturbation V makes it nonintegrable for $\epsilon \neq 0$) consider the Henon-Heiles Hamiltonian

$$H = \frac{1}{2} (p_x^2 + p_y^2) + \frac{1}{2} (\omega_x^2 x^2 + \omega_y^2 y^2) + \epsilon \left(x^2 y - \frac{1}{3} y^3 \right), \quad (4.1)$$

a well-known paradigm for classical chaos with applications in galactic dynamics and condensed matter. For $\epsilon = 0$ it obviously reduces to a 2D linear harmonic oscillator and becomes integrable. As we know, nonintegrability does not always imply chaos; indeed, this is a typical system with mixed phase space, with both chaotic and regular orbits. Chaotic orbits proliferate only when the perturbation is larger than some finite ϵ_c ; they are almost absent at low energies, numerous at intermediate energies and again absent at very high energies when the potential energy is negligible compared to the kinetic energy [28, 29]. For such a quasi-integrable system our idea is to apply elementary perturbation theory in the occupation number basis to estimate OTOC starting from (3.4). We will present the perturbation theory in a fully general way, for an arbitrary Hamiltonian $H_0 + \epsilon V$, and some of the conclusions will also turn out to be quite general. Only at the end we will show the quantitative results for the Henon-Heiles system (4.1).

Let us write the perturbative expression for OTOC. Replacing the initial basis states $|n\rangle$ with the first-order⁵ perturbatively corrected states $|n + \delta n\rangle$ and introducing likewise the perturbative corrections $\delta a_{mn}, \delta b_{mn}$ for the matrix elements of A and B , the equation (2.3) becomes

$$\begin{aligned}
c_{mn}^{(1)} &= c_{mn} + \sum_k (\delta a_{mk} b_{kn} + a_{mk} \delta b_{kn}) e^{-iE_{km}t} - \sum_k (\delta b_{mk} a_{kn} + b_{mk} \delta a_{kn}) e^{-iE_{nk}t} = \\
&= c_{mn} + \sum_{kl} (\delta_{ml} a_{lk} b_{kn} + \delta_{kl}^* a_{lm}^* b_{kn} + \delta_{kl} b_{ln} a_{mk} + \delta_{nl}^* b_{lk}^* a_{mk}) e^{-iE_{km}t} + (a \leftrightarrow b) e^{-iE_{nk}t} = \\
&= c_{mn} + \left(\delta \cdot A \cdot B + A^\dagger \cdot \delta^\dagger \cdot B + A \cdot \delta \cdot B + A \cdot B^\dagger \cdot \delta^\dagger \right)_{mn} e^{-iE_{km}t} + (A \leftrightarrow B)_{mn} e^{-iE_{nk}t}.
\end{aligned} \tag{4.2}$$

Now we insert this result into (2.2) and apply the Cauchy-Schwarz-Bunyakovski inequality:

$$\begin{aligned}
C^{(1)}(t) &= \frac{1}{Z} \sum_{mn} e^{-\beta E_n} |c_{mn}^{(1)}|^2 \leq \\
&\leq \frac{1}{Z} \sum_{mn} e^{-\beta E_n} \left(|c_{mn}|^2 + |\delta \cdot A \cdot B + A^\dagger \cdot \delta^\dagger \cdot B + A \cdot \delta \cdot B + A \cdot B^\dagger \cdot \delta^\dagger|_{mn}^2 + (A \leftrightarrow B) \right) = \\
&= C(t) + \frac{1}{Z} \left(4\text{Tr} \left(B^\dagger \cdot A^\dagger \cdot \delta^\dagger \cdot \tilde{\rho}^2 \cdot \delta \cdot A \cdot B \right) + 4\text{Tr} \left(B^\dagger \cdot \delta \cdot A \cdot \tilde{\rho}^2 \cdot A^\dagger \cdot \delta^\dagger \cdot B \right) \right) \leq \\
&\leq C_\infty + \frac{8}{Z} |\tilde{\rho}|^2 |A|^2 |B|^2 |\delta|^2 \equiv C_\infty + \delta C_\infty,
\end{aligned} \tag{4.3}$$

where $\tilde{\rho} \equiv \text{diag}(\exp(-\beta E_n))$ is the non-normalized density matrix. In the above derivation, we have also used the definition of trace and the definition of thermal expectation values $\langle A \rangle \equiv \text{Tr}(\rho \cdot A)$. The norm of a matrix is defined as usual by $|A|^2 \equiv \text{Tr} A^\dagger A$. This estimate manifestly replaces $C(t)$ by its asymptotic (maximum) value, as we have replaced the terms containing the time-dependent phase factors by their time-independent norms.

In order to move further, notice first that $|\tilde{\rho}|^2 = \sum_n \exp(-2\beta E_n) = Z(2\beta)$ for a canonical ensemble with the diagonal density matrix that we consider here. This means, from (4.3), that the temperature dependence is encapsulated in the ratio $Z(2\beta)/Z(\beta)$. The prefactor will again differ between dense and sparse A, B and V . For sparse matrices, we can write $|A|^2 \sim Na^2$ whereas for dense matrices we have $|A|^2 \sim N^2 a^2$, assuming that all matrix elements have some characteristic magnitude a . Obviously, if this is not true the outcome will be more complicated, but it seems this does not influence the temperature dependence. For concreteness we assume sparse A and B . For sparse V with nonzero elements of order v concentrated near the diagonal (this is true for the Henon-Heiles Hamiltonian and in general for perturbations expressed as low-degree polynomials in coordinates and momenta), we can estimate $|\delta|^2 \sim Nv^2/\omega^2$. Here we assume an approximately equidistant spectrum of H_0 with frequency (neighboring level spacing) ω . This yields:

$$C_\infty^{(1)} \sim C_\infty + \frac{Z(2\beta)}{Z(\beta)} N^3 a^2 b^2 \frac{v^2}{\omega^2}. \tag{4.4}$$

⁵The whole argument works the same way also for higher-order perturbation theory; we assume first order just for simplicity.

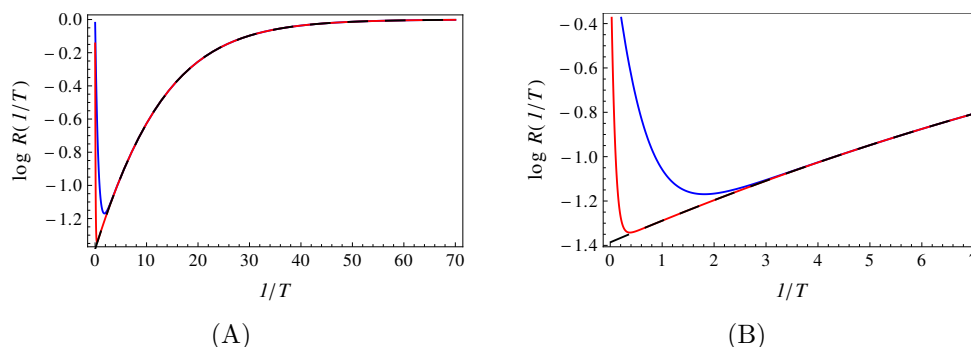


Figure 5. The thermal dependence factor of $\log C_\infty$ for weak perturbative chaotic systems, given by $Z(2\beta)/Z(\beta)$ with $\beta = 1/T$ the inverse temperature, here given for a 2D linear harmonic oscillator with the frequencies $\omega_x = \omega_y = 0.1$, with $N = 20$ levels (blue) and with $N = 150$ levels (red). We also plot the sum over $N = \infty$ levels (black). In (A) we zoom in at high β /low temperatures, and in (B) we focus on smaller β /larger T . The $N = 150$ plot is already quite close to the monotonic $N = \infty$ dependence but at high temperatures there is always a region decaying with β , before the approximately linear $\log R(1/T)$ dependence sets in, just like in the numerical results. At very low temperatures the ratio saturates, as we see in the panel (A). This ensures that our estimate for the saturated OTOC has a finite limit at zero temperature. The overall scale is arbitrary as the R factor is always multiplied by various other factors.

For a dense perturbation V , the only factor that changes is $|\delta|$. Assuming again the utterly simple situation where all matrix elements of V are roughly equal v , we have $\delta_{mn} \sim v/E_{mn} \sim v/(\omega(m-n))$, which yields a series that can be summed analytically. However, we will not pursue this further as the temperature dependence is universal in all cases, given by the simple ratio of the partition functions:

$$C_\infty \propto R(\beta) \equiv \frac{Z(2\beta)}{Z(\beta)} \rightarrow \frac{\sum_{j=1}^N e^{-2\beta\omega}}{\sum_{j=1}^N e^{-\beta\omega}} \rightarrow \frac{e^{\beta\omega}}{1 + e^{\beta\omega}}. \quad (4.5)$$

The first simplification holds when H_0 is a 1D harmonic oscillator, and the second one when $N \rightarrow \infty$. But the basic result (the ratio of partition functions) always holds. We are in fact more interested in the 2D harmonic oscillator, which is the integrable part of the Henon-Heiles Hamiltonian. For that case, we plot the sum (for finite N) in figure 5. Of course, the analysis of the function $R(\beta)$ is trivial — we plot it in the figure merely to emphasize the qualitative agreement with the numerics.

As a final remark, what we have found is the correction of the OTOC plateau δC_∞ . There is still the unperturbed value of C_∞ for the integrable Hamiltonian H_0 . We know that this can be nonzero and even quite large because of local instability [6, 7, 9]. We are mainly interested in the opposite situation, when the scrambling chiefly comes from chaos so that OTOC does not grow when $H = H_0$. In this case $C^{(1)} \approx \delta C_\infty$ and the temperature dependence is primarily determined by (4.4). In the next subsection we will look both at the Henon-Heiles system where this holds, and a perturbed inverse chaotic oscillator where H_0 is unstable.

4.2 Weak chaos: examples and numerics

As our main example we can now study the Henon-Heiles system of eq. (4.1). Starting from the nonperturbed Hamiltonian (2D harmonic oscillator), we will express the nonzero elements of $c_{mn}(t)$ exactly, i.e. we will not use the estimates (4.3) as the perturbation is quite simple and amenable to analytic treatment. Denoting a basis state by the quantum number $n = (n_x, n_y)$, we can write the amplitudes $c_{n_x n_y n'_x n'_y}$ as products of amplitudes of the two decoupled subsystems $c_{n_x n_y n'_x n'_y} = C_{n_x n'_x} C_{n_y n'_y}$, with

$$\begin{aligned} C_{n_x n_x} &= -n_x \omega_x \cos t \\ C_{n_x, n_x-2} &= \frac{i}{2} \sqrt{n_x-1} \left(\sqrt{n_x+1} e^{-i\omega_x t} - \sqrt{n_x+2} e^{i\omega_x t} \right) \\ C_{n_x, n_x+2} &= \frac{i}{2} \sqrt{n_x+1} \left(\sqrt{n_x-1} e^{i\omega_x t} - \sqrt{n_x-2} e^{-i\omega_x t} \right), \end{aligned} \tag{4.6}$$

and all other elements are zero; for the y quantum numbers the coefficients are the same with $(n_x, \omega_x) \mapsto (n_y, \omega_y)$. For nonzero ϵ , the off-diagonal matrix elements can be represented exactly as

$$c_{n_x n'_x n_y n'_y}(t) = \epsilon \delta_{|n_x-n'_x|-2} \delta_{|n_y-n'_y|-1} \sqrt{m_x(m_x-1)} \sqrt{m_y+1}, \quad m_{x,y} \equiv \min(n_{x,y}, n'_{x,y}). \tag{4.7}$$

The state vectors are now easily calculated in textbook perturbation theory. We have compared the analytic calculation to the numerics and find that they agree within a relative error ≤ 0.04 ; therefore, one may safely use (4.6)–(4.7) in order to speed up the computations and avoid numerical diagonalization of large matrices.

The magnitude of the plateau value of $C(t)$, computed by long-time averaging similarly to the random matrix calculations in section 3, are given in figure 6. At large T values, C_∞ decays with $1/T$, at intermediate values it shows an exponential growth with $1/T$ just like $R(1/T)$ in figure 5, and as the temperature goes to zero it reaches a finite value. In figure 7 we consider a system with much reduced state space, with $N = 25$. We expect that for small N the existence of two regimes is more clearly visible, and that the crossover temperature is higher. This is indeed what happens, although the exact form of the function $C_\infty(1/T)$ is not very well described by the analytical result. As we have made many crude approximations, this is not surprising: our analytical result still explains the existence of two regimes and the crossover between them.⁶ One unexpected finding is that the high-temperature regime is apparently universal for all perturbation strengths and scales as $C_\infty \propto \exp(-4\pi/T)$. This is probably specific for the Henon-Heiles system; we do not understand it at present.

It is instructive to look at the energy level statistics of the Henon-Heiles system for the same parameters that we have used for the OTOC calculation, in order to understand the relation of OTOC to chaos. In figure 8 we plot the histograms of the neighboring level

⁶One might regard such truncation of the state space as artificial and unphysical. It is clearly just a technical step in order to show the effect we seek for more clearly, however in principle it can be realized by introducing an additional external potential. In other contexts, e.g. finite spin systems, a finite Hilbert space is perfectly natural.

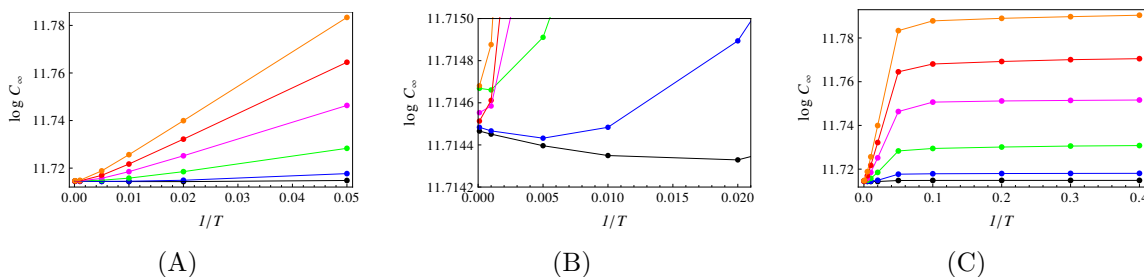


Figure 6. (A) The saturated kinematic OTOC value C_∞ for a range of inverse temperatures $\beta = 10^{-4}, 10^{-3}, 5 \times 10^{-3}, 0.01, 0.02, 0.05$ and a range of perturbation strengths $\epsilon = 1, 2, 5, 10, 15, 20$ (black, blue, green, magenta, red, orange). Here we see the scaling $C_\infty \propto \exp(c/T)$, with c growing with ϵ . In the (B) panel we zoom in the high-temperature region, to show that for $\epsilon \leq 5$ there is also the other regime where C_∞ grows with T . Since the number of points in this interval is small it is not easy to judge the form of T -dependence. In (C) we focus on the opposite regime, at very low temperature, showing that C_∞ saturates as $T \rightarrow 0$. This is again in accordance with the $\beta \rightarrow \infty$ limit of $Z(2\beta)/Z(\beta)$. The system is truncated to $N = 144$ levels.

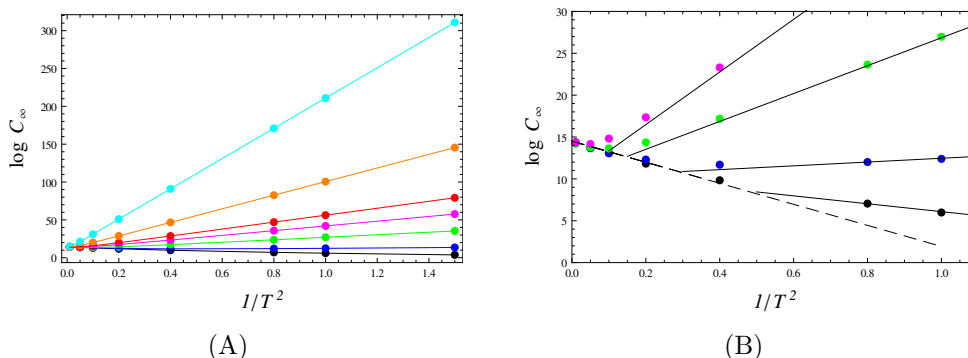


Figure 7. The saturated kinematic OTOC value $C_\infty(T)$ for the truncated Henon-Heiles model with $N = 25$ levels. The perturbation strength is $\epsilon = 0.1, 0.3, 0.5, 0.7, 0.9, 1.5, 3.0$ (black, blue, green, magenta, red, orange, cyan). Already in (A) we see that for $\epsilon \leq 0.5$ there is a finite crossover temperature T_c so that C_∞ grows with temperature for $T > T_c$. Since T_c goes down when the Hilbert space is reduced, we can observe the high-temperature regime very clearly and see that it collapses to a universal law $C_\infty \sim \exp(-4\pi/T)$. This is seen in the panel (B) where we zoom in at the interesting region.

spacing. Even for large ϵ , the regular (Poisson) component is dominant over the chaotic (Wigner-Dyson) component. In other words, the classically mixed phase space, with the increasing chaotic component, is almost completely regular in the quantum regime; quantum chaos is “weaker”, as is often found in the literature [33]. For us, the fact that the system is outside the Wigner-Dyson regime means that indeed the behavior of C_∞ is a good litmus test of quantum dynamics, behaving (at low temperatures) as $\exp(1/T^2)$ or $\exp(1/T)$ for strong or weak chaos respectively.⁷ Indeed, we would not expect that a system which is well described by perturbation theory shows strong level repulsion.

⁷In fact, this is only true provided that the scrambling is chaos-related, i.e. that the integrable limit with $\epsilon = 0$ and $H = H_0$ does not scramble significantly. We will come to this issue in the next paragraph.

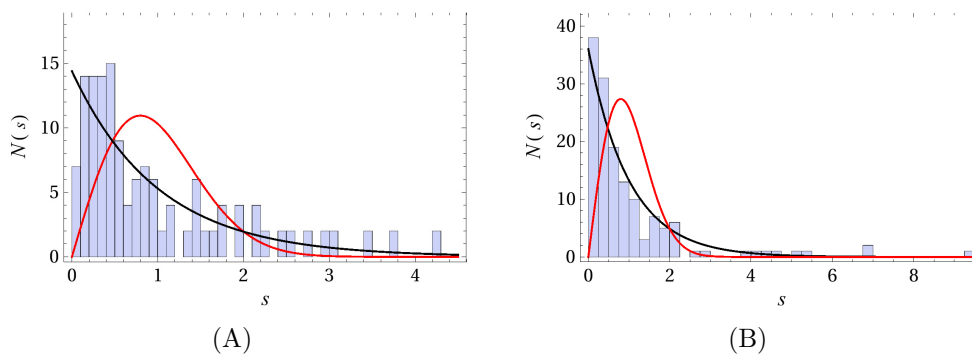


Figure 8. Distribution of neighboring energy level spacings $N(s)$ for the Henon-Heiles Hamiltonian with $\epsilon = 0.1$ (A) and $\epsilon = 1.5$ (B). In each plot we compare the level distribution to the Poisson law ($\exp(-s)$) and the Wigner-Dyson law for orthogonal matrices ($s \exp(-\pi s^2/4)$). The distribution is dominantly Poissonian even for large perturbations, although there is a small admixture of Wignerian statistics. The perturbative dynamics of the Henon-Heiles system is always weakly chaotic in quantum mechanics (despite being classically strongly chaotic for large enough ϵ).

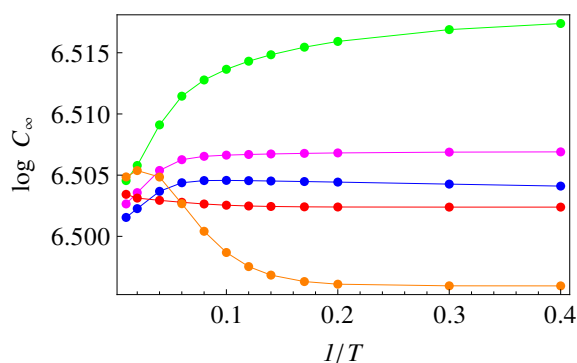


Figure 9. Temperature dependence of the asymptotic OTOC C_∞ for the inverse Henon-Heiles Hamiltonian, with $\omega_x^2 = 4\omega_y^2 = -1$, and perturbation strength $\epsilon = 0, 0.1, 0.5, 0.9, 1.5$ (blue, green, magenta, red, orange). The curves are more or less flat and without a clear trend, as the scrambling is rooted in local instability, not chaos.

Finally, it is instructive to look at the inverse Henon-Heiles system, with $(\omega_x^2, \omega_y^2) \mapsto (-\omega_x^2, -\omega_y^2)$, so that H_0 is the inverse harmonic oscillator. As already found in the literature, scrambling is significant already at $\epsilon = 0$, and this contribution dominates even at high ϵ , at all temperatures. In other words, neither the perturbation nor the temperature have a significant influence over C_∞ . This is fully in accordance with the result (4.3) and the morale is that OTOC directly describes scrambling, and chaos only indirectly, through the scrambling, *if the scrambling originates mainly from chaos*; if not, OTOC is largely insensitive to chaos. Therefore, the temperature dependence of the OTOC value, derived from the assumptions about the dynamics (perturbative chaos or strong, random-matrix chaos) cannot be seen when there is a strong non-chaotic component to scrambling (figure 9).

4.3 Strong chaos: numerics and the return to random matrices

As a final stroke, we will now examine a strongly chaotic system which is also relevant for black hole scrambling and related problems in high energy physics. This is the bosonic sector of the D0 brane matrix model known as the BMN (Berenstein-Maldacena-Nastase) model [30], obtained as a deformation of the BFSS (Banks-Fischler-Shenker-Susskind) model [31] by a mass term and a Chern-Simons term. This model has been studied in detail in the context of non-perturbative string and M theory. It is known to describe the dynamics of M theory on pp-waves and is also related to the type IIA string theory at high energies; for details one can look at the original papers or the review [32]. Following [33–36], we focus solely on the bosonic sector which is enough to have strongly chaotic dynamics with equations of motion that are not too complicated. The quantum-mechanical Hamiltonian of the BMN bosonic sector reads:⁸

$$H = \text{Tr} \left(\frac{1}{2} \Pi^i \Pi^i - \frac{1}{4} [X^i, X^j] [X^i, X^j] + \frac{1}{2} \nu^2 X^a X^a + \frac{1}{8} \nu^2 X^\alpha X^\alpha + w \varepsilon_{abc} X^a X^b X^c \right) \\ i \in \{1 \dots 9\}, \quad a, b, c \in \{1, 2, 3\}, \quad \alpha \in \{4 \dots 9\}, \quad (4.8)$$

where Π^i are the canonical momenta, X^i the canonical variables, ε_{abc} is the Levi-Civita tensor, and $\nu^2 > 0$ is the mass deformation which also determines the size of the Chern-Simons deformation (the last term in (4.8)). Following [35], we will study the “mini-BMN” model with $X^\alpha = 0$, so we effectively only have three degrees of freedom. The matrices $X_{1,2,3}$ and $P_{1,2,3}$ are $N \times N$ matrices. For this example we have to abandon the master formulas for OTOC (2.2)–(2.3) as it is very difficult to find the quantities c_{mn} — for this we would have to perform exact diagonalization of the Hamiltonian (4.8). Instead, we follow [36] and write a truncated system of equations directly for the expectation values $\langle X^a \rangle$ and $\langle P^a \rangle$ and the two-point correlators $\langle X^a X^b \rangle$, $\langle \Pi^a X^b \rangle$ and $\langle \Pi^a \Pi^b \rangle$, where the expectation value is obtained through the trace over the density matrix: $\langle X^a \rangle \equiv \text{Tr}(\rho X^a)$. The equations read (for their derivation see [36]):

$$\partial_t \langle X^a \rangle = \frac{1}{N} \langle \Pi^a \rangle \\ \frac{1}{N} \partial_t \langle \Pi^a \rangle = \langle X^b \rangle \langle X^b \rangle \langle X^a \rangle - 2 \langle X^b \rangle \langle X^a \rangle \langle X^b \rangle + \langle X^a \rangle \langle X^b \rangle \langle X^b \rangle + \nu^2 \langle X^a \rangle + w \varepsilon_{abc} \langle X^b \rangle \langle X^c \rangle + \\ + \left(X^a \langle X^b X^b \rangle - \langle X^b X^b \rangle X^a + X^b \langle X^a X^b \rangle - \langle X^a X^b \rangle X^b + w \varepsilon_{abc} \langle X^b X^c \rangle \right), \quad (4.9)$$

where the last line contains the leading quantum corrections: all possible terms with a single contraction of the classical equation of motion, and the summation over repeated indices is understood. The equations of motion for the two-point correlators are obtained again by writing the classical equations of motion for the bilinears $\Pi^a \Pi^b$, $\Pi^a X^b$ and $X^a X^b$

⁸Do not confuse with the classical action considered in [33, 35].

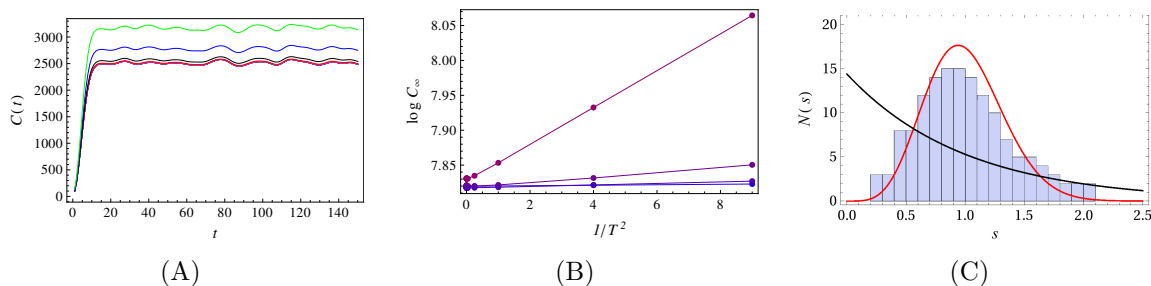


Figure 10. (A) Time dynamics $C(t)$ for the truncated quantum-mechanical mini-BMN model, with $\nu = 0.1, 0.3, 0.5, 1.0$ (red, black, blue, red), shows the expected pattern of growth followed by an oscillating plateau. (B) Temperature dependence of the saturated value $C_{\infty}(T)$ for the same values of ν (blue to red) has the same $\exp(1/T^2)$ scaling as the random matrix ensembles. The level spacing statistics, shown in (C) for $\nu = 0.5$, is indeed quite close to the Gaussian unitary ensemble (full red curve) and clearly at odds with the Poisson statistics (full black curve), confirming that this system is within the scope of our random matrix calculation.

and taking all possible single contractions in each term. This yields:

$$\partial_t \langle X^a X^b \rangle = \frac{1}{N} \left(\langle \Pi^a X^b \rangle + \langle X^a \Pi^b \rangle \right) \quad (4.10)$$

$$\begin{aligned} \partial_t \langle \Pi^a X^b \rangle = & \frac{1}{N} \langle \Pi^a \Pi^b \rangle + N \langle X^a X^b \rangle \langle X^c X^c \rangle - N \langle X^c X^c \rangle \langle X^a X^b \rangle + N \langle X^b X^c \rangle \langle X^a X^c \rangle - \\ & - N \langle X^a X^c \rangle \langle X^b X^c \rangle + \nu^2 \langle X^a X^b \rangle \end{aligned} \quad (4.11)$$

$$\begin{aligned} \partial_t \langle \Pi^a \Pi^b \rangle = & N \langle X^a X^b \rangle \langle X^c X^c \rangle - N \langle X^c X^c \rangle \langle X^a X^b \rangle + N \langle X^b X^c \rangle \langle X^a X^c \rangle - \\ & - N \langle X^a X^c \rangle \langle X^b X^c \rangle + \nu^2 \langle X^a X^b \rangle + (a \leftrightarrow b). \end{aligned} \quad (4.12)$$

As explained in [36], this truncated system is obtained by assuming a Gaussian approximation for the wavefunctions. Therefore, we solve the truncated quantum dynamics of the mini-BMN model — essentially a toy model, but it will serve our purpose. Now that we have set the stage, we can express the kinematic OTOC as $C(t) = \langle X^a \Pi^b \rangle - \langle \Pi^a X^b \rangle$ and study its dynamics. The outcome is given in figure 10. We are essentially back to the random matrix regime of section 3 — there is a clear scaling $C_{\infty} \sim \exp(1/T^2)$ (we do not see the other regime, but again it may well be there for sufficiently high temperatures), and the level distribution is a near-perfect fit to the Wigner-Dyson curve. Therefore, if a Hamiltonian is strongly chaotic, then both the level distribution and the OTOC plateau are well described by the random matrix theory.

5 Discussion and conclusions

In this paper we have formulated a somewhat unexpected indicator of quantum chaos, useful mainly in few-body (few-degrees-of-freedom) systems. While OTOC has become the quintessential object in the studies of quantum chaos and information transport, characterized mainly by its growth rate — the (quantum) Lyapunov exponent, in our examples its growth pattern tends to be quite nonuniversal and “noisy” (in the sense that it depends sensitively on the system at hand and the operators we look at). Our analytic

treatment of OTOC dynamics is quite sketchy, however both analytical and numerical results strongly suggest there is no clear exponential growth. At first glance, one might think that this finding is completely at odds with the established wisdom, however this is not true. In the literature, exponential growth is mainly characteristic for systems with a classical gravity dual (and reaches its maximum when the dual contains a thermal black hole horizon). There are abundant examples of quantum chaotic systems which do not have an exponentially growing OTOC (we especially like [19] but there are many other published examples). The exponential growth follows, in the AdS/CFT picture, from the shock wave dynamics in a classical gravity background, and need not exist when the background is not classical or when the gravity dual does not exist at all. This is precisely what happens here: the Henon-Heiles Hamiltonian is certainly nothing like a strongly coupled large N field theory, while the truncated mini-BMN model comes closer (it is actually related to discretized Yang-Mills) but we tackle it at finite N and thus away from the fast scrambling dual. For random matrices, our findings for $C(t)$ are in line with the rigorous results of [13]. As pointed out in that work, the crucial difference between random matrices and strongly coupled field theories is that the former have no notion of locality neither in time nor in space. In our small systems, the spatial locality is irrelevant anyway but if the system is not sufficiently chaotic there will still be long-term temporal correlations in dynamics (this indeed gives rise to different scaling regimes for strong and weak chaos).

On the other hand, what we have found is that the long-time OTOC behavior, when it becomes essentially stationary, with a complex oscillation pattern, is surprisingly regular — behaving as $\exp(1/T^2)$ and $\exp(1/T)$ respectively in strong and weak chaos. This indicator seems to have a stronger connection to quantum chaos in the sense of level statistics than the Lyapunov exponent; in all examples we have studied the $\exp(1/T^2)$ regime and the Wigner-Dyson level distribution go hand in hand. At very high temperatures we detect also a different regime, when the OTOC plateau grows with temperature. This regime seems less universal, and we do not understand it very well. One might think that the plateau value should not carry any useful information; it is often laconically stated that OTOC reaches saturation when the initial perturbation has spread all over the system and that this saturation value is unity when OTOC is appropriately normalized. This is roughly true, however “spreading all over the system” is not a rigorous notion — depending on the system and the operators A, B in OTOC, the perturbation may never spread completely due to symmetry constraints, specific initial conditions, quasi-integrals of motion etc. Such factors are particularly important in finite systems (quantum mechanics as opposed to quantum field theory) that we study. Looking at the figures, one sees that differences in the asymptotic OTOC value C_∞ tend to be small, and C_∞ tends to be about the same to an order of magnitude in all cases. We conjecture that such differences would dwindle to zero in the field limit.

A simple intuitive explanation for the falloff of asymptotic OTOC with temperature is the following: we expect that higher temperatures lead to faster information spreading and quicker equilibration. Therefore, it is logical that the plateau value will be lower, so that the system needs less time to reach it, i.e. it needs less time to equilibrate.

We note in passing that we have confirmed that scrambling can originate from at least two distinct mechanisms: local instability and chaos, so in the former case the relation of OTOC to chaos is largeley lost. This is a known fact in many examples already [6, 7, 9, 16, 26] and we emphasize it here merely as a reminder to the reader that the OTOC-chaos connection is really a relation of three elements: OTOC-scrambling-chaos, and if the second link is missing no attempt should be made to understand chaotic dynamics from OTOC.

We conclude with some speculations. The OTOC plateau value, as we found, is a rather universal function of temperature, and it is essentially a finite-size fluctuation of the correlation function, when the system is small enough that the relative size of fluctuations does not go to zero. We may then look for universality and the connections to chaotic dynamics in other similar quantities, e.g. the average fluctuation of the expectation value of some operator during thermalization. Such a quantity remains nonzero also in AdS/CFT at large N , and may relate our results to the more familiar fast scrambling, strongly correlated holographic systems.

Acknowledgments

This work has made use of the excellent Sci-Hub service. Work at the Institute of Physics is funded by the Ministry of Education, Science and Technological Development and by the Science Fund of the Republic of Serbia, under the Key2SM project (PROMIS program, Grant No. 6066160).

A Detailed structure and calculation of OTOC for Gaussian orthogonal ensembles

In this appendix we consider the calculation of OTOC for random matrix systems in some more detail, and describe the detailed structure of the correlation function $C(t)$. Let us first denote, for the sake of brevity:

$$\sum_{\text{tot}} \equiv \sum_{n,m} \sum_{k,k'} \sum_{i_1,i_2} \sum_{j_1,j_2} \sum_{i'_1,i'_2} \sum_{j'_1,j'_2}, \quad \mathbf{C} \equiv c_{j_1}^k \dots c_{j'_2}^m.$$

Denote also the products of matrix elements of the operators A, B entering the expression (3.4) by $\chi_1, \chi_2, \chi_3, \chi_4$. Now the expression for $\langle C(t) \rangle$ can be written as:

$$\begin{aligned} \langle C(t) \rangle = & \sum_{\text{tot}} \int P(\{c\}) d^{N^2} \{c\} \int \mathcal{P}(\{E\}) d^N E e^{-\beta E_n} \mathbf{C} \\ & \times \left(\chi_1 e^{i(E_{k'}-E_k)t} + \dots + \chi_4 e^{i(E_m+E_n-E_{k'}-E_k)t} \right) \end{aligned} \quad (\text{A.1})$$

As we have noticed in the main text, the integral over \mathbf{C} yields just a numerical constant. Let us therefore evaluate the energy integral $I_1 = \int d^N \{E\} \mathcal{P}(\{E\}) e^{-\beta E_n} \chi_1 e^{i(E_{k'}-E_k)t}$. We have:

$$I_1 = \int \int \dots \int dE_n \int dE_k \int dE_{k'} \mathcal{P}(\{E\}) e^{-\beta E_n} \times \chi_1 e^{i(E_{k'}-E_k)t}. \quad (\text{A.2})$$

The absolute values of the differences can be written out in the obvious way:

$$\prod_{n < m} |E_n - E_m| = \sum_i (-1)^{\pi(i)} E_1^{\alpha_{1,i}} E_2^{\alpha_{2,i}} \dots E_N^{\alpha_{N,i}}, \quad (\text{A.3})$$

where all $\alpha_{i,j}$ are some (positive) integer exponents and $\pi(i)$ is the appropriate sign factor 0 or 1. Therefore, I_1 can be reorganized as:

$$I_1 = \sum_j \chi_1 \prod_{i \neq n, k, k'} \int E_i^{\alpha_{i,j}} e^{-E_i^2} dE_i \int dE_n e^{-\beta E_n} e^{-E_n^2} \int dE_k e^{iE_{k'} t} e^{-E_{k'}^2} \int dE_k e^{-iE_k t} e^{-E_k^2}. \quad (\text{A.4})$$

Note that the part $\prod_{n < m} |E_n - E_m|$, is not essential for the general behavior, since the singular integral $\int E_i^{\alpha_{i,j}} e^{-E_i^2} dE_i$ is either some constant (if α is even), or zero if α is odd. Otherwise for $i \neq j$

$$\int \prod_{l < i < l'} E_i e^{-E_j^2} dE_j = \text{const.} \times \prod_{l < i < l'} E_i. \quad (\text{A.5})$$

Therefore, we only focus on calculating integrals of the form

$$\int dE_n \int dE_k \int dE_{k'} e^{-\beta E_n} \times e^{-E_n^2 - E_k^2 - E_{k'}^2} \chi_1 e^{i(E_{k'} - E_k)t}, \quad (\text{A.6})$$

which yields the closed-form expression for the temperature dependence of I_1 :

$$I_1 \sim \delta_{k,k'} e^{\beta^2/4} + (1 - \delta_{k,k'}) e^{\beta^2/4} e^{-t^2/2}, \quad (\text{A.7})$$

where $\delta_{k,k'}$ is the Kronecker delta, reminding us that the main contribution comes from the terms with $E_k = E_{k'}$ which generically means $k = k'$. It is clear that a similar calculation holds for the other parts of $\langle C(t) \rangle$. This produces the temperature scaling found in the main text for random matrices, of the form $\langle C(t) \rangle \sim e^{1/4T^2}$. But the time dependence is more complicated. In order to see this, we look at the structure of the polynomial factors in I_1 in some more detail. We see immediately that $\langle C(t) \rangle$ will also have dependence on t^{2n}, β^n . Start from

$$\int E_i^{\alpha_i} e^{-iE_i t} e^{-E_i^2} dE_i = e^{-t^2/4} \int (u - it/2)^{\alpha_i} e^{-u^2} du, \quad (\text{A.8})$$

where $E_i = u - it/2$. Let us look at two cases: α_i even and α_i odd. For any α_i the polynomial will have the form:

$$(u - it/2)^{\alpha_i} = \sum_{j=0}^{\alpha_i} \gamma_j u^j (it/2)^{\alpha_i - j}. \quad (\text{A.9})$$

Assume first that α_i is even. This means that j and $\alpha_i - j$ are of same parity. For even j the Gaussian integral evaluates to some constant, but we will also have the prefactor of $(it/2)^{\alpha_i - j}$, for all even $j \leq \alpha_i$. The odd powers (j odd) will disappear because of the symmetric domain of integration. For α_i odd, j and $\alpha_i - j$ will be of different parity so again, only even j give a nonzero integral. In conclusion, the integral (A.6) with polynomial prefactors included will have the form:

$$\int E_i^{\alpha_i} e^{-iE_i t} e^{-E_i^2} dE_i = e^{-t^2/4} Q(t^{2n}), \quad (\text{A.10})$$

where $Q(t^{2n})$ is a real polynomial depending on even powers of t , and $2n \leq \alpha_i$. Alternatively, for α_i odd, we get:

$$\int E_i^{\alpha_i} e^{-iE_i t} e^{-E_i^2} dE_i = i e^{-t^2/4} R(t^{2n+1}), \quad (\text{A.11})$$

where $R(t^{2n+1})$ is a real polynomial depending on odd powers of t , and $2n + 1 \leq \alpha_i$. Analogous logic holds for the β dependence. Now we look back at I_1 :

$$I_1 = \text{const.} \int E_n^{\alpha_n} dE_n \int E_k^{\alpha_k} dE_k \int E_{k'}^{\alpha_{k'}} dE_{k'} e^{-\beta E_n} \times e^{-E_n^2 - E_k^2 - E_{k'}^2} \chi_1 e^{i(E_{k'} - E_k)t}. \quad (\text{A.12})$$

When we write out the products of energies, we have the following types of monomials in the resulting polynomial:

1. QQU
2. QRU
3. QQV
4. QRV ,

with the prefactor $\delta_{k,k'} e^{-t^2/2} e^{\beta^2/4}$. Here, Q, R are polynomials of t and are U/V are polynomials of even/odd powers of β respectively. Note however that QR and RQ give the same structure after integration.

The other integral appearing when writing out the master formula for OTOC is

$$K_n = \int E_n^{\alpha_n} e^{-\beta E_n} e^{-iE_n t} e^{-E_n^2} dE_n. \quad (\text{A.13})$$

According to the same logic as for I_1 , it is not hard to get the equivalent form of K_n (leaving out the exponentially decaying terms):

$$K_n = e^{\beta^2/4} e^{i\beta t/2} \int \sum_j \gamma_j u^j (\beta/2 + it/2)^{\alpha_n - j} e^{-u^2} du. \quad (\text{A.14})$$

Now we will use the fact that OTOC is a real function, as we can see also from the definition (2.2). Therefore, all imaginary parts must vanish. From this fact we reach a few important conclusions:

1. In the structure of I_1 , the combination QR is impossible, thus we will only have polynomials of t with an even exponent, and no restriction for polynomials of β as it is a real integral, and no term has to vanish.
2. In the structure of K_n , when we have the factor $\cos(\beta t/2)$, only even powers of t and arbitrary powers of β can survive.
3. In the structure of K_n when we have the factor $\sin(\beta t/2)$, only odd powers of t and arbitrary powers of β can survive.

The conclusion of the above analysis gives us a rough idea of what the $\langle C(t) \rangle$ looks like:

$$\langle C(t) \rangle = e^{\frac{\beta^2}{4}} W_0(\sigma\beta) + e^{\frac{\beta^2}{4}} \left(\cos\left(\frac{\beta t}{2}\right) Q(t^{2n}) W_1(\sigma\beta) + \sin\left(\frac{\beta t}{2}\right) R(t^{2n+1}) W_2(\beta^n) \right), \tag{A.15}$$

where W_0, W_1, W_2 are arbitrary polynomials of β . This is the form found also in the main text, with the exception that in the main text we have rescaled the combination βt as $\beta t/\sigma^2$ in order to have a dimensionless expression.

A.1 The large matrix limit

In the limit $N \rightarrow \infty$ we can say more on the structure of OTOC. We can first schematically rewrite (A.15) together with any exponentially suppressed corrections as

$$\langle C(t) \rangle = e^{\frac{\beta^2}{4}} Q(t^{2n}) W(\sigma\beta) \left(L_1 + L_2 e^{-t^2/2} \right). \tag{A.16}$$

Here we have first absorbed all time and β dependence of (A.15) into the functions Q and W respectively, and then we have included the exponentially suppressed correction coming from the $k \neq k'$ terms in the integrals I_1 and K_n . By L_1, L_2 we denote the constant (time- and temperature-independent) factors. In general one can write L_1 as

$$L_1 = \sum_{j=1}^N \sum_{i=0}^j c_i \binom{j}{i} \tag{A.17}$$

We can easily estimate the second sum. Namely,

$$\left(\sum_{i=0}^j c_i \binom{j}{i} \right)^2 \leq \left(\sum_{i=0}^j c_i^2 \right) \left(\sum_{i=0}^j \binom{j}{i}^2 \right), \tag{A.18}$$

by the Cauchy-Schwarz-Bunyakovski inequality. Next, the well known formula $\sum_{i=0}^j \binom{j}{i}^2 = \binom{2j}{j}$ yields

$$\sum_{i=0}^j c_i \binom{j}{i} \leq \text{const.} \times \sqrt{\binom{2j}{j}}. \tag{A.19}$$

To get rid of the binomial coefficient we will use the Stirling's formula and get

$$\sqrt{\binom{2j}{j}} = \sqrt{\frac{(2j)!}{j!j!}} \approx \sqrt{\frac{\sqrt{4\pi j} \frac{(2j)^{2j}}{e^{2j}}}{2\pi j \frac{(j^j)^2}{(e^j)^2}}} \approx \text{const.} \times \frac{2^j}{j^{1/4}}. \tag{A.20}$$

Finally we reach the result:

$$L_1 \approx \text{const.} \times \sum_{j=1}^N \frac{2^j}{j^{1/4}} \approx \text{const.} \times \frac{2^{N+1}}{N^{1/4}}, \tag{A.21}$$

for $N \rightarrow \infty$. Exactly the same logic goes for L_2 .

In the large matrix limit it is possible to show explicitly what we know has to happen: OTOC reaches a plateau. Looking at (A.16), the condition to reach the plateau for times longer than some scale t_0 is

$$e^{-\frac{t^2}{2}} Q(t^{2n}) (L_1 + L_2 e^{-t^2/2}) = \text{const.} \quad t > t_0. \quad (\text{A.22})$$

It is more convenient to look at the forms given in (A.15). First let us look at the condition $Q(t^{2n}) = \text{const.} \times e^{t^2/2}$. The exponential term can be represented as a series; equating it with $Q(t^{2n})$ we get

$$\sum_j \alpha_j t^{2j} = \text{const.} \times \sum_j \frac{t^{2j}}{2^j j!}, \quad (\text{A.23})$$

thus, we need $\alpha_j \sim \frac{1}{2^j j!}$, which we know is the case from (A.9). For the second term the situation is similar:

$$\sum_j \beta_j t^{2j} = \text{const.} \times \sum_j \frac{t^{2j}}{j!}, \quad (\text{A.24})$$

so we need to have $\beta_j \sim \frac{1}{j!}$; this is true by $\cos(\beta t/2) = Q(t^{2n})W(\beta^{2n})$ and $\sin(\beta t/2) = Q(t^{2n+1})W(\beta^{2n+1})$, since the terms in the Taylor expansions of the left-hand sides behave as $\sim \frac{1}{j!}$.

We can also look at the opposite limit in which $t \rightarrow 0$. Let us rearrange (A.16):

$$\langle C(t) \rangle = L'_1 Q(t^{2n}) e^{-t^2/2} + L'_2 Q(t^{2n}) e^{-t^2}. \quad (\text{A.25})$$

Now, simply using the definition of Q and expanding into a series we get:

$$\langle C(t) \rangle = L'_1 \left(1 - \frac{t^2}{2} + o(t^4) \right) (q_0 + q_1 t^2) + L'_2 \left(1 - t^2 + o(t^4) \right) (q_0 + q_1 t^2). \quad (\text{A.26})$$

After some algebra we get:

$$\langle C(t) \rangle = Q_0 + Q_1 t^2 + Q_2 t^4 + o(t^4) = P(t). \quad (\text{A.27})$$

We see now that OTOC behaves in a very simple way for early times; this expansion is also consistent with the result (3.16) of [13].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [[arXiv:1503.01409](https://arxiv.org/abs/1503.01409)] [[INSPIRE](https://inspirehep.net/literature/1503014)].
- [2] S.H. Shenker and D. Stanford, *Stringy effects in scrambling*, *JHEP* **05** (2015) 132 [[arXiv:1412.6087](https://arxiv.org/abs/1412.6087)] [[INSPIRE](https://inspirehep.net/literature/1412608)].
- [3] S. Kundu, *Extremal chaos*, *JHEP* **01** (2022) 163 [[arXiv:2109.08693](https://arxiv.org/abs/2109.08693)] [[INSPIRE](https://inspirehep.net/literature/2109086)].

- [4] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065 [[arXiv:0808.2096](#)] [[INSPIRE](#)].
- [5] N. Lashkari, D. Stanford, M. Hastings, T. Osborne and P. Hayden, *Towards the Fast Scrambling Conjecture*, *JHEP* **04** (2013) 022 [[arXiv:1111.6580](#)] [[INSPIRE](#)].
- [6] E.B. Rozenbaum, L.A. Bunimovich and V. Galitski, *Early-Time Exponential Instabilities in Nonchaotic Quantum Systems*, *Phys. Rev. Lett.* **125** (2020) 014101 [[arXiv:1902.05466](#)] [[INSPIRE](#)].
- [7] T. Xu, T. Scaffidi and X. Cao, *Does scrambling equal chaos?*, *Phys. Rev. Lett.* **124** (2020) 140602 [[arXiv:1912.11063](#)] [[INSPIRE](#)].
- [8] K. Hashimoto, K. Murata and R. Yoshii, *Out-of-time-order correlators in quantum mechanics*, *JHEP* **10** (2017) 138 [[arXiv:1703.09435](#)] [[INSPIRE](#)].
- [9] K. Hashimoto, K.-B. Huh, K.-Y. Kim and R. Watanabe, *Exponential growth of out-of-time-order correlator without chaos: inverted harmonic oscillator*, *JHEP* **11** (2020) 068 [[arXiv:2007.04746](#)] [[INSPIRE](#)].
- [10] F. Haake, S. Gnutzman and M. Kuś, *Quantum signatures of chaos*, Springer-Verlag, Berlin, Germany (2019).
- [11] M.L. Mehta, *Random matrices*, Academic, New York, U.S.A. (2004).
- [12] J.S. Cotler et al., *Black Holes and Random Matrices*, *JHEP* **05** (2017) 118 [*Erratum ibid.* **09** (2018) 002] [[arXiv:1611.04650](#)] [[INSPIRE](#)].
- [13] J. Cotler, N. Hunter-Jones, J. Liu and B. Yoshida, *Chaos, Complexity, and Random Matrices*, *JHEP* **11** (2017) 048 [[arXiv:1706.05400](#)] [[INSPIRE](#)].
- [14] J. Cotler and N. Hunter-Jones, *Spectral decoupling in many-body quantum chaos*, *JHEP* **12** (2020) 205 [[arXiv:1911.02026](#)] [[INSPIRE](#)].
- [15] W. Kirkby, D.H.J. O’Dell and J. Mumford, *False signals of chaos from quantum probes*, [arXiv:2108.09391](#) [[INSPIRE](#)].
- [16] T. Akutagawa, K. Hashimoto, T. Sasaki and R. Watanabe, *Out-of-time-order correlator in coupled harmonic oscillators*, *JHEP* **08** (2020) 013 [[arXiv:2004.04381](#)] [[INSPIRE](#)].
- [17] A. Bhattacharyya, W. Chemissany, S.S. Haque, J. Murugan and B. Yan, *The Multi-faceted Inverted Harmonic Oscillator: Chaos and Complexity*, *SciPost Phys. Core* **4** (2021) 002 [[arXiv:2007.01232](#)] [[INSPIRE](#)].
- [18] W. Kłobus, P. Kurzyński, M. Kuś, W. Laskowski, R. Przybycień and K. Życzkowski, *Transition from order to chaos in reduced quantum dynamics*, *Phys. Rev. E* **105** (2022) 034201 [[arXiv:2111.13477](#)] [[INSPIRE](#)].
- [19] B. Bertini, P. Kos and T. Prosen, *Exact Spectral Form Factor in a Minimal Model of Many-Body Quantum Chaos*, *Phys. Rev. Lett.* **121** (2018) 264101 [[arXiv:1805.00931](#)] [[INSPIRE](#)].
- [20] S. Pappalardi, A. Russomanno, B. Žunkovič, F. Iemini, A. Silva and R. Fazio, *Scrambling and entanglement spreading in long-range spin chains*, *Phys. Rev. B* **98** (2018) 134303 [[arXiv:1806.00022](#)] [[INSPIRE](#)].
- [21] K. Okuyama, *Spectral form factor and semi-circle law in the time direction*, *JHEP* **02** (2019) 161 [[arXiv:1811.09988](#)] [[INSPIRE](#)].

- [22] Q. Hummel, B. Geiger, J.D. Urbina and K. Richter, *Reversible quantum information spreading in many-body systems near criticality*, *Phys. Rev. Lett.* **123** (2019) 160401 [[arXiv:1812.09237](#)] [[INSPIRE](#)].
- [23] H. Goto and T. Kanao, *Chaos in coupled Kerr-nonlinear parametric oscillators*, *Phys. Rev. Res.* **3** (2021) 043196 [[arXiv:2110.04019](#)] [[INSPIRE](#)].
- [24] N.D. Varikuti and V. Madhok, *Out-of-time ordered correlators in kicked coupled tops and the role of conserved quantities in information scrambling*, [arXiv:2201.05789](#) [[INSPIRE](#)].
- [25] A.V. Kirkova, D. Porras and P.A. Ivanov, *Out-of-time-order correlator in the quantum Rabi model*, *Phys. Rev. A* **105** (2022) 032444 [[arXiv:2201.06340](#)] [[INSPIRE](#)].
- [26] J.R.G. Alonso, N. Shammah, S. Ahmed, F. Nori and J. Dressel, *Diagnosing quantum chaos with out-of-time-ordered-correlator quasiprobability in the kicked-top model*, [arXiv:2201.08175](#) [[INSPIRE](#)].
- [27] K. Hashimoto, K. Murata, N. Tanahashi and R. Watanabe, *A bound on energy dependence of chaos*, [arXiv:2112.11163](#) [[INSPIRE](#)].
- [28] J. Aguirre, J.C. Vallejo and M.A.F. Sanjuán, *Wada basins and chaotic invariant sets in the Hénon-Heiles system*, *Phys. Rev. E* **64** (2001) 066208.
- [29] H.E. Kandrup, C. Siopis, G. Contopoulos and R. Dvorak, *Diffusion and scaling in escapes from two-degree-of-freedom Hamiltonian systems*, *Chaos* **9** (1999) 381 [[astro-ph/9904046](#)] [[INSPIRE](#)].
- [30] D.E. Berenstein, J.M. Maldacena and H.S. Nastase, *Strings in flat space and pp waves from $N = 4$ super Yang-Mills*, *JHEP* **04** (2002) 013 [[hep-th/0202021](#)] [[INSPIRE](#)].
- [31] T. Banks, W. Fischler, S.H. Shenker and L. Susskind, *M theory as a matrix model: A Conjecture*, *Phys. Rev. D* **55** (1997) 5112 [[hep-th/9610043](#)] [[INSPIRE](#)].
- [32] K.L. Zarembo and Y.M. Makeenko, *An introduction to matrix superstring models*, *Phys. Usp.* **41** (1998) 1 [[INSPIRE](#)].
- [33] Y. Asano, D. Kawai and K. Yoshida, *Chaos in the BMN matrix model*, *JHEP* **06** (2015) 191 [[arXiv:1503.04594](#)] [[INSPIRE](#)].
- [34] G. Gur-Ari, M. Hanada and S.H. Shenker, *Chaos in Classical D0-Brane Mechanics*, *JHEP* **02** (2016) 091 [[arXiv:1512.00019](#)] [[INSPIRE](#)].
- [35] X. Han and S.A. Hartnoll, *Deep Quantum Geometry of Matrices*, *Phys. Rev. X* **10** (2020) 011069 [[arXiv:1906.08781](#)] [[INSPIRE](#)].
- [36] P.V. Buividovich, M. Hanada and A. Schäfer, *Quantum chaos, thermalization, and entanglement generation in real-time simulations of the Banks-Fischler-Shenker-Susskind matrix model*, *Phys. Rev. D* **99** (2019) 046011 [[arXiv:1810.03378](#)] [[INSPIRE](#)].

The bound on chaos for closed strings in Anti-de Sitter black hole backgrounds

Mihailo Čubrović

*Scientific Computing Lab, Center for the Study of Complex Systems,
Institute of Physics Belgrade, University of Belgrade,
Pregrevica 118, 11080 Belgrade, Serbia*

E-mail: mcubrovic@gmail.com

ABSTRACT: We perform a systematic study of the maximum Lyapunov exponent values λ for the motion of classical closed strings in Anti-de Sitter black hole geometries with spherical, planar and hyperbolic horizons. Analytical estimates from the linearized variational equations together with numerical integrations predict the bulk Lyapunov exponent value as $\lambda \approx 2\pi T n$, where n is the winding number of the string. The celebrated bound on chaos stating that $\lambda \leq 2\pi T$ is thus systematically modified for winding strings in the bulk. Within gauge/string duality, such strings apparently correspond to complicated operators which either do not move on Regge trajectories, or move on subleading trajectories with an unusual slope. Depending on the energy scale, the out-of-time-ordered correlation functions of these operators may still obey the bound $2\pi T$, or they may violate it like the bulk exponent. We do not know exactly why the bound on chaos can be modified but the indication from the gauge/string dual viewpoint is that the correlation functions of the dual gauge operators never factorize and thus the original derivation of the bound on chaos does not apply.

KEYWORDS: Gauge-gravity correspondence, AdS-CFT Correspondence, Integrable Field Theories, Black Holes

ARXIV EPRINT: [1904.06295](https://arxiv.org/abs/1904.06295)

Contents

1	Introduction	1
2	String dynamics in static black hole backgrounds	5
2.1	Fixed points and near-horizon dynamics	7
3	Lyapunov exponents and the bound on chaos	10
3.1	Variational equations and analytical estimates of Lyapunov exponents	11
3.1.1	Thermal horizon	11
3.1.2	Away from the horizon	12
3.1.3	Extremal horizon	13
3.1.4	Lyapunov time versus event time	14
3.1.5	Dimensionful constants	14
3.2	Numerical checks	15
4	Toward a physical interpretation of the modified bound	15
4.1	Dual gauge theory interpretation	15
4.1.1	Operators dual to a ring string?	16
4.1.2	Planetoid string	18
4.2	The limits of quasiclassicality	20
4.3	Ring string scattering amplitude and the relation to OTOC	20
4.3.1	Eikonal approximation	21
4.3.2	Beyond the eikonal approximation: waves on the string	23
5	Discussion and conclusions	25
A	Summary of the numerics	28

1 Introduction

Sharp results like inequalities and no-go theorems are often the cornerstones of our understanding of physical phenomena. Besides being appealing and captivating, they are easy to test as they provide a sharp prediction on a certain quantity, and we can often learn a lot by understanding the cases when such bounds need to be generalized or abandoned. The upper bound on the Lyapunov exponent (the rate of the growth of chaos), derived in [1] inspired by hints found in several earlier works [2–7], is an example of such a result, which is related to the dynamics of nonstationary correlation functions and provides insight into the deep and important problem of thermalization and mixing in strongly coupled systems.

It is clear, as discussed also in the original paper [1], that there are cases when the bound does not apply: mainly systems in which the correlation functions do not factorize even at arbitrarily long times, and also systems without a clear separation of short timescales (or collision times) and long timescales (or scrambling times). A concrete example of bound violation was found in [8] for a semiclassical system with a conserved angular momentum (inspired by the Sachdev-Ye-Kitaev (SYK) model [9–12]) and in [13], again for a SYK-inspired system. In the former case, the reason is clear: the orbits that violate the bound are precisely those that cannot be treated semiclassically, so the violation just signals that the model used becomes inaccurate; in the latter case things are more complicated and the exact reason is not known. Finally, in [14] systematic higher-order quantum corrections to the bound are considered. The bound is in any case a very useful benchmark, which can tell us something on long-term dynamics of the system at hand, i.e. if some bound-violating mechanisms are at work or not.

Although the bound on chaos is mainly formulated for field theories in flat spacetime, it has an intimate connection to gravity: the prediction is that fields with gravity duals saturate the bound. This makes dynamics in asymptotically anti-de Sitter (AdS) spacetimes with a black hole particularly interesting: they have a field theory dual,¹ and black holes are conjectured to be the fastest scramblers in nature [2, 3], i.e., they minimize the time for the overlap between the initial and current state to drop by an order of magnitude. Some tests of the bound for the motion of particles in the backgrounds of AdS black holes and an additional external potential were already made [15]; the authors find that the bound is systematically modified for particles hovering at the horizon and interacting with *higher spin* external fields. When the external field becomes scalar, the exact bound by Maldacena, Shenker and Stanford is recovered (as shown also in [16]).

The idea of this paper is to study the bound on chaos in the context of *motion of strings in AdS black hole geometries*. Asymptotically AdS geometry is helpful not only because of the gauge/gravity duality, but also for another reason: AdS asymptotics provide a regulator, i.e., put the system in a box, making its dynamics more interesting (in asymptotically flat space, most orbits immediately escape to infinity with no opportunity to develop chaos). Now why consider strings instead of geodesics? Because geodesics are not the best way to probe the chaos generated by black holes: we know that geodesics in AdS-Schwarzschild, AdS-Reissner-Nordstrom and AdS-Kerr backgrounds (and also in all axially symmetric and static black hole geometries) are integrable, and yet, since the horizon in all these cases has a finite Hawking temperature, there should be some thermalization and chaos going on. The logical decision is therefore to go for string dynamics, which is practically always nonintegrable in the presence of a black hole. We look mainly at the Lyapunov exponents and how they depend on the Hawking temperature. We will see that the bound of [1] is surprisingly relevant here, even though the bound was formulated for *field theories* with a *classical gravity dual*, whereas we look at the *bulk* dynamics of *strings*, which go beyond the realm of Einstein gravity. At first glance, their Lyapunov exponents should not saturate (let alone violate) the bound; in fact, at first glance, it is not

¹Of course, one should be careful when it comes to details; it is known that for some field contents in the bulk the boundary theory does not exist.

obvious at all how to relate the Lyapunov exponent of classical bulk orbits to the result [1], which defines the Lyapunov exponent in terms of the out-of-time ordered correlation functions (OTOC).² An important discovery in relation to this issue was made in [17], where the authors consider a holographically more realistic string (open string dual to a quark in Brownian motion in a heat bath), compute the Lyapunov exponent in dual field theory, and find that it exactly saturates the bound. However, their *world-sheet* theory, i.e., their induced metric itself looks somewhat like gravity on AdS_2 ; therefore, close connection to the Einstein gravity result is understandable. Our situation is different not only because the ring string configurations have worldsheet actions very different from Einstein gravity but also because we look mainly at the Lyapunov exponents of the bulk orbits.³ We will eventually look also at the OTOC in dual field theory and find that the “quantum” Lyapunov exponents do not in general coincide with the classical bulk values. However, the subject of OTOC functions is more complicated as it requires one to consider the backreaction on the background, and studying the behavior of the ring string in such backreacted geometry is in general more difficult than for the open string of [17]. Therefore, we mostly leave the OTOC and quantum Lyapunov exponent for future work.

At this point we come to another question, distinct but certainly related to the chaos bound: the story of (non)integrability in various curved spacetimes. For point particles (i.e., motion on geodesics) it is usually not so difficult to check for integrability, and symmetries of the problem usually make the answer relatively easy. However, integrability in string theory remains a difficult topic. Most systematic work was done for top-down backgrounds, usually based on the differential Galois theory whose application for string integrability was pioneered in [19]. Systematic study for various top-down configurations was continued in [20–22]; [21] in particular provides the results for strings in a broad class of brane backgrounds, including Dp-brane, NS1 and NS5 brane configurations. The bottom line is that integrable systems are few and far apart, as could be expected. Certainly, $\text{AdS}_5 \times S^5$ is an integrable geometry, as could be expected from its duality to the (integrable) supersymmetric Yang-Mills field theory. In fact, direct product of AdS space and a sphere is integrable in any dimension, which is obvious from the separability of the coordinates. But already a marginal deformation destroys integrability; a specific example was found analytically and numerically in [23], for the β -deformation of super-Yang-Mills and its top-down dual. More information can be found, e.g., in the review [24].

The first study of integrability in a black hole background was [25], where the nonintegrability of string motion in asymptotically flat Schwarzschild black hole background was shown. In [26] the first study for an AdS black hole background (AdS-Schwarzschild) was performed, putting the problem also in the context of AdS/CFT correspondence. In [27] the work on top-down backgrounds was started, considering the strings

²In addition, the scrambling concept of [2, 4–7] is more complex; it is about the equilibration of the black hole and its environment after something falls in. In other words, it necessarily includes the perturbation of the black hole itself. We do not take into account any backreaction so we cannot compute the scrambling time, only the Lyapunov exponent.

³Another example where the bound is modified (by a factor of 2) in a theory that goes beyond Einstein gravity is [18].

on the $\text{AdS} \times T^{1,1}$ geometry generated in a self-consistent top-down way. For the top-down AdS-Sasaki-Einstein background the nonintegrability was proven analytically [19]. Finally, AdS-soliton and AdS-Reissner-Nordstrom were also found to be nonintegrable in [28, 29]. So most well-known in AdS/CFT have nonintegrable string dynamics: AdS-Schwarzschild, AdS-Reissner-Nordstrom, AdS soliton and AdS-Sasaki-Einstein.⁴ Other results on (non)integrability can be found in [30–33]; the list is not exhaustive.

Apart from the usual spherical static black holes (neutral and charged), we consider also non-spherical horizons with constant curvature. Among them are also the zero-curvature black branes, with infinite planar horizons, which are most popular in applied holography. But it is known that more general horizons can be embedded in AdS space (in general not in Minkowski space). Such black holes are usually called topological black holes, first constructed in [34–37] and generalized in [38]. The term topological is in fact partly misleading, as the backgrounds considered in some of the original papers [35] and also in our paper are not necessarily of higher topological genus: besides spherical and planar horizons, we mainly consider an infinite, topologically trivial hyperbolic horizon with constant negative curvature (pseudosphere).⁵

The reader might wonder how important the non-spherical black holes are from the physical viewpoint. In fact, as shown in the aforementioned references, they arise naturally in spaces with negative cosmological constant, i.e., in AdS spaces, for example in the collapse of dust [39], and the topological versions are easily obtained through suitable gluings (identifications of points on the orbit of some discrete subgroup of the total symmetry group) of the planar or pseudospherical horizon. Another mechanism is considered in [34], where topological black holes are pair-created from instanton solutions of the cosmological C-metric (describing a pair of black holes moving with uniform acceleration). More modern work on constant-curvature black holes and some generalizations can be found in [40–42], and AdS/CFT correspondence was applied to topological black holes in [43]. But our main motivation for considering non-spherical black holes is methodological, to maximally stretch the testing ground for the chaos bound and to gain insight into various chaos-generating mechanisms. In hindsight, we find that hyperbolic are roughly speaking most chaotic, because moving on a manifold of negative curvature provides an additional chaos-generating mechanism, in addition to the black hole.

The plan of the paper is the following. In the next section we write down the equations of motion for a closed string in static black hole background, inspect the system analytically and numerically and show that dynamics is generically non-integrable. In the third section we compute the Lyapunov exponents numerically and estimate them analytically, formulating a generalized bound in terms of the local temperature and the string winding number. The fourth section is a rather speculative attempt to put our results in the context of the dual field theory and the derivation of the original bound from [1]; we will also try to clarify the relation of the bulk classical Lyapunov exponent to the decay rates of OTOC functions in dual field theory. The last section sums up the conclusions.

⁴In [26, 29] it was shown that Reissner-Nordstrom black holes in asymptotically flat space are also nonintegrable.

⁵In fact, constant-curvature black holes would be a more suitable term than topological black holes.

2 String dynamics in static black hole backgrounds

A constant curvature black hole in $N + 1$ spacetime dimensions is a geometry of constant curvature with the metric [34–36]

$$\begin{aligned}
 ds^2 &= -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\sigma_{N-1}^2 \\
 f(r) &= r^2 + k - \frac{2m}{r^{N-2}} + \frac{q^2}{r^{2N-4}},
 \end{aligned}
 \tag{2.1}$$

where $d\sigma_{N-1}^2$ is the horizon manifold, which has curvature k , and m and q define the mass and charge of the black hole. It is a vacuum solution of the Einstein equations with constant negative cosmological constant and thus interpolates to AdS space with radius 1. From now on let us stick to $N = 3$ unless specified otherwise. For $k = 1$ we have the familiar spherical black hole. For $k = 0$ we get the planar horizon (black brane) popular in AdS/CFT applications.⁶ Finally, for $k = -1$ the horizon is an infinite hyperbolic sheet (pseudosphere), with the symmetry group $SO(2, 1)$.⁷ Notice that k can always be rescaled together with the coordinates on σ_2 thus we only consider $k = -1, 0, 1$. The metric of the horizon surface takes the form

$$d\sigma_2^2 = d\phi_1^2 + \text{sink}^2 \phi_1 d\phi_2^2, \tag{2.2}$$

with $\text{sink}(x) = \sin x$ for $k = 1$, $\text{sink}(x) = x$ for $k = 0$ and $\text{sink}(x) = \sinh(x)$ for $k = -1$.

A closed string with tension $1/\alpha'$ on the worldsheet (τ, σ) with target space X^μ and the metric $G_{\mu\nu}$ is described by the Polyakov action:

$$S = -\frac{1}{2\pi\alpha'} \int d\tau d\sigma \sqrt{-h} h^{ab} G_{\mu\nu}(X) \partial_a X^\mu \partial_b X^\nu + \epsilon^{ab} B_{\mu\nu}(X) \partial_a X^\mu \partial_b X^\nu. \tag{2.3}$$

In our black hole backgrounds we always have $B_{\mu\nu} = 0$ so we can pick the gauge $h^{ab} = \eta^{ab} = \text{diag}(-1, 1)$. This gives the Virasoro constraints

$$G_{\mu\nu} \left(\dot{X}^\mu \dot{X}^\nu + X'^\mu X'^\nu \right) = 0, \quad G_{\mu\nu} \dot{X}^\mu X'^\nu = 0, \tag{2.4}$$

where we introduce the notation $\dot{X} \equiv \partial_\tau X$, $X' \equiv \partial_\sigma X$. The first constraint is the Hamiltonian constraint $H = 0$. We consider closed strings, so $0 \leq \sigma \leq 2\pi$. From the second constraint the following ansatz is consistent (of course, it is not the only one possible):

$$\mathcal{T} = \mathcal{T}(\tau), \quad R = R(\tau), \quad \Phi_1 = \Phi_1(\tau), \quad \Phi_2 = n\sigma. \tag{2.5}$$

We denote the (dynamical) target-space coordinates $X_\mu(\tau, \sigma)$ by capital letters $\mathcal{T}, R, \Phi_1, \Phi_2$, to differentiate them from the notation for spacetime coordinates t, r, ϕ_1, ϕ_2 in the metric (2.1). The form (2.5) was tried in most papers exploring the integrability and chaos

⁶With periodic identifications on σ_2 one gets instead a toroidal horizon.

⁷If we identify the points along the orbits of the little group of $SO(2, 1)$, we get a genus g surface with $g \leq 2$, and the horizon becomes compact and topologically nontrivial, hence the term topological black holes for this case.

of strings [19, 25–29]. It is not an arbitrary ansatz: the winding of Φ_2 follows from the equations of motion, i.e., from the fact that Φ_2 is a cyclic coordinate, leading to the solution $\ddot{\Phi}_2 = 0$. Since Φ_2 has trivial dynamics, from now on we will denote $\Phi \equiv \Phi_1$. The equations of motion follow from (2.3):

$$\partial_\tau (f\dot{\mathcal{T}}) = 0 \Rightarrow E \equiv f\dot{\mathcal{T}} = \text{const.} \tag{2.6}$$

$$\ddot{R} + \frac{f'}{2f}(E^2 - \dot{R}^2) + fR (\dot{\Phi}^2 - n^2 \text{sink}^2 \Phi) = 0 \tag{2.7}$$

$$\ddot{\Phi} + \frac{2\dot{R}}{R}\dot{\Phi} + \frac{n^2}{2}\text{sink}(2\Phi) = 0. \tag{2.8}$$

Clearly, the stationarity of the metric yields the first integral E with the informal meaning of mechanical energy for the motion along the R and Φ coordinates (it is not the total energy in the strict sense). The system is more transparent in Hamiltonian form, with the canonical momenta $P_{\mathcal{T}} = -E = -f\dot{\mathcal{T}}$, $P_R = \dot{R}/f$, $P_\Phi = R^2\dot{\Phi}$:⁸

$$H = \frac{f}{2}P_R^2 + \frac{1}{2R^2}P_\Phi^2 + \frac{n^2}{2}R^2\text{sink}^2\Phi - \frac{E^2}{2f} = 0, \tag{2.9}$$

the second equality being the Virasoro constraint. We thus have a 2-degrees-of-freedom system (due to the integral of motion E , i.e., the cyclic coordinate \mathcal{T}), with a constraint, effectively giving a 1.5-degrees-of-freedom system, moving on a three-dimensional manifold in the phase space (R, P_R, Φ, P_Φ) . Notice that the motion along a geodesic is obtained for $n = 0$; in this case, the system is trivially separable and becomes just motion in a central potential. For nonzero n , the Hamiltonian (2.9) is not separable and the system is nonintegrable.⁹ On the other hand, for a point particle all constant-curvature black holes have a full set of integrals of motion leading to the integrability of geodesics: for the sphere, the additional integrals (besides E) are L^2 and L_z from $\text{SO}(3)$, and for the pseudosphere these are K^2 and K_z from $\text{SO}(2,1)$. For the planar black hole we obviously have $P_{x,y}$, the momenta, as the integrals of motion. Of course, if we consider compactified surfaces, the symmetries become discrete and do not yield integrals of motion anymore. Therefore, truly topological black holes are in general nonintegrable even for geodesics.¹⁰

⁸In this and the next section we put $\alpha' = 1/\pi$, as we only consider classical equations of motion, which are independent of α' . In section 4, when calculating the quantities of the dual gauge theory, we restore α' as it is related to the 't Hooft coupling, a physical quantity.

⁹One can prove within Picard-Vessiot theory that no canonical transformation exists that would yield a separable Hamiltonian, so the system is nonintegrable. We will not derive the proof here, as it is not very instructive; the nonintegrability of the spherical case was already proven in [26, 29], and the existence of nonzero Lyapunov exponents will *de facto* prove the nonintegrability for the other cases. One extra caveat is in order for the planar case. For $k = 0$ and $\text{sink}\Phi = \Phi$, the Hamiltonian is still not separable, and dynamics is nonintegrable. One could change variables in the metric (2.1) as $(\phi_1, \phi_2) \mapsto (\phi'_1 = \phi_1 \cos \phi_2, \phi'_2 = \phi_1 \sin \phi_2)$, and the string with the wrapping $\Phi'_2 = n\sigma$ would provide an integrable system, with the separable Hamiltonian $H' = \frac{f}{2}P_R^2 + \frac{1}{2R^2}P_{\Phi'}^2 + \frac{n^2}{2}R^2 - \frac{E^2}{2f}$. But that is a *different* system from (2.9): even though a change of variables is clearly of no physical significance, the wrapping $\Phi'_2 = n\sigma$ is physically different from $\Phi_2 = n\sigma$. Integrability clearly depends on the specific string configuration.

¹⁰For special, fine-tuned topologies and parameters, one finds integrable cases (even for string motion!) but these are special and fine-tuned; we will consider these cases elsewhere as they seem peripheral for our main story on the chaos bound.

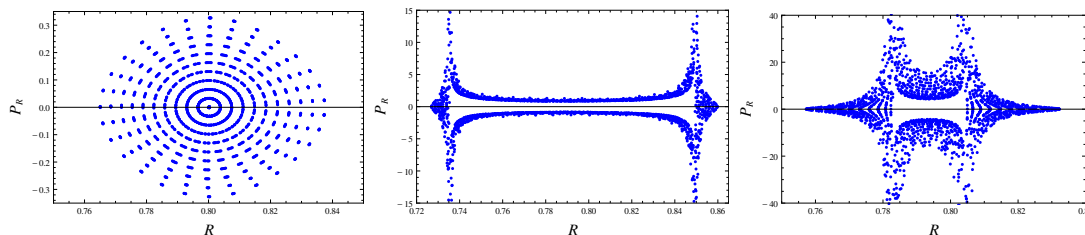


Figure 1. Poincaré section (R, P_R) for orbits starting at the apparent horizon (removed for a distance of 10^{-4} from the event horizon), at increasing temperatures $T = 0.00, 0.05, 0.10$, for a planar black hole with $m = 1$ and charge parameter q determined by the temperature. The coordinate and momentum are in units of AdS radius.

2.1 Fixed points and near-horizon dynamics

For a better overall understanding of chaos in string motion, let us sketch the general trends in dynamics first. For spherical black holes, this job was largely done in [26, 29, 44] and for similar geometries also in [27, 28]. We will emphasize mainly the properties of near-horizon dynamics that we find important for the main story.

Typical situation can be grasped from figure 1, where the Poincaré sections of orbits starting near the horizon are shown for increasing temperatures of the horizon, as well as figures 2 and 3 where we show typical orbits in the $x - y$ plane for different temperatures and initial conditions.

1. Higher temperatures generally increase chaos, with lower and lower numbers of periodic orbits (continuous lines in the Poincaré section in figure 1) and increasing areas covered with chaotic (area-filling) orbits. This is also obvious from the figure 2.
2. Orbits closer to the horizon are more chaotic than those further away; this will be quantified by the analysis of the Lyapunov exponents. This is logical, since the equations of motion for strings in pure AdS space are integrable, and far away from the horizon the spacetime probed by the string becomes closer and closer to pure AdS. An example of this behavior is seen in figure 3(A).
3. The previous two trends justify the picture of the thermal horizon as the generator of chaos. However, for an extremal or near-extremal *hyperbolic* horizon there is a slight discrepancy — in this case, moving away from the horizon increases the chaos. In other words, there is yet another mechanism of chaos generation, independent of the temperature and not located precisely at the horizon, which is subleading and not very prominent, except when it is (almost) the only one, i.e., when the horizon is (near-)extremal. This is demonstrated in figure 3(B).

When we come to the consideration of the Lyapunov exponents, we will identify the horizon-induced scrambling and the chaotic scattering as the chaos-inducing mechanisms at work for $r \rightarrow r_h$ and for intermediate r , respectively.

Consider now the radial motion from the Hamiltonian (2.9). Radial motion exhibits an effective attractive potential $E^2/2f$ which diverges at the horizon. The Φ -dependent terms

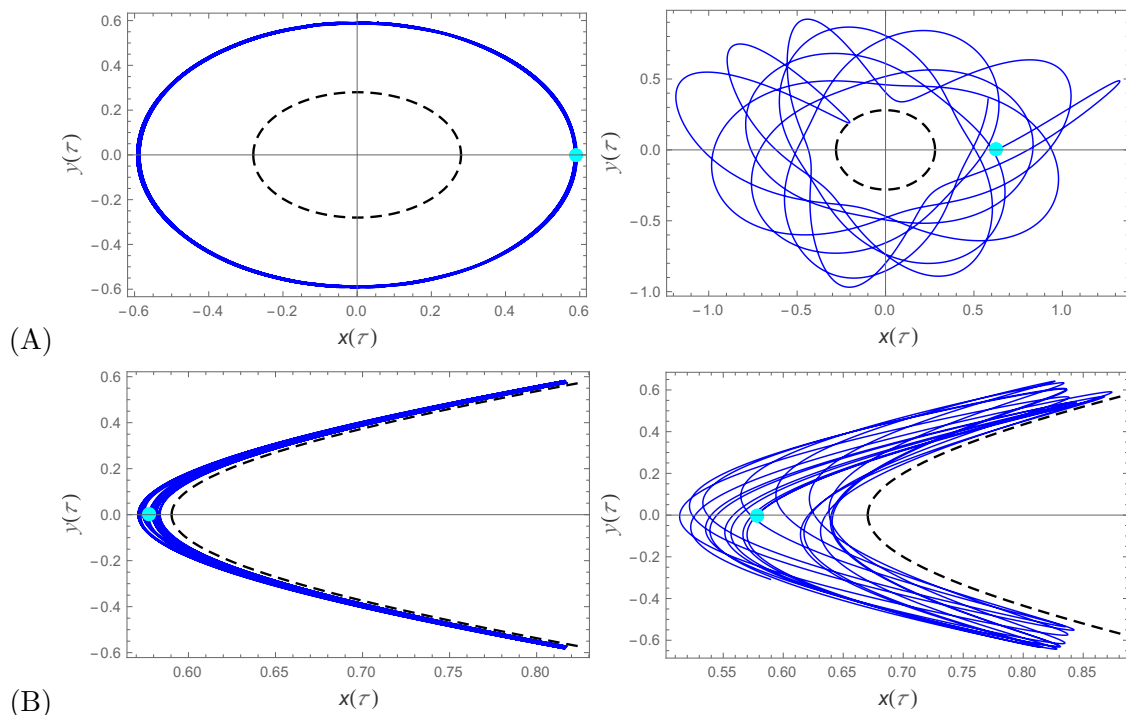


Figure 2. Thermal horizon as the generator of chaos. We show the orbits in the vicinity of the spherical (A) and hyperbolic (B) horizon, at $T = 0.01$ (left) and $T = 0.10$ (right); obviously, hot horizons generate more chaos than cold ones. The light blue dot is the initial condition of the orbit (the position of the point on the string with $\Phi = 0$ at $\tau = 0$).

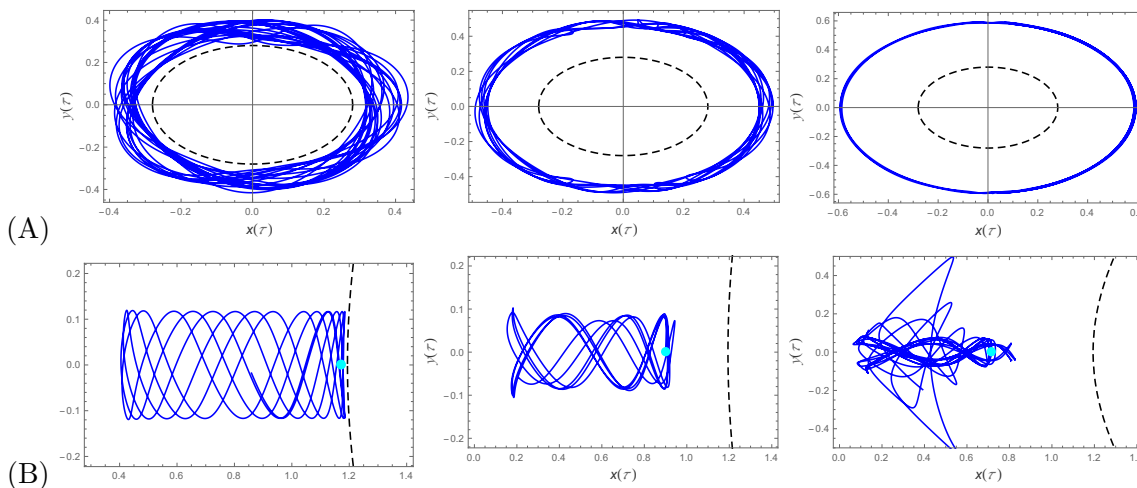


Figure 3. Thermal horizon and hyperbolic scattering as generators of chaos. In (A) and (B), we show the orbits in the vicinity of the spherical and hyperbolic horizon, respectively, at the small temperature $T = 0.01$ and starting at increasing distances from the horizon. In (A), the further from the horizon, the more regular the orbit becomes. But in the hyperbolic geometry (B), the thermally-generated chaos is negligible; instead, the orbit becomes chaotic as it explores larger and larger area of the hyperbolic manifold. Hence for hyperbolic horizons, an additional, non-thermal generator of chaos exists: it is the hyperbolic scattering. Light blue dots are again the initial positions of the string origin ($\Phi = 0$).

proportional to R^2 and $1/R^2$ are repulsive and balance out the gravitational attraction to some extent but they remain finite for all distances. For R large, the repulsion proportional to n^2 dominates so for large enough distances the string will escape to infinity. For intermediate distances more complex behavior is possible: the string might escape after some number of bounces from the black hole, or it might escape after completing some (non-periodic, in general) orbits around the black hole. The phase space has invariant planes given by $(R, P_R, \Phi, P_\Phi) = (R_0 + E\tau, E/f_0, N\pi, 0)$, with $R_0 = \text{const.}$ and $f_0 \equiv f(R_0)$ and N an integer. It is easy to verify this solution by first plugging in $\dot{\Phi} = 0$ into (2.8) to find Φ ; eq. (2.7) and the constraint (2.9) then reduce to one and the same condition $\dot{R}^2 = E^2$. We discard the solution with the minus sign (with $R = R_0 - E\tau$) as R is bounded from below. Pictorially, this solution means that a string with a certain orientation just moves uniformly toward the black hole and falls in, or escapes to infinity at uniform speed, all the while keeping the same orientation. Besides, there is a trivial fixed point at infinity, $(R, P_R, \Phi, P_\Phi) = (\infty, 0, N\pi, 0)$, found also in [26, 29].

We are particularly interested if a string can hover at a fixed radial slice $R = r_0 = \text{const.}$. Let us start from the spherical case. Inserting $R = r_0, \dot{R} = 0$ into eq. (2.8) leads to the solution in terms of the incomplete Jacobi sine integral sn (Jacobi elliptic function of the first kind, Jacobi E -function), and two integration constants to be determined. The other equation, (2.7), is a first-order relation for $\dot{\Phi}$ acting as a constraint. Solving it gives a Jacobi elliptic function again, with one undetermined constant, and we can match the constants to obtain a consistent solution:

$$\sin \Phi(\tau) = \text{sn} \left(\frac{E\sqrt{|f'_0|}}{\sqrt{2r_0 f_0}} \tau, \frac{2n^2 r_0 f_0^2}{E^2 |f'_0|} \right). \tag{2.10}$$

The value of r_0 is found from the need to satisfy also the Hamiltonian constraint. The constraint produces a Jacobi elliptic function with a different argument, and the matching to (2.10) reads

$$2f(r_0) + r_0 f'(r_0) = 0. \tag{2.11}$$

This turns out to be a cubic equation independent of the black hole charge, as the terms proportional to q cancel out. It has one real solution, *which is never above the horizon*. The solution approaches the horizon as $f'(r_0)$, approaches zero, and $r = r_h$ is obviously a solution of (2.11) for $f'(r_h) = 0$. However, the $r \rightarrow r_h$ limit is subtle in the coordinates we use because some terms in equations of motion diverge, so we need to plug in $f(r) = 0$ from the beginning. Eqs. (2.6), (2.8) then imply $\dot{R} = E$, i.e., there is no solution at constant R except for $E = 0$. This is simply because the energy is infinitely red-shifted at the horizon, i.e., E scales with f (eq. (2.6)), thus indeed unless $\dot{\mathcal{T}} \rightarrow \infty$, which is unphysical, we need $E = 0$. Now solving eq. (2.7) gives the same solution as before, of the form $\text{sn}(C_1\tau, C_2)$, with undetermined constants $C_{1,2}$, which are chosen so as to establish continuity with the solution (2.10). For an extremal horizon of the form $f \sim a(r - r_h^2) \equiv a\epsilon^2$, a smooth and finite limit is obtained by rescaling $E \mapsto E\epsilon^2$. Now expanding the sn function in ϵ produces simply a linear function at first order in ϵ :

$$\Phi(\tau) = E\tau/\sqrt{ar_0} + O(\epsilon^3). \tag{2.12}$$

Therefore, a string can hover at the extremal horizon, at strict zero temperature, when its motion (angular rotation) becomes a simple linear winding with a single frequency. Such an orbit is expected to be linearly stable, and in the next section we show it is also stable according to Lyapunov and thus has zero Lyapunov exponent. Finally, from (2.7) and (2.11) the radial velocity \dot{R} in the vicinity of a non-extremal horizon behaves as:

$$\dot{R}^2 \approx E^2 + 4\pi T r_h (r - r_h)^2, \tag{2.13}$$

meaning that \dot{R} grows quadratically as the distance from the horizon increases. This will allow us to consider near-horizon dynamics at not very high temperatures as happening at nearly constant radius: the string only slowly runs away.

For a hyperbolic horizon the calculation is similar, changing $\sin \mapsto \sinh$ in the solution (2.10). The constraint (2.11) is also unchanged (save for the sign of k in the redshift function), and the final conclusion is the same: the string can only balance at the zero temperature horizon (but now such a horizon need not be charged, as we mentioned previously). The zero temperature limit is the same linear function (2.12). For a planar horizon things are different. For $\dot{R} = 0$, we get simply harmonic motion $\Phi = C_1 \cos n\tau + C_2 \sin n\tau$, which is consistent with the constraint $H = 0$. But eq. (2.7) implies exponential motion instead, $D_1 \sinh n\tau + D_2 \cosh n\tau$. Obviously, there is no way to make these two forms consistent. Accordingly, no hovering on the horizon (nor at any other fixed radial slice) is possible for a planar black hole. But the same logic that lead to (2.13) now predicts oscillating behavior:

$$R(\tau) \approx E^2 + 4\pi T r_h (r - r_h)^2 (n^2 \cos^2 n\tau - \sin^2 n\tau). \tag{2.14}$$

Therefore, even though there are no orbits at all which stay at exactly constant R , we now have orbits which oscillate in the vicinity of the horizon forever. Averaging over long times now again allows us to talk of a string that probes some definite local temperature, determined by the average distance from the horizon.

The point of this (perhaps tedious and boring) qualitative analysis of possible orbits is the following. No orbits at fixed distance from the horizon are possible, *but* at low temperatures a string that starts near the horizon will spend a long time in the near-horizon area. Therefore, we can study the influence of the low-temperature horizon as the main chaos-generating mechanism by expanding the variational equations for the Lyapunov exponents in the vicinity of the horizon, This we shall do in the next section.

3 Lyapunov exponents and the bound on chaos

In general, Lyapunov exponents are defined as the coefficients λ of the asymptotic exponential divergence of initially close orbits; in other words, of the variation δX of a coordinate X :

$$\lambda \equiv \lim_{t \rightarrow \infty} \lim_{\delta X(0) \rightarrow 0} \frac{1}{t} \log \frac{|\delta X(t)|}{|\delta X(0)|}, \tag{3.1}$$

and the variation is expected to behave as $\delta X \sim \delta X(0) \exp(\lambda t)$ for t large and $\delta X(0)$ small enough in practice. This definition makes sense for classical systems; in quantum

mechanics, the linearity of the state vector evolution guarantees zero exponent but the intuition that initially small perturbations eventually grow large in a strongly coupled system remains when we look at appropriately defined correlation functions, like the OTOC used in [1]. We should first make the following point clear. In a classical nonlinear system, the presence of deterministic chaos leads to positive Lyapunov exponents even in absence of temperature or noise. Quantum mechanically, as we explained, the linearity of evolution means that exponential divergence is only possible in a thermal state, and this situation leads to the temperature bound on the Lyapunov exponents. This is easy to see upon restoring dimensionful constants, when the bound from [1] takes the form $\lambda \leq 2\pi k_B T/\hbar$, and indeed in a classical system where $\hbar \rightarrow 0$ no bound exists. In the context of our work, which effectively reduces to the classical Hamiltonian (2.9) which has no gravitational degrees of freedom, it is not *a priori* clear if one should expect any connection to the bound on chaos: instead of a QFT correlation function or its gravity dual, we have classical dynamics, and the Hawking temperature of the black hole is not the local temperature probed by the string. But we will soon see that analytical and numerical estimates of λ nevertheless have a form similar to the chaos bound of [1].

Before we proceed one final clarification is in order. One might worry that the Lyapunov exponents are gauge-dependent, as we consider equations of motion in terms of the worldsheet coordinate τ , and for different worldsheet coordinates the variational equations would be manifestly different; in other words, the definition (3.1) depends on the choice of the time coordinate (denoted schematically by t in (3.1)). Indeed, the value of λ clearly changes with coordinate transformations, however it has been proven that the *positivity* of the largest exponent (the indicator of chaos) is gauge-invariant; the proof was derived for classical general relativity [45] and carries over directly to the worldsheet coordinate transformations. This is all we need, because we will eventually express the τ -exponent in terms of proper time for an inertial observer, making use of the relation $\dot{\tau} = -E/f$. This could fail if a coordinate change could translate an exponential solution into an oscillating one (because then λ drops to zero and it does not make sense to re-express it units of proper time); but since we know that cannot happen we are safe.

3.1 Variational equations and analytical estimates of Lyapunov exponents

3.1.1 Thermal horizon

Consider first a thermal black hole horizon at temperature T , with the redshift function behaving as $f = 4\pi T(r - r_h) + O((r - r_h)^2)$. Variational equations easily follow from (2.6)–(2.7):

$$\delta\ddot{R} - \frac{E^2}{(R - r_h)^2}\delta R - 4\pi T \left(\dot{\Phi}^2 - n^2 \text{sink}^2\Phi \right) \delta R - 8\pi T(R - 2r_h)R\dot{\Phi}\delta\dot{\Phi} + 4\pi n^2 TR \text{sink}(2\Phi)\delta\Phi = 0 \quad (3.2)$$

$$\delta\ddot{\Phi} + n^2 \text{sink}(2\Phi) + \frac{2}{r_h}\dot{\Phi}\delta\dot{R} = 0, \quad (3.3)$$

with on-shell solutions $R(\tau), \Phi(\tau)$. This system looks hopeless, but it is not hard to extract the leading terms near the horizon which, as we explained, makes sense at low temperatures.

Therefore, we start from the solutions (2.10), (2.12), (2.14), adding a small correction $(r_0, \Phi(\tau)) \rightarrow (r_0 + \Delta R(\tau), \Phi(\tau) + \Delta\Phi(\tau))$. Then we expand in inverse powers of $r_0 - r_h$, and express the angular combinations $\dot{\Phi}^2 \pm \text{sink}^2\Phi$ making use of the constraint (2.9). When the dust settles, the leading-order equations simplify to:

$$\delta\ddot{R} - \left(16(\pi T)^3 \frac{n^2}{E^2}(r_0 - r_h) - 32(\pi T)^3 \frac{Cn}{E^2\phi_0}(r_0 - r_h)^2 \right) \delta R = 0 \quad (3.4)$$

$$\delta\ddot{\Phi} + n^2 \langle \text{cosk}^2(2\Phi) \rangle \delta\Phi = 0, \quad (3.5)$$

where $C = C(k, E)$ is a subleading (at low temperature) correction whose form differs for spherical, planar and hyperbolic horizons. From the above we read off that angular motion has zero Lyapunov exponent (the variational equation is oscillatory, because $\langle \text{cosk}^2(2\Phi) \rangle \geq 0$) but the radial component has an exponent scaling as

$$\tilde{\lambda}(T) \sim 4\sqrt{(\pi T)^3(r_0 - r_h)} \frac{n}{E} \left(1 - (r - r_h) \frac{C}{\phi_0 n} \right). \quad (3.6)$$

Now we have calculated the Lyapunov exponent in worldsheet time τ . The gauge-invariant quantity, natural also within the black hole scrambling paradigm, is the proper Lyapunov exponent λ , so that $1/\lambda$ is the proper Lyapunov time for an asymptotic observer. To relate $\tilde{\lambda}$ to λ , we remember first that the Poincare time t is related to the worldsheet time τ through (2.6) as $|dt| \sim E/f \times d\tau$. Then we obtain the proper time as $t_p = t\sqrt{-g_{00}} = t\sqrt{f}$, where near the thermal horizon we can write $f \approx 4\pi T(r - r_h)$. This gives¹¹

$$\lambda(T) \sim 2\pi T n \left(1 - \epsilon \frac{C}{\phi_0 n} \right). \quad (3.7)$$

At leading order, we get the estimate $2\pi T n$, with the winding number n acting as correction to the original bound.

3.1.2 Away from the horizon

At intermediate radii we can do a similar linear stability analysis starting from $f \sim r^2 + k + A/r$ where A is computed by series expansion (with just the AdS term $r^2 + k$ in f , without the leading black hole contribution A/r , we would trivially have integral motion and zero λ ; but this approximation applies at large, not at intermediate distances). In this case the equations of motion yield $R \sim \tau\sqrt{E^2 - 1}$, and the variational equations, after some algebra, take the form

$$\delta\ddot{R} - \frac{2}{R}(k + R^2)\delta\dot{\Phi} + E \left(\frac{3kR^2}{R^2 + k} + 1 \right) \delta R = 0. \quad (3.8)$$

One can show again that $\delta\dot{\Phi}$ is always bounded in absolute value, thus the third term determines the Lyapunov exponent. The exponent vanishes for $k > -1/3$ (because the equations becomes oscillatory) and for $k \leq -1/3$ we get

$$\lambda \sim \sqrt{-(3k + 1)E}. \quad (3.9)$$

¹¹We introduce the notation $\epsilon \equiv r - r_h$.

Since the curvature only takes the values $-1, 0, 1$, the prediction (3.9) always holds for hyperbolic horizons. Notice that this same term (the third term in (3.8)) appears as subleading in the near-horizon expansion, so we can identify it with $C(k, E)$ and write (3.7) as $\lambda(T) \sim 2\pi T n (1 - \epsilon |(3k + 1)E| / (\phi_0 n))$. This holds for any k , and we see that $C \leq 0$; thus the bound is only approached from below as it should be.

In absence of negative curvature, i.e., for $k > 0$, we have vanishing C at leading order in $1/R$ but subleading contributions still exist, so both the slight non-saturation of the limit $2\pi T n$ near-horizon (for small ϵ) and a parametrically small non-zero Lyapunov exponent at intermediate distances will likely appear, which we see also in the numerics. That the motion is chaotic on a pseudosphere (negative curvature) is of course no surprise; it is long known that both particles and waves have chaotic scattering dynamics on pseudospheres [46]. We dub this contribution the scattering contribution to the Lyapunov exponent, as opposed to the scrambling contribution. It is largely independent of temperature and largely determined by the geometry of the spacetime away from the horizon.

3.1.3 Extremal horizon

For an extremal horizon we replace f by $f \sim a(r - r_h)^2 = a\epsilon^2$, and plug in this form into the variational equations. Now the result is (for concreteness, for the spherical horizon)

$$\delta\ddot{R} - \left(\frac{a^2 \epsilon^4 r_h^2 n^2}{2a\epsilon r_h - 2a\epsilon^2} \right) \delta R = 0 \tag{3.10}$$

$$\delta\ddot{\Phi} + n^2 \langle \text{cosk}(2\Phi) \rangle \delta\Phi = 0, \tag{3.11}$$

leading to a vanishing exponent value:

$$\tilde{\lambda}(T) \sim \sqrt{ar_h/2n\epsilon^3} \rightarrow 0. \tag{3.12}$$

Obviously, this also means $\lambda = 0$ — there is no chaos at the extremal horizon. This is despite the fact that the string motion in this case is still nonintegrable, which is seen from the fact that no new symmetries or integrals of motion arise in the Hamiltonian in this case. The horizon scrambling is proportional to temperature and does not happen at $T = 0$, but the system is still nonintegrable and the chaos from other (scattering) origins is still present. In particular, the estimate (3.8)–(3.9) remains unchanged.

The estimates (3.7), (3.9), (3.12) are the central sharp results of the paper. We can understand the following physics from them:

1. At leading order, we reproduce (and saturate) the factor $2\pi T$ of the Maldacena-Shenker-Stanford bound, despite considering classical dynamics only.
2. The bound is however multiplied by the winding number n of the ring string. The spirit of the bound is thus preserved but an extra factor — the winding number — enters the story.
3. Taking into account also the scattering chaos described by (3.9), the results are in striking accord with the idea of [2]: there are two contributions to chaos, one

proportional to the black hole temperature and solely determined by the scrambling on the horizon, with the universal factor $2\pi T$ expected from the concept of black holes as the fastest scramblers in nature, and another determined by the (slower) propagation of signals from the horizon toward the AdS boundary, which we call the scattering term, as it is determined also by dynamics at large distances.

4. For a particle ($n = 0$), we correctly get $\lambda = 0$, as the geodesics are integrable.
5. The temperature appearing in (3.7) is always the Hawking temperature of the black hole T .

In the next section, when we consider the AdS/CFT interpretation, we will try to shed some more light on where the modification of the bound $2\pi T \mapsto 2\pi Tn$ comes from.

3.1.4 Lyapunov time versus event time

In the above derivations we have left one point unfinished. We have essentially assumed that $R(\tau) \approx \text{const.} = r$ and treated the difference $\epsilon = r - r_h$ as a fixed small parameter. This is only justified if the local Lyapunov time $1/\tilde{\lambda}$ is much shorter than the time to escape far away from r_h and the horizon, or to fall into the black hole. In other words, it is assumed that the Lyapunov time is much shorter than the “lifetime” of the string (let us call it event time t_E). Now we will show that this is indeed so. For the spherical black hole, upon averaging over the angle Φ , we are left with a one-dimensional system

$$\dot{R}^2 + R^2 f(R) \frac{E^2 f'(R)}{R f^2(R)} = E^2, \tag{3.13}$$

which predicts the event time as

$$t_E \sim \int_{r_0}^{r_h, \infty} \frac{dR}{\sqrt{|E - E f'(R) R f^2(R)|}} \approx \frac{\pi r_h}{\sqrt{2}} \frac{1}{\sqrt{4\pi T \epsilon n}} \approx \frac{\pi r_h}{\sqrt{2}} \times \frac{\tilde{\lambda}^{-1}}{\epsilon}. \tag{3.14}$$

In other words, the event times are roughly by a factor $1/\epsilon$ longer than Lyapunov times, therefore our estimate for λ should be valid. In (3.14), we have considered both the infalling orbits ending at r_h , and the escaping orbits going to infinity (for the latter, we really integrate to some $r_\infty > r_0$ and then expand over $1/r_\infty$). An orbit will be infalling or escaping depending on the sign of the combination under the square root, and to leading order both cases yield a time independent of r_0 (and the cutoff r_∞ for the escaping case). The hyperbolic case works exactly the same way, and in the planar case since $R(\tau)$ oscillates the event time is even longer (as there is no uniform inward or outward motion). For extremal horizons, there is no issue either as $r = r_h$ is now the fixed point.

3.1.5 Dimensionful constants

One might wonder what happens when dimensionful constants are restored in our results for the Lyapunov exponents like (3.7) or (3.9): the original chaos bound really states $\lambda \leq 2\pi k_B T/\hbar$, and we have no \hbar in our system so far. The resolution is simple: the role of \hbar is played by the inverse string tension $2\pi\alpha'$, which is obvious from the standard form

of the string action (2.3); the classical string dynamics is obtained for $\alpha' \rightarrow 0$. Therefore, the dimensionful bound on chaos for our system reads $\lambda = 2\pi k_B T n / 2\pi\alpha' = k_B T n / \alpha'$. Another way to see that α' takes over the role of \hbar in the field-theory derivation [1] is that the weight in computing the correlation functions for a quantum field is given by the factor $\exp(-1/\hbar \int \mathcal{L})$, whereas for a string the amplitudes are computed with the weight $\exp(-1/2\pi\alpha' \int \mathcal{L})$. In the next section, we will also look for the interpretation in the framework of dual field theory. In this context, α' is related to the number of degrees of freedom in the gauge dual of the string, just like the Newton's constant G_N is related to the square of the number of colors N^2 in the gauge dual of a pure gravity theory. But the issues of gauge/string correspondence deserve more attention and we treat them in detail in section 4.

3.2 Numerical checks

We will now inspect the results (3.7), (3.9), (3.12) numerically. Figure 4 tests the basic prediction for the horizon scrambling, $\lambda \approx 2\pi T n$ at low temperatures: both the n -dependence at fixed temperature (A), and the T -dependence at fixed n (B) are consistent with the analytical prediction. All calculations were done for the initial condition $\dot{R}(0) = 0$, and with energy E chosen to ensure a long period of hovering near the horizon. The temperatures are low enough that the scattering contribution is almost negligible. In figure 5 we look at the scattering term in more detail. First we demonstrate that at zero temperature, the orbits in non-hyperbolic geometries are regular (A): the scattering term vanishes at leading order, and the scrambling vanishes at $T = 0$. In the (B) panel, scattering in hyperbolic space at intermediate radial distances gives rise to chaos which is independent of the winding number, in accordance to (3.7). To further confirm the logic of (3.7), one can look also at the radial dependence of the Lyapunov exponent: at zero temperature, there is no chaos near-horizon (scrambling is proportional to T and thus equals zero; scattering only occurs at finite $r - r_h$), scattering yields a nonzero λ at intermediate distances and the approach to pure AdS at still larger distances brings it to zero again; at finite temperature, we start from $\lambda = 2\pi T n$ near-horizon, observe a growth due to scattering and fall to zero approaching pure AdS.

4 Toward a physical interpretation of the modified bound

4.1 Dual gauge theory interpretation

The ring string wrapped along the σ coordinate is a very intuitive geometry from the viewpoint of bulk dynamics. However it has no obvious interpretation in terms of the gauge/gravity duality, and the Hamiltonian (2.9) itself, while simple-looking, is rather featureless at first glance: essentially a forced nonlinear oscillator, it does not ring a bell on why to expect the systematic modification of the Maldacena-Shenker-Stanford bound and what the factor n means. Thus it makes sense to do two simple exercises: first, to estimate the energy and spin of the operators corresponding to (2.5) to understand if it has to do with some Regge trajectory; second, to consider some other string configurations, with a more straightforward connection to the operators in gauge theory. Of course, finite

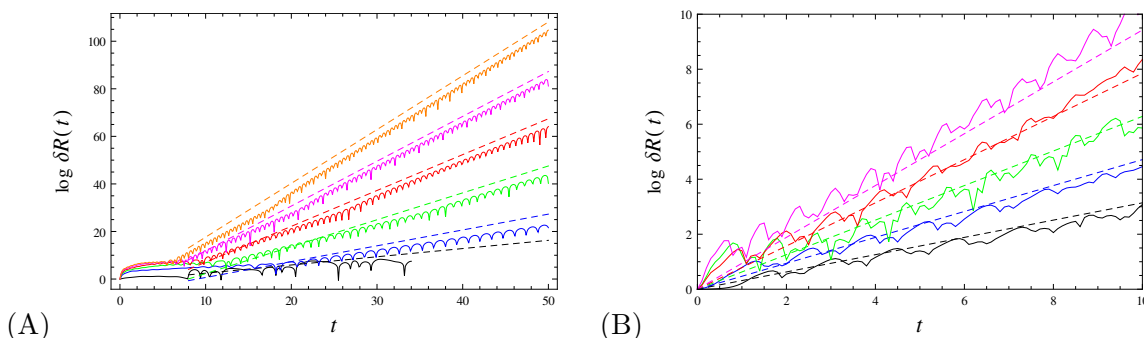


Figure 4. (A) Logarithm of the relative variation of the coordinate R , for a spherical AdS-Reissner-Nordstrom black hole, for a fixed temperature $T = 0.04$ and increasing winding numbers $n = 1, 2, 3, 4, 5, 6$ (black, blue, green, red, magenta, orange). Full lines are the numerical computational of the function $\log(\delta X(\tau)/\delta X(0)) = \lambda\tau$, so their slopes equal the Lyapunov exponents λ . Dashed lines show the analytically predicted bound $\log \delta X = 2\pi T n \tau + \log X_0$. Numerically computed variations almost saturate the bounds denoted by the dashed lines. The calculation for $n = 1$ is stopped earlier because in this case the orbit falls in into the black hole earlier than for higher n . (B) Same as (A) but for a hyperbolic AdS-Schwarzschild black hole, at fixed $n = 1$ and increasing temperature $T = 0.050, 0.075, 0.100, 0.125, 0.150$ (black, blue, green, red, magenta), again with analytically predicted bounds shown by the dashed lines. For the two highest temperatures (red, magenta) the computed slopes are slightly above the bound probably because the near-horizon approximation does not work perfectly well. The short-timescale oscillations superimposed on the linear growth, as well as the nonlinear regime before the linear growth starts in the panel (A) are both expected and typical features of the variation δR (Lyapunov exponents are defined asymptotically, for infinite times).

temperature horizons are crucial for our work on chaos, and saying *anything* precise about the gauge theory dual of a string in the black hole background is extremely difficult; we will only build some qualitative intuition on what our chaotic strings do in field theory, with no rigorous results at all.

Let us note in passing that the ring string configurations considered so far are almost insensitive to spacetime dimension. Even if we uplift from the four-dimensional spacetime described by (t, r, ϕ_1, ϕ_2) to a higher-dimensional spacetime $(t, r, \phi_1, \phi_2, \dots, \phi_{N-2})$, with the horizon being an $N - 2$ -dimensional sphere/plane/pseudosphere, the form of the equations of motion does not change if we keep the same ring configuration, with $\Phi_1 = \Phi_1(\tau, \sigma)$, $\Phi_2 = n\sigma$, $\Phi_3 = \text{const.}, \dots, \Phi_{N-2} = \text{const.}$ — this is a solution of the same eqs. (2.6)–(2.8) with the same constraint (2.9). The difference lies in the redshift function $f(r)$ which depends on dimensionality. This, however, does not change the main story. We can redo the calculation of the radial fixed point from the second section, to find a similar result — a string can oscillate or run away/fall slowly in the vicinity of a horizon, and the variational equations yield the same result for the Lyapunov exponent as before. It is really different embeddings, i.e., different Polyakov actions, that might yield different results.

4.1.1 Operators dual to a ring string?

We largely follow the strategy of [47] in calculating the energy and the spin of the string and relating them to the dual Yang-Mills theory. In fact, the ring string is quite close to

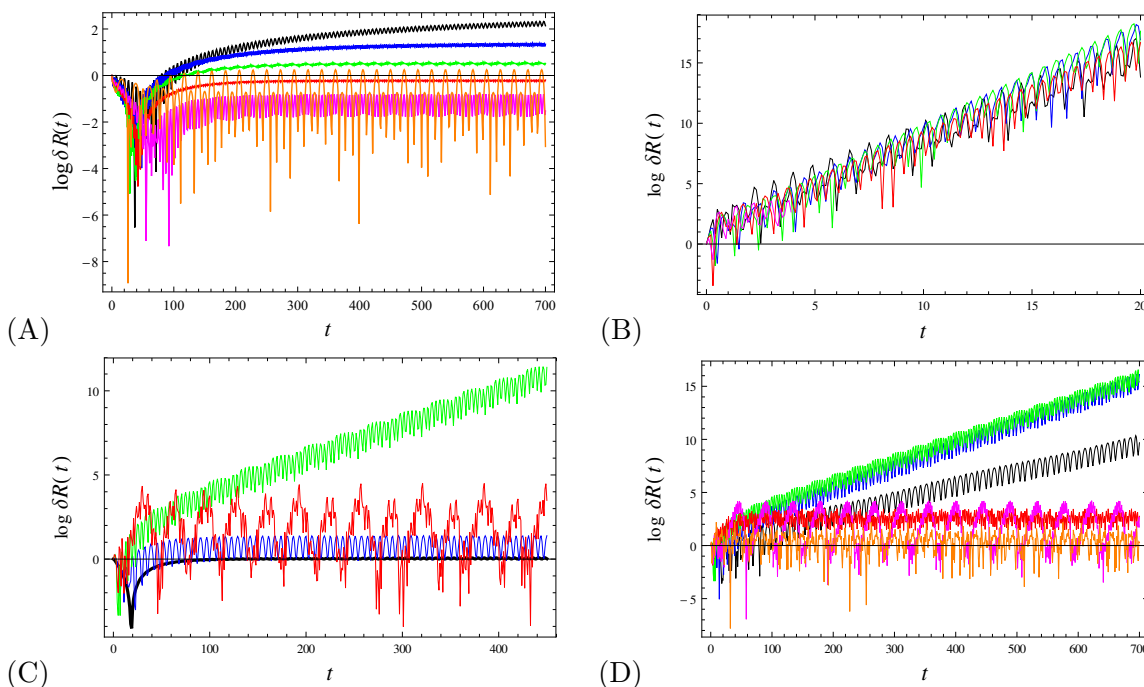


Figure 5. (A) Logarithm of the radial variation δR for near-horizon orbits with $n = 1, 2, 3, 4, 5, 6$ (black, blue, green, red, magenta, orange) in a planar extremal Reissner-Nordstrom geometry. All curved asymptote to a constant, i.e., (almost) zero slope, resulting in $\lambda \approx 0$. (B) Same as previous for an extremal hyperbolic black hole. Now the Lyapunov exponent is nonzero, and equal for all winding numbers: in absence of thermal scrambling, the chaos originates solely from scattering, which is independent of n . (C) The Lyapunov exponent in zero-temperature hyperbolic black hole background for $n = 1$ and $r = r_h, 1.1r_h, 1.2r_h, 1.3r_h$ (black, blue, green, red) starts at zero (no scrambling, no scattering), grows to a clear nonzero value for larger radii due to scattering, and again falls to zero for still larger distances, as the geometry approaches pure AdS (D) Lyapunov exponent in $T = 0.02$ hyperbolic black hole background for $n = 1$ and $r = r_h, 1.1r_h, 1.2r_h, 1.3r_h, 1.4r_h, 1.5r_h$ (black, blue, green, red, magenta, orange) starts at the scrambling value (black), reaches its maximum when both scrambling and scattering are present (blue, green) and then falls to zero when AdS is approached (red, magenta, orange).

what the authors of [47] call the oscillating string, *except* that we allow one more angle to fluctuate independently (thus making the system nonintegrable) and, less crucially, that in [47] only the winding number $n = 1$ is considered.

Starting from the action for the ring string (2.3), we write down the expressions for energy and momentum:

$$\mathcal{E} = \frac{1}{2\pi\alpha'} \int d\tau \int d\sigma P_{\mathcal{T}} = \frac{E}{\alpha'} \int_{\phi_1}^{\phi_2} \frac{d\Phi}{\dot{\Phi}} \quad (4.1)$$

$$\mathcal{S} = \frac{1}{2\pi\alpha'} \int d\tau \int d\sigma P_{\Phi} = \frac{1}{\alpha'} \int_{\phi_1}^{\phi_2} \frac{d\Phi}{\dot{\Phi}} R^2(\Phi) \dot{\Phi}, \quad (4.2)$$

where the second worldsheet integral gives simply $\int d\sigma = 2\pi$ as R, Φ do not depend on σ , and we have expressed $d\tau = d\Phi/\dot{\Phi}$; finally, the canonical momentum is conserved, $P_{\mathcal{T}} = E$,

and in the expression for the spin we need to invert the solution $\Phi(\tau)$ into $\tau(\Phi)$ in order to obtain the function $R(\Phi)$. We are forced to approximate the integrals. Expressing $\dot{\Phi}$ from the Hamiltonian constraint (2.4), we can study the energy in two regimes: small amplitude $\phi_0 \ll \pi$ which translates to $E/T \ll 1$, and large amplitude $\phi_0 \sim \pi$, i.e., $E/T \sim 1$. For these two extreme cases, we get:

$$\mathcal{E} \approx \frac{4r_0 \sqrt{f(r_0)}}{\alpha'} \phi_0 = \frac{4E}{\alpha'n}, \quad \phi_0 \ll \pi \tag{4.3}$$

$$\mathcal{E} \approx \frac{\pi E}{\alpha'n}, \quad \phi_0 \sim \pi \tag{4.4}$$

For the spin similar logic gives

$$\mathcal{S} \approx \frac{8r_0 E}{\alpha' \sqrt{f(r_0)}} \phi_0 = \frac{8E^2}{\alpha'n} \frac{1}{f(r_0)} \approx \frac{8E^2}{\alpha'n} \frac{1}{4\pi T \epsilon}, \quad \phi_0 \ll \pi \tag{4.5}$$

$$\mathcal{S} \approx \frac{4E^2}{\alpha'n} \sqrt{\frac{2f'(r_0)r_0}{f^3(r_0)}} \approx \frac{8E^2}{\alpha'n} \frac{\sqrt{2\pi}}{4\pi T \epsilon}, \quad \phi_0 \sim \pi. \tag{4.6}$$

The bottom line is that in both extreme regimes (and then presumably also in the intermediate parameter range) we have $\mathcal{E} \propto E/\alpha'n$ and $\mathcal{S} \propto E^2/\alpha'nT\epsilon$; as before $\epsilon = r - r_h$ and it should be understood as a physical IR cutoff (formally, for $r \rightarrow r_h$ the spin at finite temperature diverges; but we know from section 2 that in fact no exact fixed point at constant r exists, and the average radial distance is always at some small but finite distance ϵ). Therefore, we have $\mathcal{E}^2 \propto \mathcal{S}/\alpha'nT\epsilon$.

The presence of temperature in the above calculation makes it hard to compare the slope to the familiar Regge trajectories. But in absence of the black hole, when $f(r) = 1$, we get

$$\mathcal{E} = 4E/\alpha'n, \quad \mathcal{S} = 8E^2/\alpha'n \Rightarrow \mathcal{E}^2 = 2\mathcal{S}/\alpha'n. \tag{4.7}$$

For $n = 1$, this is precisely the leading Regge trajectory. For higher n the slope changes, and we get a different trajectory. Therefore, the canonical Lyapunov exponent value $\lambda = 2\pi T$ precisely corresponds to the leading Regge trajectory. We can tentatively conclude that the winding string at finite temperature describes complicated thermal mixing of large-dimension operators of different dimensions and spins, and these might well be sufficiently nonlocal that the OTOC never factorizes and the bound from [1] does not apply.

4.1.2 Planetoid string

In this subsection we consider so-called planetoid string configurations, also studied in [47] in the zero-temperature global AdS spacetime and shown to reproduce the leading Regge trajectory in gauge theory. This is again a closed string in the same black hole background (2.1) but now the solution is of the form¹²

$$\mathcal{T} = e\tau, \quad R = R(\sigma), \quad \Phi_1 = \Phi_1(\tau), \quad \Phi_2 = \Phi_2(\sigma), \tag{4.8}$$

¹²The authors of [47] work mostly with the Nambu-Goto action but consider also the Polyakov formulation in the conformal gauge; we will stick to the Polyakov action from the beginning for notational uniformity with the previous section. For the same reason we keep the same coordinate system as in (2.1).

where the auxiliary field e is picked so as to satisfy the conformal gauge, and any additional coordinates Φ_3, Φ_4, \dots and $\Theta_1, \Theta_2, \dots$ are fixed. The Lagrangian

$$L = -\frac{1}{2f} (R')^2 - \frac{e^2}{2} f + \frac{R^2}{2} \left(-\dot{\Phi}_1^2 + \sin^2 \Phi_1 \Phi_2'^2 \right) \quad (4.9)$$

has the invariant submanifold $\Phi_1 = \omega\tau, \Phi_2 = \text{const.}$ when the dynamics becomes effectively one-dimensional, the system is trivially integrable and, in absence of the black hole, it is possible to calculate exactly the energy and spin of the dual field theory operator. This is the integrable case studied in [47, 48], and allowing Φ_2 to depend on σ seems to be the only meaningful generalization, because it leads to another submanifold of integrable dynamics with $R = r_0 = \text{const.}, \Phi_2 = n\sigma$ and the pendulum solution for Φ_1 :

$$\sin \Phi_1(\tau) = \text{sn} \left(\ell\tau, -\frac{n^2}{\ell^2} \right), \quad (4.10)$$

where $\ell^2 = \dot{\Phi}_1^2 - n^2 \sin^2 \Phi_1$ is the adiabatic invariant on this submanifold. With two integrable submanifolds, a generic orbit will wander between them and exhibit chaos. The variational equations can be analyzed in a similar fashion as in the previous section. Here, the chaotic degree of freedom is $\Phi_1(\tau)$, with the variational equation

$$\delta\ddot{\Phi}_1 - \Phi_2'^2 \cos(2\Phi_1) = 0, \quad (4.11)$$

which in the near-horizon regime yields the Lyapunov exponent

$$\lambda = 2\pi T n, \quad (4.12)$$

in the vicinity of the submanifold (4.10). In the vicinity of the other solution ($\Phi_1 = \omega\tau, \Phi_2 = \text{const.}$), we get $\lambda = 0$. Chaos only occurs in the vicinity of the winding string solution, and the winding number again jumps in front of the universal $2\pi T$ factor.

Now let us see if this kind of string reproduces a Regge trajectory. In the presence of the black hole the calculation results in very complicated special functions, but we are only interested in the leading scaling behavior of the function $\mathcal{E}^2(\mathcal{S})$. Repeating the calculations from (4.1)–(4.2), we first reproduce the results of [47] in the vicinity of the solution $\Phi_1 = \omega\tau$: for short strings, we get $\mathcal{E} \sim 2/\omega T, \mathcal{S} \sim 2/\omega^2 T^2$ and thus $\mathcal{E}^2 \propto 2\mathcal{S}$, precisely the result for the leading Regge trajectory. Now the Regge slope does not depend on the temperature (in the short string approximation!). This case, as we found, trivially satisfies the original chaos bound ($\lambda = 0$, hence for sure $\lambda < 2\pi T$). In the vicinity of the other solution, with $R = r_0$, things are different. Energy has the following behavior:

$$\mathcal{E} \sim \frac{8\pi T}{\alpha' n}, \quad \ell \ll 1 \quad (4.13)$$

$$\mathcal{E} \sim \frac{8\pi^2 T}{\alpha' \ell}, \quad \ell \gg 1. \quad (4.14)$$

For the spin, the outcome is

$$\mathcal{S} \sim \frac{2r_0^2 \ell}{\alpha' n}, \quad \ell \ll 1 \quad (4.15)$$

$$\mathcal{S} \sim \frac{2r_0^2}{\alpha'}, \quad \ell \gg 1, \quad (4.16)$$

so in this case there is no Regge trajectory at all, i.e., no simple relation for $\mathcal{E}^2(\mathcal{S})$ because the scale r_0 and the quantity ℓ show up in the $\mathcal{E}^2(\mathcal{S})$ dependence even at zero temperature.

In conclusion, the strings that can violate the chaos bound have a strange Regge behavior in the gauge/string duality, in this case in a more extreme way than for the ring strings (even for $n = 1$ no Regge trajectory is observed). The strings which have $\lambda = 0$ and thus trivially satisfy the bound on the other hand obey the leading Regge trajectory.

4.2 The limits of quasiclassicality

One more thing needs to be taken into account when considering the modification of the chaos bound. Following [8], one can suspect that the violating cases are not self-consistent in the sense that they belong to the deep quantum regime when semiclassical equations (in our case for the string) cease to be valid and quantum effects kill the chaos. For a ring string this seems not to be the case. To check the consistency of the semiclassical limit, consider the energy-time uncertainty relation $\Delta\mathcal{E}\Delta t \geq 1$. The energy uncertainty is of the order of $E/\alpha'n$ as we found in (4.3)–(4.4), and the time uncertainty is precisely of the order of the Lyapunov time $1/2\pi Tn$; the uncertainty relation then gives $E \geq 2\pi Tn^2\alpha'$. On the other hand, we require that the spin \mathcal{S} should be large in the classical regime: $\mathcal{S} \gg 1$. This implies $E^2 \gg 4\pi T\epsilon n\alpha'$ or, combining with the uncertainty relation, $Tn^3\alpha' \gg \epsilon$. Roughly speaking, we need to satisfy simultaneously $Tn^2 \leq 1/\alpha'$ and $Tn^3 \gg \epsilon/\alpha'$, which is perfectly possible: first, we need to have small enough α' (compared to Tn^2), as could be expected for the validity of the semiclassical regime; second, we need to have sufficiently large $n/\epsilon \gg 1$, which can be true even for $n = 1$ for small ϵ , and for sure is satisfied for sufficiently large n even for $\epsilon \sim 1$. In conclusion, there is a large window when the dynamics is well-described by the classical equations (and this window even grows when $n \gg 1$ and the violation of the chaos bound grows).

4.3 Ring string scattering amplitude and the relation to OTOC

So far our efforts to establish a field theory interpretation of a ring string in black hole background have not been very conclusive, which is not a surprise knowing how hard it is in general to establish a gauge/string correspondence in finite-temperature backgrounds and for complicated string geometries. Now we will try a more roundabout route and follow the logic of [4–6], constructing a gravity dual of the OTOC correlation function, which has a direct interpretation in field theory; it defines the correlation decay rate and the scrambling time of some boundary operator. In [17] this formalism was already applied to study the OTOC of field theory operators (heavy quarks) dual to an open string in BTZ black hole background, hanging from infinity to infinity through the horizon in eikonal approximation. That case has a clear interpretation: the endpoints of the string describe the Brownian motion of a heavy quark in a heat bath. As we already admitted, we do not have such a clear view of what our case means in field theory, but we can still construct the out-of-time ordered correlator corresponding to whatever complicated boundary operator our string describes.

We will be deliberately sketchy in describing the basic framework of the calculation as it is already given in great detail in [4–6]. The idea is to look at the correlation func-

tion $\langle\langle \hat{V}_{x_1}(t_1)\hat{W}_{x_2}(t_2)\hat{V}_{x_3}(t_3)\hat{W}_{x_4}(t_4)\rangle\rangle$ of some operators V, W at finite temperature (hence the expectation value $\langle\langle(\dots)\rangle\rangle$ includes both quantum-mechanical and thermal ensemble averaging). The time moments need not be ordered; we are often interested in the case $\Re t_1 = \Re t_3 \equiv 0, \Re t_2 = \Re t_4 = t$.¹³ This correlation function corresponds to the scattering amplitude between the in and out states of a perturbation sourced from the boundary. The propagation of the perturbation is described by the bulk-to-boundary propagators K . The perturbation has the highest energy at the horizon since the propagation in Schwarzschild time becomes a boost in Kruskal coordinates, and the perturbation, however small at the boundary, is boosted to high energy in the vicinity of the horizon. In the Kruskal coordinates defined the usual way:

$$U = -e^{\frac{t-r_*}{2r_h}}, \quad V = e^{\frac{t+r_*}{2r_h}}, \quad r_* = \int_r^\infty \frac{dr}{f(r)}, \quad (4.17)$$

the scattering amplitude becomes

$$D = \prod_{i=1}^4 \int d^2 p_i \int d^2 x_i K^*(p_3; x_3) K^*(p_4; x_4) K(p_1; x_1) K(p_2; x_2)_{\text{out}}(p_3^U, p_4^V; x_3, x_4 | p_1^U, p_2^V; x_1, x_2)_{\text{in}}. \quad (4.18)$$

The propagators are expressed in terms of the Kruskal momenta $p_i = (p_i^U, p_i^V)$ and the coordinates $x_i = (x_i^1, x_i^2)$ in the transverse directions. The in-state is defined by (p_3^U, x^3) at $U = 0$, and by (p_4^V, x^4) at $V = 0$, and analogously for the out-state. The form of the propagators is only known in the closed form for a BTZ black hole (in 2+1 dimensions), but we are happy enough with the asymptotic form near the horizon. For simplicity, consider a scalar probe of zero bulk mass, i.e., the conformal dimension $\Delta = D$, and at zero black hole charge, i.e., for a Schwarzschild black hole. The propagator then behaves as ($\tilde{\omega} \equiv \omega/4\pi T$):

$$K(p^U, p^V) \sim \frac{\pi}{\sinh(\frac{\pi}{T})} \frac{1 - e^{-\pi\tilde{\omega}}}{\Gamma(-i\tilde{\omega}) \Gamma(i\tilde{\omega})} \frac{e^{-i\tilde{\omega}t}}{(p^U)^{1+i\tilde{\omega}} + (p^V)^{1-i\tilde{\omega}}} e^{i(p^U V + p^V U)}. \quad (4.19)$$

The task is thus to calculate the amplitude (4.18) with the propagators (4.19). In the eikonal approximation used in most of the literature so far, the problem boils down to evaluating the classical action at the solution. However, it is not trivial to justify the eikonal approximation for a ring string. Let us first suppose that the eikonal approximation works and then we will see how things change if it doesn't.

4.3.1 Eikonal approximation

If the energy in the local frame near the horizon is high enough, then we have approximately $p_1^U \approx p_3^U \equiv p, p_2^V \approx p_4^V \equiv q$ so that $p_1^V \approx p_2^U \approx p_3^V \approx p_4^U \approx 0$, and for a short enough scattering event (again satisfied if the energy and thus the velocity is high enough) the coordinates are also roughly conserved, therefore the amplitude $\langle\text{out}|\text{in}\rangle$ is diagonal and can be written as a phase shift $\exp(i\delta)$. The point of the eikonal approximation is that the

¹³In the Schwinger-Keldysh finite-temperature formalism the time is complex, with the imaginary time axis compactified to the radius of the inverse temperature.

shift δ equals the classical action. The action of the ring configuration is

$$S = \frac{1}{2\pi\alpha'} \int d\tau \int d\sigma \left(\frac{\dot{R}^2}{2} (\dot{\Phi}_1^2 - n^2 \sin^2 \Phi_1) + \frac{\dot{R}^2}{2f} + \frac{f}{2} \dot{\mathcal{T}}^2 \right). \quad (4.20)$$

We will consider again the string falling slowly in the vicinity of the horizon (see eqs. (2.10)–(2.14)) and put $\dot{R} \rightarrow 0, R(\tau) \approx r_0, r_0 - r_h \ll r_h$. Now we need to pass to the Kruskal coordinates and then introduce the new variables $\mathbb{T} = (V + U)/2, X = (V - U)/2$. In these coordinates the near-horizon geometry in the first approximation is Minkowskian and we can easily expand around it as required for the eikonal approximation. The action and the energy (to quartic order in the fluctuations) are now

$$S = \frac{1}{2\pi\alpha'} \int d\tau \int d\sigma \left[\frac{1}{2} (-\dot{\mathbb{T}}^2 + \dot{X}^2 + r_0^2 \dot{\Phi}^2 + r_0^2 n^2 \sin^2 \Phi^2) \left(1 + \frac{\mathbb{T}^2 - X^2}{2} \right) \right] \quad (4.21)$$

$$\mathcal{E} = \frac{1}{2\pi\alpha'} \int d\tau \int d\sigma \frac{\dot{\mathbb{T}}}{(1 - \mathbb{T}^2 + X^2)^2}. \quad (4.22)$$

As a sanity check, for $n = 0$ the fluctuations of the (\mathbb{T}, X) variables in the action (4.21) are the same as in [17], although we use a different worldsheet parametrization. The dynamics of the angle Φ crucially depends on the winding number. One consequence is that the on-shell action is nontrivial already at quadratic order. For the solution (2.10) — the slowly-moving near-horizon string — we can assume $\dot{\mathbb{T}}, \dot{X} \ll \dot{\Phi}$, so the equations of motion yield as approximate on-shell solutions

$$\mathbb{T} = \mathbb{T}_0 e^{nr_0\tau/\sqrt{2}}, \quad X = X_0 e^{-nr_0\tau/\sqrt{2}}, \quad (4.23)$$

so that, as the perturbation dies out, the string approaches the locus $\mathbb{T}_0 = 0 \Rightarrow U = -V \Rightarrow t \rightarrow \infty$. Inserting (4.23) into (4.21), we obtain, after regularizing the action:

$$S^{(0)} = \frac{nr_0}{2\alpha'} \mathbb{T}_0^2 + \dots \quad (4.24)$$

$$\mathcal{E}^{(0)} = \frac{\sqrt{2}}{\alpha'} \mathbb{T}_0 + \dots \quad (4.25)$$

Therefore, $S^{(0)} = (\mathcal{E}^{(0)})^2 \times nr_h \alpha' / 4$ (where we have plugged in $r_0 \approx r_h$): the action is proportional to the square of energy, which equals $\mathcal{E}^2 = pq$ in the center-of-mass frame. This is perfectly in line with the fast scrambling hypothesis. Plugging in $\delta = S^{(0)}$ into the amplitude in (4.18) and rescaling

$$T_{13} \equiv e^{2\pi T t_1} - e^{2\pi T t_3^*}, \quad T_{24} \equiv e^{-2\pi T t_4^*} - e^{-2\pi T t_2} \quad (4.26)$$

$$p^U = \frac{p}{i} \frac{1}{T_{13}}, \quad p^V = \frac{q}{i} \frac{1}{T_{24}} \quad (4.27)$$

we obtain:

$$D = N_\omega^4 \left(e^{2\pi T t_1} - e^{2\pi T t_3^*} \right)^2 \left(e^{-2\pi T t_4^*} - e^{-2\pi T t_2} \right)^2 \int \frac{dp}{p^2} \int \frac{dq}{q^2} e^{-p-q-i\frac{pq}{T_{13}T_{24}}\frac{\alpha' nr_h}{4}}, \quad (4.28)$$

with N_ω containing the first two factors in (4.19) which only depend on ω and T . Introducing the change of variables $p = Q \sin \gamma, q = Q \cos \gamma$, we can reduce (4.28) to an exponential integral. With the usual contour choice for OTOC $\Im t_i = -\epsilon_i, \Re t_1 = \Re t_3 = 0, \Re t_2 = \Re t_4 = t$, we end up at leading order with

$$D \sim 1 + 2i\alpha' n r_h e^{2\pi T t} \Rightarrow \lambda_{\text{OTOC}} \sim 2\pi T, t_* \sim \frac{1}{2\pi T} \log \frac{1/\alpha'}{n r_h}. \quad (4.29)$$

Therefore, the Lyapunov time as defined by the OTOC in field theory precisely saturates the predicted bound $2\pi T$, and in the eikonal approximation is not influenced by the winding number n . On the other hand, the scrambling time t_* is multiplied by a factor of $\log(1/\alpha' n)$ (the horizon radius can be rescaled to an arbitrary value by rescaling the AdS radius, thus we can ignore the factor of r_h). The factor $1/\alpha'$ appears also in [17] and plays the role of a large parameter, analogous to the large N^2 factor in large- N field theories: the entropy of the string (the number of degrees of freedom to be scrambled) certainly grows with $1/\alpha'$. For a ring string, this factor is however divided by n , as the number of excitations is reduced by the implementation of the periodic winding boundary condition. Therefore, the winding of the ring string indeed speeds up the chaotic diffusion, by speeding up the scrambling. However, the faster scrambling is not seen in the timescale of local divergence which, unlike the classical Lyapunov exponent, remains equal to $2\pi T$; it is only seen in the timescale on which the perturbation permeates the whole system.

In conclusion, *the violation of the Maldacena-Shenker-Stanford limit for the bulk Lyapunov exponent in AdS space in the eikonal approximation likely corresponds to a decrease of scrambling time in dual field theory, originating from reduction in the number of degrees of freedom.*

4.3.2 Beyond the eikonal approximation: waves on the string

What is the reason to worry? Even if the scattering is still elastic and happens at high energies and momenta (therefore the overlap of the initial and final state is diagonal in the momenta), it might not be diagonal in the coordinates if the string oscillations are excited during the scattering. These excitations might be relevant for the outcome.¹⁴ However, the quantum mechanics of the string in a non-stationary background is no easy matter and we plan to address it in a separate work. In short, one should write the amplitude (4.18) in the worldsheet theory and then evaluate it in a controlled diagrammatic expansion. For the black hole scrambling scenario, the leading-order stringy corrections are considered in [6]; the Regge (flat-space) limit is the pure gravity black hole scrambling with the Lyapunov exponent $2\pi T$ and scrambling time determined by the large N . We need to do the same for the string action (4.21) but, as we said, we can only give a rough sketch now.

¹⁴With an open string hanging from the boundary to the horizon as in [17] this is not the case, since it stretches along the radial direction and the scattering event — which is mostly limited to near-horizon dynamics because this is where the energy is boosted to the highest values — remains confined to a small segment of the string, whereas any oscillations propagate from end to end. However, a ring string near the horizon is *wholly in the near-horizon region all the time*, and the string excitations may happily propagate along it when the perturbation reaches the area $UV \approx 0$.

The amplitude (4.18) is given by the worldsheet expectation value

$$\mathcal{A} = \prod_i \int d^2 z_i \langle \hat{V}(z_1, \bar{z}_1) \hat{W}(z_2, \bar{z}_2) \hat{V}(z_3, \bar{z}_3) \hat{W}(z_4, \bar{z}_4) \rangle \quad (4.30)$$

with the action (4.20), or (4.21) in the target-space coordinates (\mathbb{T}, X) accommodated to the shock-wave perturbation. Here, we have introduced the usual complex worldsheet coordinates $z = \tau + i\sigma, \bar{z} = \tau - i\sigma$. We thus need to compute a closed string scattering amplitude for the tachyon of the Virasoro-Shapiro type, but with nontrivial target-space metric and consequently with the vertex operators more complicated than the usual plane-wave form. This requires some drastic approximations. We must first expand the non-Gaussian functional integral over the fields $\mathbb{T}(z, \bar{z}), X(z, \bar{z}), \Phi(z, \bar{z})$ perturbatively, and then we can follow [6] and [49] and use the operator-product expansion (OPE) to simplify the vertex operators and decouple the functional integral over the target-space coordinates from the worldsheet integration. First we can use the worldsheet reparametrization to fix as usual $z_1 = \infty, z_2 = z, z_3 = 1, z_4 = 0$. The most relevant regime is that of the highly boosted perturbation near the horizon, with $|z| \sim 1/s$. At leading order in the expansion over \mathbb{T}, X , the action (4.21) decouples the Gaussian functional integral over the (\mathbb{T}, X) coordinates from the pendulum dynamics of the Φ coordinate. We can just as easily use the (U, V) dynamics, with $1/2(\dot{\mathbb{T}}^2 - \dot{X}^2) \mapsto -2\dot{U}\dot{V}$; this is just a linear transformation and the functional integral remains Gaussian. The states in U and V coordinates are just the plane waves with $p_1 = p_3 = p, p_2 = p_4 = q$, but the Φ states are given by some nontrivial wavefunctions $\psi(\Phi)$. Altogether we get

$$\begin{aligned} \mathcal{A} &= \int d^2 z \int DUDVD\Phi \exp \left[-\frac{1}{2\pi\alpha'} \int d^2 z' \left(-2\dot{U}\dot{V} + r_h^2 \left(\dot{\Phi}^2 + n^2 \sin^2 \Phi \right) \right) \right] \hat{V}_1 \hat{W}_2 \hat{V}_3 \hat{W}_4 \\ \hat{V}_{1,3} &= g(U_{1,3}) e^{\mp i p U_1} \psi^\mp(\Phi_{1,3}), \quad \hat{W}_{2,4} = g(V_{2,4}) e^{\mp i q V_{2,4}} \psi^\mp(\Phi_{2,4}), \end{aligned} \quad (4.31)$$

where we denote by the index $i = 1, 2, 3, 4$ the coordinates depending on z_i and the coordinates in the worldsheet action in the first line depend on z' which is not explicitly written out to save space. The higher-order metric corrections in U and V give rise to the weak non-plane-wave dependence of the vertices on U and V , encapsulated in the functions g above. We will disregard them completely, in line with considering the decoupled approximation of the metric as written explicitly in the action in (4.31). The functional integral over U, V is easily performed but the Φ -integral is formidable. However, for small $|z|$, we can expand the ground state solution (2.10) in z, \bar{z} , which corresponds to the linearized oscillatory behavior and the functional integral becomes Gaussian: $\dot{\Phi}^2 + n^2 \sin^2 \Phi \mapsto \dot{\Phi}^2 + n^2 \Phi^2$. With the effective potential for the tachyon $V_{\text{eff}}(\Phi) = n^2 \Phi^2$, the worldsheet propagator takes the form

$$G^\Phi(z, \bar{z}, z', \bar{z}') = K_0(n|z - z'|) \sim \log \frac{n|z - z'|^2}{2}. \quad (4.32)$$

For the plane wave states we take the ansatz $\psi(\Phi) = e^{i\ell\Phi}$, where $\ell = l - i\nu$, with $l \in \mathbb{Z}$ being the angular momentum and $0 < \nu \ll 1$ the correction from the interactions (fortunately we will not need the value of ν). The worldsheet propagator for the flat (U, V) coordinates

has the standard logarithmic form. Now we use the fact that $1/|z| \sim s = pq$ to expand the vertices for \hat{W}_2 and \hat{W}_4 in OPE. The OPE reads

$$:\hat{W}_2\hat{W}_4: \sim \exp(iqz\partial V_2 + iq\bar{z}\bar{\partial}V_2) \exp(i\ell z\partial\Phi_2 + i\ell\bar{z}\bar{\partial}\Phi_2) |z|^{-2-\frac{2\pi\alpha'}{r_h^2}(\ell^2-n^2/2)}, \quad (4.33)$$

which follow from the action of the Laplace operator on the state $e^{i\ell\Phi}$. This finally gives

$$\mathcal{A} = \text{const.} \times \int d^2z : \hat{W}_2\hat{W}_4 : \exp\left(-\frac{\pi\alpha'}{2}pq \log|1-z|^2\right) \exp\left[\frac{\pi\alpha'}{r_h^2}\ell^2(G^\Phi(z)+G^\Phi(1-z))\right]. \quad (4.34)$$

The above integral results in a complicated ratio of the ${}_1F_1$ hypergeometric functions and gamma functions. We still have three possible poles, as in the Virasoro-Shapiro amplitude. In the stringy regime at large pq , the dominant contribution must come from $\ell \sim l = 0$, for the other pole brings us back to the purely gravitational scattering, with $S \propto pq$, whereby the local scrambling rate remains insensitive to n , as we have shown in the eikonal approximation. The stringy pole yields the momentum-integrated amplitude

$$D \sim \int \frac{dp}{p^2} \int \frac{dq}{q^2} \exp\left[-p - q - (pqe^{-2\pi T t})^{1+\frac{\pi\alpha'}{r_h^2}n^2}\right] \sim 1 + \text{const.} \times e^{2\pi T(1+\pi\alpha'n^2)}$$

$$\lambda_{\text{OTOC}} \sim 2\pi T(1 + \pi\alpha'n^2), \quad (4.35)$$

showing that the Lyapunov scale $2\pi T$ is modified (we again take $r_h = 1$ for simplicity). We conclude that *in the strong stringy regime the Lyapunov exponent in dual field theory behaves as $2\pi(1 + \pi\alpha'n^2)T$, differing from the expected chaos bound for nonzero winding numbers n .* Thus, if the classical gravity eikonal approximation does not hold, the modification of the bulk Lyapunov exponent also has an effect on the OTOC decay rate in field theory.

Once again, the above reasoning has several potential loopholes: (1) we completely disregard the higher-order terms in the metric, which couple that radial and transverse dynamics (2) we assume only small oscillations in Φ (3) we disregard the corrections to vertex operators (4) we disregard the corrections to the OPE coefficients. These issues remain for future work.

5 Discussion and conclusions

Our study has brought us to a sharp formal result with somewhat mystifying physical meaning. We have studied classical chaos in the motion of closed strings in black hole backgrounds, and we have arrived, analytically and numerically, at the estimate $\lambda = 2\pi T n$ for the Lyapunov exponent, with n being the winding number of the string. This is a correction by the factor of n of the celebrated chaos bound $\lambda \leq 2\pi T$. However, one should think twice before connecting these things. From the bulk perspective, what we have is different from classical gravity — it includes string degrees of freedom, and no gravity degrees of freedom. Therefore, the fast scrambler hypothesis that the black holes in Einstein gravity exactly saturate the bound is not expected to be relevant for our system

anyway, but the question remains why the bound is modified *upwards* instead of simply being unsaturated (in other words, we would simply expect to get $\lambda < 2\pi T$). The twist is that the Lyapunov exponent in the bulk is related to but in general *distinct* from the Lyapunov exponent in field theory, usually defined in terms of OTOC. Apparently, one just should not uncritically apply the chaos bound proven for the correlation function decay rates in flat-space quantum fields to worldsheet classical string dynamics.

Therefore, it might be that the field theory Lyapunov time does not violate the bound at all. The timescale of OTOC decay for a field theory dual to the fluctuating string is calculated in [17]: OTOC equals the expectation value of the scattering operator for bulk strings with appropriate boundary conditions. The field-theory Lyapunov time is then determined by the phase shift of the collision. In particular, [17] finds the saturated bound $\lambda = 2\pi T$ as following from the fact that the phase shift is proportional to the square of the center-of-mass energy. On the other hand, [6] predicts that the Lyapunov exponent is lower than the bound when stringy effects are considered. We have done first a completely classical calculation of OTOC and have found, expectedly perhaps, that the $2\pi T$ bound is exactly obeyed. Then we have followed the approximate scheme of [49] to include the one-loop closed string tachyon amplitude as the simplest (and hopefully representative enough?!) stringy process. For a ring string background, this gives an *increased* value for the field-theory Lyapunov rate, yielding some credit to the interpretation that complicated string configurations encode for strongly nonlocal operators, which might indeed violate the bound. But as we have explained, the approximations we took are rather drastic. We regard a more systematic study of loop effects in string chaos as one of the primary tasks for future work.

To gain some more feeling on the dual field theory, we have looked also at the Regge trajectories. In one configuration, the strings that violate the bound n times are precisely those whose Regge trajectory has the slope n times smaller than the leading one (and thus for $n = 1$ the original bound is obeyed and at the same time we are back to the leading Regge trajectory). In another configuration, the strings that violate the bound describe no Regge trajectory at all. However, it is *very* hard to say anything precise about the gauge theory operators at finite temperature. Deciphering which operators correspond to our strings is an important but very ambitious task; we can only dream of moving toward this goal in very small steps. What we found so far makes it probable that complicated, strongly non-local operators correspond to the bound-violating strings, so that (as explained in the original paper [1]) their OTOC cannot be factorized and the bound is not expected to hold.¹⁵

¹⁵In relation to the gauge/string duality it is useful to look also at the gauge theories with N_f flavors added, which corresponds to the geometry deformed by N_f additional D-branes in the bulk. In [50] it was found that the system becomes nonintegrable in the presence of the flavor branes (expectedly, as it becomes non-separable), but the Lyapunov exponent does not grow infinitely with the number of flavors, saturating instead when the number of colors N_c and the number of flavors become comparable. This is expected, as the D3-D7 brane background of [50] formally becomes separable again when $N_f/N_c \rightarrow \infty$ (although in fact this regime cannot be captured, the calculation of the background ceases to be valid in this case). In our case the winding number n is a property of the string solution, not geometry, and the Hamiltonian (2.9) seems to have no useful limit for $n \rightarrow \infty$, thus we do not expect the estimate $2\pi T n$ will saturate.

Preparing the final version of the paper, we have learned also of the work [52] where the n -point OTOCs are studied following closely the logic of [1] and the outcome is a factor of n enlargement, formally the same as our result. This is very interesting but, in the light of the previous paragraph, we have no proof that this result is directly related to ours. It certainly makes sense to investigate if the winding strings are obtained as some limit of the gravity dual for the n -point correlations functions. We know that n -point functions in AdS/CFT are a complicated business. The Witten diagrams include bulk propagators carrying higher spin fields that might in turn be obtained as string excitations. Just how far can one go in making all this precise we do not know for now.

In relation to [15, 16] one more clarifying remark should be given. In these works, particles in the vicinity of the horizon are found to exhibit chaos (either saturating the bound or violating it, depending on the spin of the background field). At first glance, this might look inconsistent with our finding that for $n = 0$, when the string degenerates to a particle, no chaos occurs; after all, we know that geodesic motion in the background of spherically symmetric black holes is integrable, having a full set of the integrals of motion. But in fact there is no problem, because in [15, 16] an additional external potential (scalar, vector, or higher-spin) is introduced that keeps the particle at the horizon, balancing out the gravitational attraction. Such a system is of course not integrable anymore, so the appearance of chaos is expected. The modification of the bound in the presence of higher-spin fields might have to do with the findings [51] that theories with higher-spin fields can only have gravity duals in very restricted situations (in particular, higher spin CFTs with a sparse spectrum and large central charge or, roughly speaking, massive higher spin fields, are problematic).

Another task on the to-do list, entirely doable although probably demanding in terms of calculations, is the (necessarily approximate) calculation of the black hole scattering matrix, i.e., the backreaction of the black hole upon scattering or absorbing a string, along the lines of [7]. In this paper we have worked in the probe limit (no backreaction), whereas the true scrambling is really the relaxation time of the black hole (the time it needs to become hairless again), which cannot be read off solely from the Lyapunov time; this is the issue we also mentioned in the Introduction, that local measures of chaos like the Lyapunov exponent do not tell the whole story of scrambling. Maybe even a leading-order (tree-level) backreaction calculation can shed some light on this question.

Acknowledgments

I am grateful to K. Schalm and M. V. Medvedyeva for helpful discussions. I also thank to D. Giataganas, L. A. Pando Zayas and J. Kasi for insightful remarks. Special thanks goes to the anonymous referee for his stimulating question which has improved the quality of the final manuscript. This work has made use of the excellent Sci-Hub service. Work at the Institute of Physics is funded by Ministry of Education, Science and Technological Development, under grant no. OI171017.

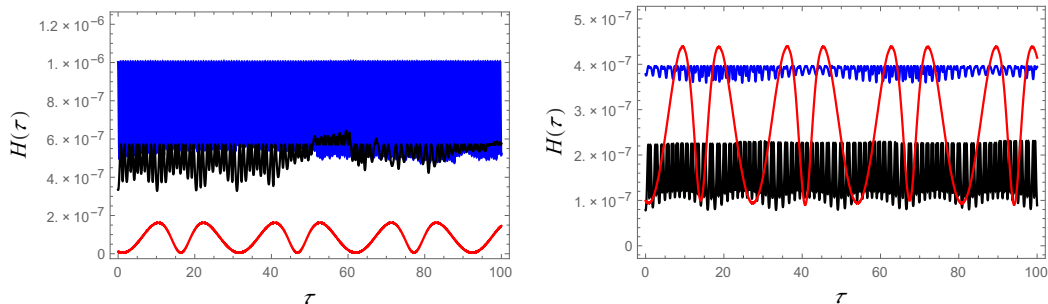


Figure 6. Check of the Hamiltonian constraint $H = 0$ during an integration for the spherical, planar and hyperbolic black hole (black, blue, red respectively), at temperature $T = 0.01$ (left) and $T = 0.10$ (right). The accuracy of the constraint is a good indicator of the overall integration accuracy, it is never above 10^{-6} and has no trend of growth but oscillates.

A Summary of the numerics

We feel it necessary to give a short account of the numerical methods used. The string equations of motion (2.6)–(2.8) present us with a system of two ordinary second-order differential equations with a constraint. This numerical calculation is not very difficult, and it would be trivial if it were not for two complicating factors. First, the constraint itself is the main complication; it is non-holonomic and cannot be easily eliminated. Second, the system is rather stiff, with \dot{R} in particular varying for several orders of magnitude. We did the integration in the `Mathematica` package, using mostly the `NDSolve` routine, and controlling both the relative and the absolute error during the calculations. The constraint problem is solved serendipitously by ensuring that the initial conditions satisfy the constraint and then adjusting the required absolute and relative error tolerance so that the constraint remains satisfied. A priori, this is a rather unlikely way to succeed but we find it works in most cases. Only in a few integrations we needed to write a routine which shoots for the condition $H = 0$ at every timestep, using the `NDSolve` routine in the solver; the shooting itself we wrote using the tangent method which is handier for this problem than the built-in routines. The usual analytic way, making use of the Lagrange multipliers, seems completely unsuitable for numerical implementation in this problem. In figure 6 we show the evolution of the constraint for a few examples, demonstrating the stability of the integration. We have also checked that the functions $R(\tau), \Phi(\tau)$ converge toward definite values as the precision and accuracy (relative and absolute error per step) are varied.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.


References

- [1] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [[arXiv:1503.01409](https://arxiv.org/abs/1503.01409)] [[INSPIRE](https://inspirehep.net/literature/1503014)].
- [2] Y. Sekino and L. Susskind, *Fast scramblers*, *JHEP* **10** (2008) 065 [[arXiv:0808.2096](https://arxiv.org/abs/0808.2096)] [[INSPIRE](https://inspirehep.net/literature/786018)].

- [3] N. Lashkari, D. Stanford, M. Hastings, T. Osborne and P. Hayden, *Towards the fast scrambling conjecture*, *JHEP* **04** (2013) 022 [[arXiv:1111.6580](#)] [[INSPIRE](#)].
- [4] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [[arXiv:1306.0622](#)] [[INSPIRE](#)].
- [5] D.A. Roberts, D. Stanford and L. Susskind, *Localized shocks*, *JHEP* **03** (2015) 051 [[arXiv:1409.8180](#)] [[INSPIRE](#)].
- [6] S.H. Shenker and D. Stanford, *Stringy effects in scrambling*, *JHEP* **05** (2015) 132 [[arXiv:1412.6087](#)] [[INSPIRE](#)].
- [7] J. Polchinski, *Black hole S matrix*, [arXiv:1505.08108](#) [[INSPIRE](#)].
- [8] T. Scaffidi and E. Altman, *Chaos in a classical limit of the Sachdev-Ye-Kitaev model*, *Phys. Rev. B* **100** (2019) 155128 [[arXiv:1711.04768](#)] [[INSPIRE](#)].
- [9] S. Sachdev and J. Ye, *Gapless spin fluid ground state in a random, quantum Heisenberg magnet*, *Phys. Rev. Lett.* **70** (1993) 3339 [[cond-mat/9212030](#)] [[INSPIRE](#)].
- [10] O. Parcolet and A. Georges, *Non-Fermi liquid regime of a doped Mott insulator*, *Phys. Rev. B* **59** (1998) 5341 [[cond-mat/9806119](#)].
- [11] A. Kitaev, *A simple model of quantum holography (part 1)*, talk at KITP, April 7, 2015, <http://online.kitp.ucsb.edu/online/entangled15/kitaev/>.
- [12] A. Kitaev, *A simple model of quantum holography (part 2)*, talk at KITP, May 27, 2015, <http://online.kitp.ucsb.edu/online/entangled15/kitaev2/>.
- [13] E. Marcus and S. Vandoren, *A new class of SYK-like models with maximal chaos*, *JHEP* **01** (2019) 166 [[arXiv:1808.01190](#)] [[INSPIRE](#)].
- [14] A.L. Fitzpatrick and J. Kaplan, *A quantum correction to chaos*, *JHEP* **05** (2016) 070 [[arXiv:1601.06164](#)] [[INSPIRE](#)].
- [15] K. Hashimoto and N. Tanahashi, *Universality in chaos of particle motion near black hole horizon*, *Phys. Rev. D* **95** (2017) 024007 [[arXiv:1610.06070](#)] [[INSPIRE](#)].
- [16] S. Dalui, B.R. Majhi and P. Mishra, *Presence of horizon makes particle motion chaotic*, *Phys. Lett. B* **788** (2019) 486 [[arXiv:1803.06527](#)] [[INSPIRE](#)].
- [17] J. de Boer, E. Lladrés, J.F. Pedraza and D. Vegh, *Chaotic strings in AdS/CFT*, *Phys. Rev. Lett.* **120** (2018) 201604 [[arXiv:1709.01052](#)] [[INSPIRE](#)].
- [18] J.R. David, S. Khetrpal and S.P. Kumar, *Local quenches and quantum chaos from perturbations*, *JHEP* **10** (2017) 156 [[arXiv:1707.07166](#)] [[INSPIRE](#)].
- [19] P. Basu and L.A. Pando Zayas, *Analytic nonintegrability in string theory*, *Phys. Rev. D* **84** (2011) 046006 [[arXiv:1105.2540](#)] [[INSPIRE](#)].
- [20] A. Stepanchuk and A.A. Tseytlin, *On (non)integrability of classical strings in p-brane backgrounds*, *J. Phys. A* **46** (2013) 125401 [[arXiv:1211.3727](#)] [[INSPIRE](#)].
- [21] Y. Chervonyi and O. Lunin, *(Non)-integrability of geodesics in D-brane backgrounds*, *JHEP* **02** (2014) 061 [[arXiv:1311.1521](#)] [[INSPIRE](#)].
- [22] C. Núñez, J.M. Penín, D. Roychowdhury and J. Van Gorsel, *The non-integrability of strings in massive type IIA and their holographic duals*, *JHEP* **06** (2018) 078 [[arXiv:1802.04269](#)] [[INSPIRE](#)].
- [23] D. Giataganas, L.A. Pando Zayas and K. Zoubos, *On marginal deformations and non-integrability*, *JHEP* **01** (2014) 129 [[arXiv:1311.3241](#)] [[INSPIRE](#)].

- [24] N. Beisert et al., *Review of AdS/CFT Integrability: An Overview*, *Lett. Math. Phys.* **99** (2012) 3 [[arXiv:1012.3982](#)] [[INSPIRE](#)].
- [25] A.V. Frolov and A.L. Larsen, *Chaotic scattering and capture of strings by a black hole*, *Class. Quant. Grav.* **16** (1999) 3717 [[gr-qc/9908039](#)] [[INSPIRE](#)].
- [26] L.A. Pando Zayas and C.A. Terrero-Escalante, *Chaos in the gauge/gravity correspondence*, *JHEP* **09** (2010) 094 [[arXiv:1007.0277](#)] [[INSPIRE](#)].
- [27] P. Basu and L.A. Pando Zayas, *Chaos rules out integrability of strings on $AdS_5 \times T^{1,1}$* , *Phys. Lett.* **B 700** (2011) 243 [[arXiv:1103.4107](#)] [[INSPIRE](#)].
- [28] P. Basu, D. Das and A. Ghosh, *Integrability lost: Chaotic dynamics of classical strings on a confining holographic background*, *Phys. Lett.* **B 699** (2011) 388 [[arXiv:1103.4101](#)] [[INSPIRE](#)].
- [29] P. Basu, P. Chaturvedi and P. Samantray, *Chaotic dynamics of strings in charged black hole backgrounds*, *Phys. Rev.* **D 95** (2017) 066014 [[arXiv:1607.04466](#)] [[INSPIRE](#)].
- [30] Y. Asano, D. Kawai, H. Kyono and K. Yoshida, *Chaotic strings in a near Penrose limit of $AdS_5 \times T^{1,1}$* , *JHEP* **08** (2015) 060 [[arXiv:1505.07583](#)] [[INSPIRE](#)].
- [31] Y. Asano, H. Kyono and K. Yoshida, *Melnikov's method in string theory*, *JHEP* **09** (2016) 103 [[arXiv:1607.07302](#)] [[INSPIRE](#)].
- [32] D. Giataganas and K. Sfetsos, *Non-integrability in non-relativistic theories*, *JHEP* **06** (2014) 018 [[arXiv:1403.2703](#)] [[INSPIRE](#)].
- [33] T. Ishii, K. Murata and K. Yoshida, *Fate of chaotic strings in a confining geometry*, *Phys. Rev.* **D 95** (2017) 066019 [[arXiv:1610.05833](#)] [[INSPIRE](#)].
- [34] R.B. Mann, *Pair production of topological anti-de Sitter black holes*, *Class. Quant. Grav.* **14** (1997) L109 [[gr-qc/9607071](#)] [[INSPIRE](#)].
- [35] D.R. Brill, J. Louko and P. Peldan, *Thermodynamics of (3+1)-dimensional black holes with toroidal or higher genus horizons*, *Phys. Rev.* **D 56** (1997) 3600 [[gr-qc/9705012](#)] [[INSPIRE](#)].
- [36] L. Vanzo, *Black holes with unusual topology*, *Phys. Rev.* **D 56** (1997) 6475 [[gr-qc/9705004](#)] [[INSPIRE](#)].
- [37] D. Birmingham, *Topological black holes in anti-de Sitter space*, *Class. Quant. Grav.* **16** (1999) 1197 [[hep-th/9808032](#)] [[INSPIRE](#)].
- [38] R.B. Mann, *Topological black holes — outside looking in*, *Annals Israel Phys. Soc.* **13** (1997) 311 [[gr-qc/9709039](#)] [[INSPIRE](#)].
- [39] W.L. Smith and R.B. Mann, *Formation of topological black holes from gravitational collapse*, *Phys. Rev.* **D 56** (1997) 4942 [[gr-qc/9703007](#)] [[INSPIRE](#)].
- [40] Y.C. Ong, *Hawking evaporation time scale of topological black Holes in anti-de Sitter spacetime*, *Nucl. Phys.* **B 903** (2016) 387 [[arXiv:1507.07845](#)] [[INSPIRE](#)].
- [41] Y. Chen and E. Teo, *Black holes with bottle-shaped horizons*, *Phys. Rev.* **D 93** (2016) 124028 [[arXiv:1604.07527](#)] [[INSPIRE](#)].
- [42] C.V. Johnson and F. Rosso, *Holographic heat engines, entanglement entropy and renormalization group flow*, *Class. Quant. Grav.* **36** (2019) 015019 [[arXiv:1806.05170](#)] [[INSPIRE](#)].
- [43] R. Emparan, *AdS/CFT duals of topological black holes and the entropy of zero-energy states*, *JHEP* **06** (1999) 036 [[hep-th/9906040](#)] [[INSPIRE](#)].

- [44] A.L. Larsen, *Chaotic string-capture by black hole*, *Class. Quant. Grav.* **11** (1994) 1201 [[hep-th/9309086](#)] [[INSPIRE](#)].
- [45] A.E. Motter, *Relativistic chaos is coordinate invariant*, *Phys. Rev. Lett.* **91** (2003) 231101 [[gr-qc/0305020](#)] [[INSPIRE](#)].
- [46] N.L. Balazs and A. Voros, *Chaos on the pseudosphere*, *Phys. Rept.* **143** (1986) 109 [[INSPIRE](#)].
- [47] S.S. Gubser, I.R. Klebanov and A.M. Polyakov, *A semi-classical limit of the gauge/string correspondence*, *Nucl. Phys. B* **636** (2002) 99 [[hep-th/0204051](#)] [[INSPIRE](#)].
- [48] H.J. de Vega and I.L. Egusquiza, *Planetoid string solutions in 3 + 1 axissymmetric spaces*, *Phys. Rev. D* **54** (1996) 7513 [[hep-th/9607056](#)] [[INSPIRE](#)].
- [49] R.C. Brower, J. Polchinski, M.J. Strassler and C.-I. Tan, *The pomeron and gauge/string duality*, *JHEP* **12** (2007) 005 [[hep-th/0603115](#)] [[INSPIRE](#)].
- [50] D. Giataganas and K. Zoubos, *Non-integrability and chaos with unquenched flavor*, *JHEP* **10** (2017) 042 [[arXiv:1707.04033](#)] [[INSPIRE](#)].
- [51] E. Perlmutter, *Bounding the space of holographic CFTs with chaos*, *JHEP* **10** (2016) 069 [[arXiv:1602.08272](#)] [[INSPIRE](#)].
- [52] P. Basu and K. Jaswin, *Higher point OTOCs and the bound on chaos*, [arXiv:1809.05331](#) [[INSPIRE](#)].

Spontaneous isotropy breaking for vortices in nonlinear left-handed metamaterialsTrivko Kukolj ^{*}*Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Serbia and Department of Physics, Faculty of Sciences, University of Novi Sad, Trg Dositeja Obradovića 4, Novi Sad, Serbia*Mihailo Čubrović [†]*Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Serbia*

(Received 15 July 2019; revised manuscript received 11 October 2019; published 25 November 2019)

We explore numerically and analytically the pattern formation and symmetry breaking of beams propagating through left-handed (negative) nonlinear metamaterials. When the input beam is a vortex with topological charge (winding number) Q , the initially circular (isotropic) beam acquires the symmetry of a polygon with Q , $2Q$, or $3Q$ sides, depending on the details of the response functions of the material. Within an effective field-theory model, this phenomenon turns out to be a case of spontaneous dynamical symmetry breaking described by a Landau-Ginzburg functional. Complex nonlinear dependence of the magnetic permittivity on the magnetic field of the beam plays a central role, as it introduces branch cuts in the mean-field solution, and permutations among different branches give rise to discrete symmetries of the patterns. By considering loop corrections in the effective Landau-Ginzburg field theory we obtain reasonably accurate predictions of the numerical results.

DOI: [10.1103/PhysRevA.100.053853](https://doi.org/10.1103/PhysRevA.100.053853)**I. INTRODUCTION**

The idea of a material with negative refraction index was first considered long before it could be realized in experiment, in the now famous paper by Veselago [1], in 1968. He considered a material with negative electric permeability ϵ and magnetic permittivity μ , and predicted a number of interesting properties in such systems, among them negative refraction. Only much later did it become possible to combine elements with negative ϵ and negative μ at a microscopic level, as a composite metamaterial. First experimental realizations were reported in [2,3]. Negativity, or left-handedness, is typically only achieved in a narrow frequency range, close to the resonant frequency of the conductive elements of the metamaterial. This was the original motivation for studying nonlinear effects in these systems. Nonlinearities can be strengthened by appropriate design at the microscopic level. The study of nonlinear phenomena in metamaterials started with [4]. This has become a broad and important field in metamaterials research [5]. Nonlinear phenomena like solitons [6,7], nonlinear surface waves [8], modulational instability [9,10], and ultrashort pulses [11] were observed. Other work in left-handed metamaterials relevant for our paper includes, among others, [12–20]. We have no intention of being exhaustive in this short review of the literature; we merely mention the results we have directly used or found particularly inspiring.

The focus of our paper is the dynamics of symmetry breaking in intensity patterns of electromagnetic waves propagating through a left-handed nonlinear metamaterial. Numerical solutions of the equations of motion reveal that circular

(usually Gaussian) input beams turn into polygonal patterns, with some discrete symmetry. This fits the textbook notion of symmetry breaking, more specifically dynamical symmetry breaking. The general theory of dynamical criticality is by now well developed [21] and has been applied to numerous systems [22]. In [22], a systematic theory of isotropy-breaking transition is presented, though mainly for periodic and quasiperiodic structures (convection in fluids, fluctuations in quasicrystals). The basic mechanism is that the system develops momentum eigenmodes of a fixed module but with multiple discrete directions on the sphere $|\mathbf{k}| = \text{const}$ in momentum space. In nonlinear negative materials, the situation is complicated by the strong frequency dependence of the magnetic permittivity but the same basic logic remains. At a fundamental level, this situation can be understood from the viewpoint of a spatially nonuniform Landau-Ginzburg theory. Quantitative accuracy is, however, hard to achieve; this requires cumbersome perturbative calculations. Ultimately, numerical work is the best way to describe the patterns in detail; they look like polygons or, occasionally, necklaces, with C_{3Q} , C_{2Q} , or C_Q symmetry, depending on the parameter regime; here, Q is the topological charge of the beam, a property we will discuss in detail in the next paragraph. The paper [10] is very important in this context: it starts from the model derived in [9] and studies mainly necklace configurations, which consist of discrete beads (spots of high intensity) distributed more or less uniformly along a circle. The authors find the same C_{3Q} symmetry that we see. Our goal is to gain a detailed understanding of the phenomenon, and move beyond single beams toward collective behavior and interactions.

We have chosen to study this phenomenon on vortices, topologically nontrivial solutions where the phase of the

^{*}trivko98@gmail.com[†]mcubrovic@gmail.com

complex electric and magnetic field winds one or more times along a closed line encircling the vortex core. Vortices appear in many systems described by a complex field, i.e., a field with $U(1)$ phase invariance [23,24]. In optics, this is just the complex beam envelope of the electric and magnetic field. The phase of any complex wave function or field can wind along some closed line around a defect, forming a vortex. Famously, vortices may coexist with the superconducting order [$U(1)$ symmetry breaking] in type-II superconductors or they may exist only in the normal phase, upon destroying the superconductivity (type I). Pattern-forming systems like fluids and soft matter often have rich vortex dynamics [22]. Other examples of vortex matter in nature arise in liquid helium [25], Bose-Einstein condensates [26], and magnetic systems [27]. In two spatial dimensions, interactions among the vortices lead to a vortex unbinding phase transition of infinite order found by Berezinsky, Kosterlitz, and Thouless for the planar XY model [28]. We study a three-dimensional metamaterial but with an elongated geometry, so we treat it as a $2 + 1$ -dimensional system (the x and y coordinates are spatial dimensions and the z direction has the formal role of time). We therefore have a similar situation to the XY model but with different equations of motion and different phenomena.

In addition to direct numerical and analytical study of the equations of motion, we also propose an effective field-theory Lagrangian which gives slightly different equations but captures the key properties of the system. The Lagrangian form makes it easier to understand some of the phenomenology we find in numerical simulations; the foundations of the symmetry breaking are obtained from this model in a natural way. Numerical work is done with original equations of motion, as they are directly grounded in the microscopic physics. The Lagrangian is just a phenomenological tool to facilitate the theoretical understanding. It is difficult (and perhaps impossible) to package the exact original equations in a Lagrangian form because the system is strongly nonlinear *and* dissipative. Dissipative systems can be encapsulated in a Lagrangian (our Lagrangian is also dissipative) but with some limitations, and there is certainly no general method to write down a Lagrangian for a broad class of dissipative systems.

The structure of the paper is the following. In the next section we describe the model of a nonlinear left-handed metamaterial, following closely the wave propagation equations used in previous research, e.g., in [4,6,7] and others, which correspond to a specific experimentally realizable metamaterial. We also formulate and motivate the field-theory model of the system. In the third section, we describe our numerical findings, above all the anisotropy of the intensity patterns. The fourth section offers the theoretical explanation for the patterns: first by a direct approximate solution of the propagation equations, and then also from field theory, which makes the physics of the symmetry breaking particularly clear. In the fifth section we briefly discuss how to check our predictions in experiment and how prominent the effects of symmetry breaking are compared to other possible instabilities in realistic metamaterials. The last section sums up the conclusions. We have included some long calculations in the Appendices.

II. WAVE EQUATIONS IN A NONLINEAR LEFT-HANDED METAMATERIAL

We adopt the model of [4,7] to describe a left-handed metamaterial with a nonlinear response. Microscopically, the material is realized as a lattice of split-ring resonators and wires. In the terahertz range, this is an experimentally well-studied system [3]. In [6], a detailed microscopic derivation is given, starting from the current transport equations in the resonator-wire system. The outcome is a nonlinearity similar to that postulated phenomenologically in [4]. We adopt essentially the same model, described by the electric permeability ϵ and the magnetic permittivity μ :

$$\epsilon(E, E^\dagger) \equiv \epsilon(|E|^2) = (\epsilon_{D0} + \alpha|E|^2) \left(1 - \frac{\omega_0^2}{\omega(\omega + i\gamma)} \right), \quad (1)$$

$$\mu(H, H^\dagger) \equiv \mu(|H|^2) = 1 + \frac{F\omega^2}{\omega_{0NL}^2(|H|^2) - \omega^2 + i\Gamma\omega}, \quad (2)$$

with $\alpha = 1$ or -1 for self-focusing or self-defocusing nonlinearity, respectively. Frequency is denoted by ω and ϵ_{D0} is the linear part of the permittivity. By F , γ , and Γ we denote the filling factor of the material and the electric and magnetic damping coefficients. Equations (1) and (2) allow us to model also the real (lossless) dielectric response by putting $\gamma = 0$. For the magnetic field, the permittivity will in general stay complex even for $\Gamma = 0$, as the nonlinear frequency of the resonator rings ω_{0NL} can always have a nonzero imaginary part. This frequency is related to the magnetic field through the relation ($X \equiv \omega_{0NL}/\omega_0$):

$$|H|^2 = \alpha A^2 \frac{(1 - X^2)[(X^2 - \Omega^2)^2 + \Omega^2\gamma^2]}{X^6}, \quad (3)$$

where $\Omega \equiv \omega/\omega_0$, ω_0 is the eigenfrequency of the rings, and A is a parameter which can be derived microscopically [4,6,7]; for our purposes, it can be treated just as a phenomenological parameter. This cubic equation yields three branches for ω_{0NL}^2 . All these branches are physical and correspond to different possible nonlinear oscillations [7]. Now the equations of motion are just the Maxwell equations in a medium described by (1) and (2), in the approximation of slowly changing beam envelopes. We assume an elongated (cylindrical or parallelepipedal) slab of metamaterial, so we can employ the paraxial beam approximation (e.g., [29]). The beam is initially collimated along the longitudinal axis z and focuses or defocuses slowly in the transverse x - y plane due to the nonlinearity of the material. The electric and magnetic field $\hat{E}(t; x, y, z)$ and $\hat{H}(t; x, y, z)$ are directed along the z axis. From now on, the speed of light is put to unity, $c = 1$. All the steps in deriving the nonlinear Schrödinger-like equation are well known so we merely state the final result here, which is quite close to the equations used in [13] in $1 + 1$ dimension, or the equations found in [9–11]. Full derivation can be found in Appendix A. The equations of motion turn out to be

$$-\frac{i}{b}\partial_z E = \nabla_\perp^2 E + [\omega^2\epsilon(|E|^2)\mu(|H|^2) - k^2]E - \frac{\nabla_\perp\mu(|H|^2)}{\mu(|H|^2)}\nabla_\perp E - i\frac{\partial_z\mu(|H|^2)}{2\mu(|H|^2)}E, \quad (4)$$

$$-\frac{i}{b}\partial_z H = \nabla_\perp^2 H + [\omega^2\epsilon(|E|^2)\mu(|H|^2) - k^2]H. \quad (5)$$

Here, $\nabla_{\perp} \equiv (\partial_x, \partial_y)$ is the nabla operator in the transverse plane, k is the wave vector along the z direction, and b is the characteristic propagation length along the z axis. Equations of motion (4) and (5) together with the equations (1) and (3) for the permittivities contain the following five parameters: ϵ_{D0} , F , Γ , γ , and ω_0 . Realistic values for all the parameters are discussed in [7]. The natural length scale of the model is dominated by the $1/\omega_0$ scale. Dimensional analysis of the terms on the right-hand side of (4) determines the length scale b in (4) and (5) as $b \sim 1/\omega_0$. Both in analytical and in numerical calculations, we express the transverse coordinates (x, y) in millimeters but the longitudinal coordinate z is often stated in units of b . This is because the length scales of all patterns in the transverse plane are similar whereas the propagation lengths along z can vary by an order of magnitude as γ and Ω are varied, so it is more natural to express them in terms of the characteristic distance b .

A. field-theoretical model

For some theoretical considerations it is useful to formulate a Lagrangian (gradient) model which captures the essential features of the equations of motion (4) and (5). As it often happens in studies of complex nonlinear pattern-forming systems, we cannot easily write the original equations in such a form. Instead, we construct a field theory which yields equations of motion somewhat different from the original ones but which still give the same phenomenology, and are able to explain the results of numerical calculations with the equations (4) and (5).

Let us think what such a field theory would look like. The terms with the gradient of magnetic permittivity obviously introduce dissipation, which physically originates from the losses in the inductive rings of the metamaterial. In general, dissipative systems do not have a Lagrangian, although a number of generalized Lagrangian approaches exist for dissipative systems: either with more general functional forms of the Lagrangian, or with a dissipative function in addition to the Lagrangian, or with extra degrees of freedom [30,31]. We will take the first, most conventional of the three approaches: we will consider a conventional Lagrangian (no dissipative function, no extra degrees of freedom) which gives slightly generalized equations of motion compared to (4) and (5), with dissipative terms for both electric and magnetic fields coming from the complex terms in the effective potential. The effective action reads

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_E + \mathcal{L}_H, \\ \mathcal{L}_E &= \frac{i}{2\mu(|H|^2)} (E \partial_z E^\dagger - E^\dagger \partial_z E) \\ &\quad + \frac{|\nabla_{\perp} E|^2}{\mu(|H|^2)} + \frac{k^2 |E|^2}{\mu(|H|^2)} - \omega^2 \epsilon(|E|^2) |E|^2, \\ \mathcal{L}_H &= \frac{i}{2\epsilon(|E|^2)} (H \partial_z H^\dagger - H^\dagger \partial_z H) \\ &\quad + \frac{|\nabla_{\perp} H|^2}{\epsilon(|E|^2)} + \frac{k^2 |H|^2}{\epsilon(|E|^2)} - \omega^2 \int_0^{HH^\dagger} dx \mu(x). \end{aligned} \quad (6)$$

The last term in \mathcal{L}_E equals $-\omega^2 \int_0^{EE^\dagger} dx \epsilon(x)$, analogously to the corresponding term in \mathcal{L}_H , but since ϵ is polynomial in

$E^\dagger E$ the integral can be solved explicitly. Now (6) gives the equations of motion:

$$-\frac{i}{b} \partial_z E = \nabla_{\perp}^2 E + [\epsilon(|E|^2) \mu(|H|^2) - k^2] E - \frac{i \partial_z \mu(|H|^2)}{\mu(|H|^2)} E - \frac{\nabla_{\perp} \mu(|H|^2)}{\mu(|H|^2)} \nabla_{\perp} E - \Phi_H, \quad (7)$$

$$-\frac{i}{b} \partial_z H = \nabla_{\perp}^2 H + [\epsilon(|E|^2) \mu(|H|^2) - k^2] H - \frac{i \partial_z \epsilon(|E|^2)}{\epsilon(|E|^2)} H - \frac{\nabla_{\perp} \epsilon(|E|^2)}{\epsilon(|E|^2)} \nabla_{\perp} H - \Phi_E, \quad (8)$$

where $\Phi_{E,H}$ are related to the fluxes of the electric and magnetic field (prime denotes the derivative of ϵ and μ with respect to their arguments $E^\dagger E$ and $H^\dagger H$):

$$\begin{aligned} \Phi_H &= \frac{\epsilon'(|E|^2)}{\epsilon^2(|E|^2)} \left(\frac{i}{2} (H \partial_z H^\dagger - H^\dagger \partial_z H) \right. \\ &\quad \left. + |\nabla_{\perp} H|^2 + k^2 |H|^2 \right), \end{aligned} \quad (9)$$

$$\begin{aligned} \Phi_E &= \frac{\mu'(|H|^2)}{\mu^2(|H|^2)} \left(\frac{i}{2} (E \partial_z E^\dagger - E^\dagger \partial_z E) \right. \\ &\quad \left. + |\nabla_{\perp} E|^2 + k^2 |E|^2 \right). \end{aligned} \quad (10)$$

These are the extra terms compared to the physical equations (4) and (5).¹ Inserting $\partial_z E^\pm$ from the equations of motion (7) and (8) into the above we derive

$$\Phi_E = \frac{\mu'}{\mu} \nabla_{\perp} \left(\frac{E \nabla_{\perp} E^\dagger - E^\dagger \nabla_{\perp} E}{\mu} \right), \quad (11)$$

and analogously for Φ_H , with $\epsilon \leftrightarrow \mu$, $E \leftrightarrow H$. This term is proportional to a total derivative, and is therefore related to the flux $(E \nabla_{\perp} E^\dagger - E^\dagger \nabla_{\perp} E)/\mu$. For slowly changing ϵ and μ , which is often the case in our system (i.e., for ϵ' , $\mu' \ll \epsilon$, μ), this term is small, which partly justifies the choice (6) for the Lagrangian. But the ultimate justification, as it frequently happens, is that *a posteriori* we will find that this model is able to explain the features observed in the numerics. Therefore we will not try to interpret the term (11) in detail.

III. GEOMETRY AND STABILITY OF VORTICES

We will now sum up our numerical results which demonstrate the breaking of the circular symmetry of the vortex beams and their decay during the propagation. We always start from a Gaussian input beam with a topological charge Q , of the form $E(r, \phi; z=0) = E_0 \times e^{-r^2/2\sigma^2} e^{iQ\phi}$ and analogously for the magnetic field, with amplitude H_0 but with the same vortex charge Q . Therefore, we always give an exact vortex as an input. The parameters of the model were chosen so that the permittivities ϵ and μ , given by (1) and (2), respectively, are of order unity. This serves to limit the dissipative effects, so that the propagation along the longitudinal direction can be clearly observed. Same phenomena are found for arbitrary values of

¹The dissipative term proportional to $\nabla_{\perp} H$ in (8) is also absent in the original equations, but that one is easy to interpret: we make both \mathcal{L}_E and \mathcal{L}_H complex, so both fields have dissipative dynamics.

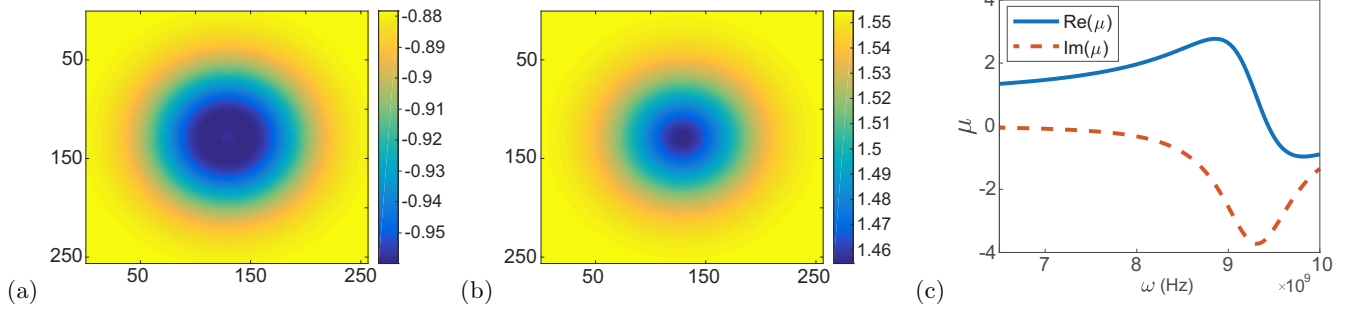


FIG. 1. The radial profile of μ for a left-handed medium (a) and a right-handed medium (b), for a vortex of charge $Q = 1$. The profiles are radially symmetric in accordance with the fact that μ depends strictly on the magnitude of the magnetic-field vector $|H|^2$. The real (blue) and imaginary (red) values of the complex permeability μ vs the frequency of the beam ω are displayed in (c). For frequencies higher than the eigenfrequency of the resonator rings ω_0 , the real part of the permeability is negative, essentially yielding a left-handed medium. The figure is made for dissipative ϵ ; for lossless ϵ the behavior is similar.

ϵ and μ but on a different length scale. We do not aim at a stamp-collecting exhaustive description of patterns for all possible parameter values, so we will focus on just a few relevant cases. We are mainly interested in left-handed materials ($\epsilon, \mu < 0$) and how they compare to right-handed ones, so for the dielectric constant we always choose the self-defocusing Kerr nonlinearity ($\alpha = -1$) with a linear part $\epsilon_{D0} = 12.8$, which has both a left-handed and a right-handed regime. To check the effects of dissipation, we either adopt $\gamma = 0$ in (1), i.e., the lossless case, or we tune γ so that $\omega_0^2/(\omega^2 + i\gamma\omega) = 1/2$. In other words, we impose either $\text{Im}\epsilon = 0$ or $\text{Im}\epsilon = \text{Re}\epsilon$. This is for simplicity and to avoid probing a huge parameter space for all possible γ values; from now on we will call these cases simply lossless ϵ and dissipative ϵ . The filling factor is $F = 0.4$ and the magnetic dampening coefficient is $\Gamma = 10^9$ Hz; these values are kept fixed in all calculations. Numerical calculations are performed with an operator split algorithm described in detail in the Appendices of [32].

The nonlinear frequency of the oscillator rings is obtained as a solution to (3). Of the three branches of the solution, we take the one that yields a negative real value of μ for $\omega > \omega_0$ (Fig. 1). We have freely taken $\omega = 9.8 \times 10^9$ Hz to represent a left-handed medium, and $\omega = 7.0 \times 10^9$ Hz to represent a right-handed medium. The transverse profiles are displayed in Fig. 1. We see there is a well-defined left-handed regime.

Now we discuss the transverse intensity profile for different initial beam configurations, with vortex input beams as explained in the beginning. We observe the following features.

(1) Circular symmetry of the vortex input always breaks down to a discrete group.

(a) In a dissipative left-handed medium, the discrete symmetry group for a vortex of charge Q is C_{3Q} , before breaking down to simple C_2 axial symmetry at longer distances [Fig. 2(a)].

(b) In a dissipative right-handed medium, the discrete symmetry group for a vortex of charge Q is C_{2Q} , before breaking to C_Q and then to C_2 axial symmetry at longer distances [Fig. 2(b)].

(c) In a lossless left-handed medium, the discrete symmetry group for a vortex of charge Q is C_{3Q} for very short distances, before quickly breaking down to C_Q and finally C_2 [Fig. 2(c)].

(d) In a lossless right-handed medium, the discrete symmetry group for a vortex of charge Q is C_{2Q} , before breaking to simple C_2 axial symmetry at longer distances [Fig. 2(d)].

(2) Vortices decay approximately exponentially as they propagate along the longitudinal axis. Figure 4 shows the intensity of the beam across the z axis, for various regimes. At early z values, total intensity may behave nonmonotonically and nonuniversally but on longer scales it decays exponentially. For different charges, the intensity curves collapse to a unique exponential function at large z . As could be expected, lossless and dissipative cases differ somewhat and collapse to different curves.

The bottom line is that there is a vocabulary of patterns with C_Q , C_{2Q} , and C_{3Q} symmetries. One of them dominates in each case (left and right handed, dissipative and lossless) but at longer propagation distances the symmetry can change, before the intensity drops to near zero from dissipation. The final stadium of C_2 symmetry is only seen at very low intensities, so it might be practically unobservable in experiment; that is why we say the vocabulary only has three possible patterns, excluding C_2 .

The findings above are further corroborated by Fig. 3, which shows the vortices with different charges $Q = 1, 2$, and 3 in the same regime [dissipative left handed (a) and lossless left handed (b)]. As claimed above, the symmetry is C_{3Q} in panel (a), and (except at small z values) C_Q in panel (b). Finally, it is obvious that there is some mixing of patterns: the polygons are never exactly regular, so the groups C_n are certainly not exact symmetries; we use the C_n nomenclature merely for convenience.

One interesting phenomenon in Fig. 2(c) is that the pattern rotates along the z axis. This can be understood as excitation of multiple angular modes (of the form e^{ilz} with various l numbers) as the beam travels along the sample. This is a well-known consequence of nonlinear terms [5,29] and typically depends on the relative strength of nonlinear mode interactions compared to energy density $|E|^2 + |H|^2$ and dissipation γ . We will not explore it in quantitative detail in this paper as it is only tangential to our main topic of radial symmetry breaking; as one can see, the structure remains the same; just the orientation changes.

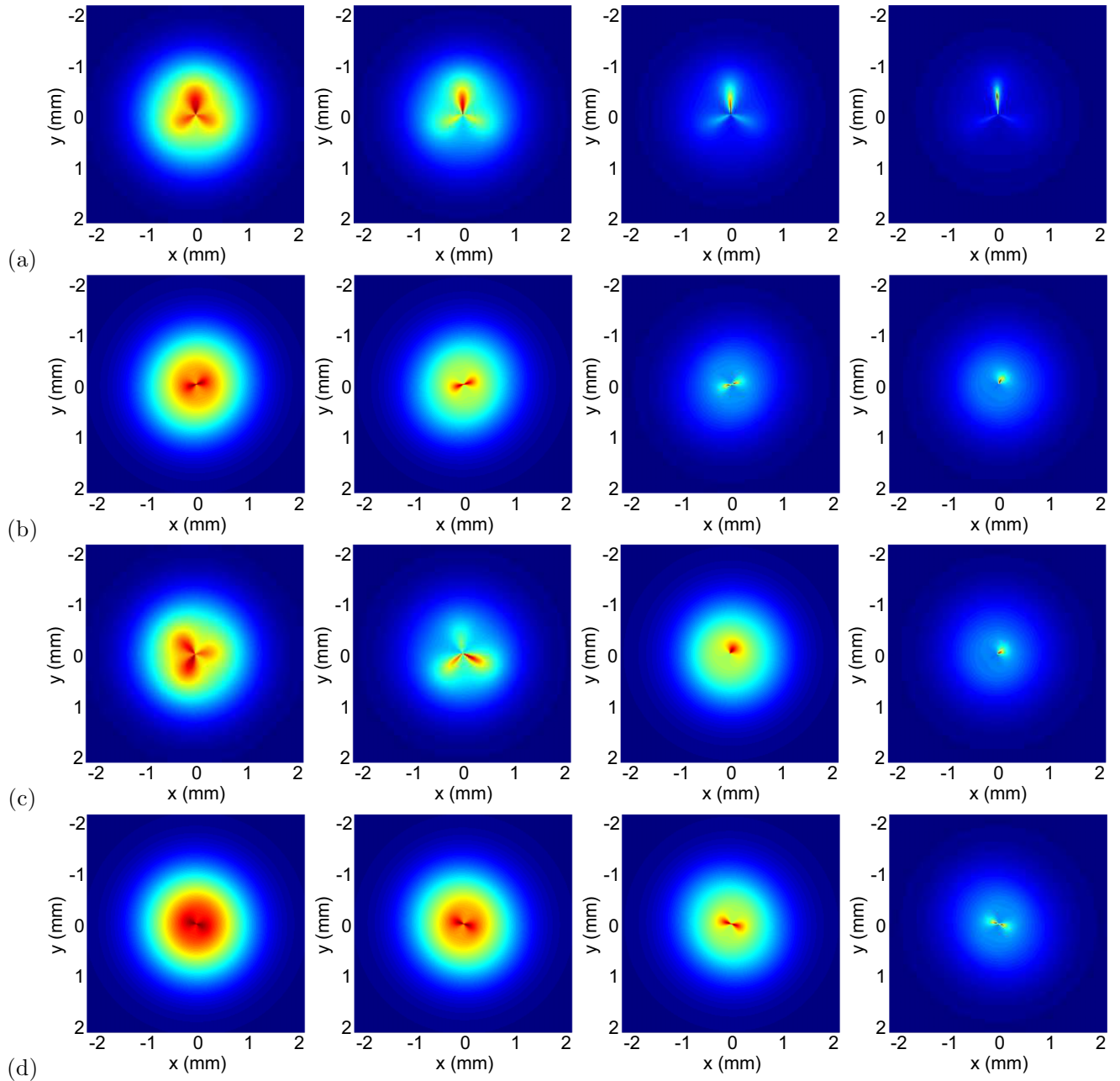


FIG. 2. The patterns for a $Q = 1$ vortex, in a left- and right-handed dissipative medium [(a) and (b), respectively], and in a left- and right-handed lossless medium [(c) and (d), respectively], at longitudinal slices $z = 2b, 4b, 6b,$ and $8b$, showing the $C_{3Q}, C_{2Q}, C_{3Q}/C_Q,$ and C_{2Q} regimes. The remaining parameters are defined in the text at the beginning of this section.

One might rightly worry that the initial conditions which contain a vortex in both electric and magnetic field are not very realistic, as in most materials the electric field dominates the optical response. Therefore, in experimental practice, one typically prepares a vortex in the electric field making use of phase masks or some other method, and the initial magnetic-field distribution is completely analytic. In Appendix B we repeat the calculations from Fig. 3 and show that the outcome is the same, including the vocabulary of patterns and their C_n shapes. Therefore, the E - H symmetric ansatz is merely a matter of convenience, and the realistic regime where $|H| \ll |E|$ is in fact covered by our paper.

Figure 4 shows that at long times the decay of intensity is universal for given dielectric dampening coefficient γ , which suggests the main mechanism of dissipation is in fact the radiative loss. This is because we deliberately chose ϵ and μ with small imaginary parts (for ϵ it can also be zero), so the losses in the medium are not so important when it comes to total energy (they are still important for being nonlinear and influencing the patterns). One important difference between the lossless medium (black and blue symbols in Fig. 4) and the dissipative medium (red, magenta) is that the former has a short interval of growing intensity, before reaching the universal regime of radiative decay. The physical reason is

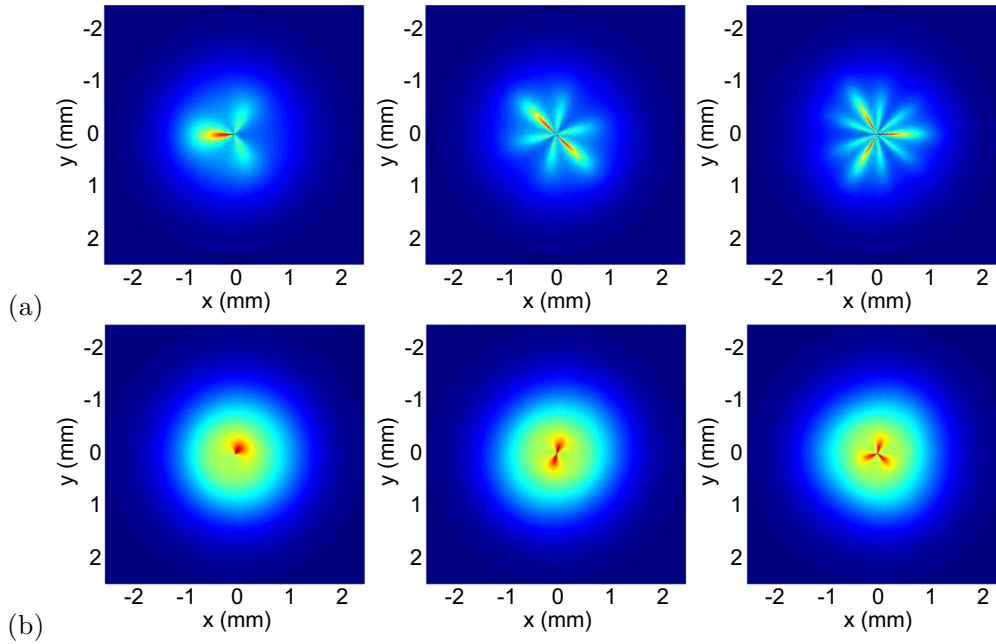


FIG. 3. The patterns for $Q = 1, 2,$ and 3 vortices (left to right), in a dissipative (a) and lossless (b) left-handed metamaterial. The behavior for three different charges confirms the previous conclusions for the type of symmetry encountered. All parameters except for the vortex charge are the same as for Fig. 2. The propagation distance is $z = 5b$ in (a) and $z = 8b$ in (b).

that the polarization, i.e., the rearrangement of charges in the self-defocusing metamaterial, reduces the overall electrostatic potential energy of the medium, and this energy becomes available to the beam, increasing its intensity. Clearly, once the radiative losses overcome the total potential energy available, the intensity decays. The growth is clearly a transient effect which cannot persist for long z intervals. A formal way to understand this is that the nonconservation of energy is encoded by the last term in (4), which can have a positive or

negative imaginary part depending on the sign of $\partial_z \mu / \mu$. At large values of z , we expect to enter a universal regime where this sign is constant, because the radiation loss dominates over nonlinearities and the exchange of energy between the beam and the medium; this is the universal decay regime in the figure.

IV. THE THEORY OF VORTEX EVOLUTION

The phenomenology described in the previous section can be understood on several levels. At the crudest level, we can introduce a variable-separation ansatz in the equations of motion and then linearize them in the amplitude (but not in the phase). This picture explains the C_{2Q} patterns, but not the C_{3Q} and C_Q regimes. It also does not explain the instabilities, that is, the changes and disappearance of patterns during the z propagation. For the full picture it is necessary to take into account the nonlinear effects through the loop corrections, i.e., to move perturbatively beyond the amplitude-linearized solution. A qualitative insight of the symmetry breaking can, however, be obtained also in a simpler and more elegant way, directly from the symmetry analysis of the model Lagrangian (6). Therefore, after finishing the amplitude-linearized analysis and the loop corrections from nonlinearity, we will obtain the same results from a unified mean-field treatment of the (nonlinear) model Lagrangian.

A note on terminology is in order. The solutions we find are not the textbook type of vortex with phase dependence solely of the type $e^{iQ\phi}$; rather, the dependence on the phase is more complicated, i.e., the phase is doing more than just the winding, but it is still true that the circulation of the phase around some point (the location of the vortex core) is an integer—the topological charge of the vortex. Such solutions are sometimes called spirals [22] whereas the term

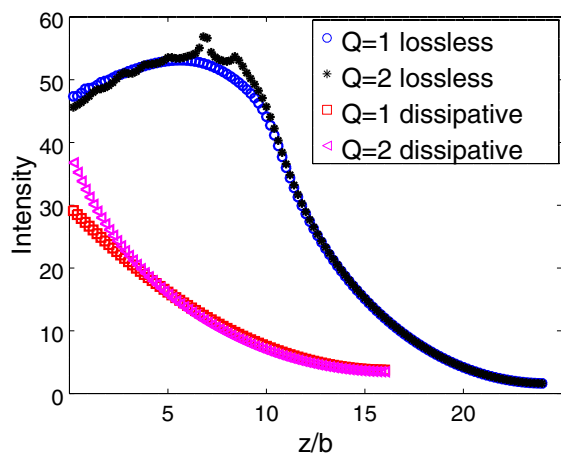


FIG. 4. Decay of the total intensity $I = \int dx \int dy (E^2 + H^2)$ in computational units for $Q = 1$ and 2 , for a lossless (blue circles, black stars) and dissipative (red squares, magenta triangles) left-handed material. At early times the behavior is complicated and nonuniversal but at late times it collapses to an exponential curve. This is expected when loss through radiation dominates. The oscillatory features of the $Q = 2$ lossless case (black) are likely due to finite numerical resolution.

“vortex” is reserved for the simple winding-phase solutions. We nevertheless stick to the widespread term “vortex” for any topologically charged solution under the fundamental group of the $U(1)$ phase symmetry.

A. Amplitude-linearized solution

We will separate variables in the equations of motion (4) and (5) [or the Lagrangian equations (7) and (8), which do not differ from the original equations at the amplitude-linearized level] and then plug in the vortex ansatz. The vortex ansatz is a solution which has a winding phase Φ with some winding number Q , for a constant (averaged) value of the permittivity $\mu_c = \text{const}$, because we ignore the nonlinear dependence of μ on $|H|$. The vortex solution of winding number (topological charge) Q in cylindrical coordinates (r, ϕ, z) can be separated into regular and vortex parts:

$$E = E_{\text{reg}} + E_{\text{vort}}. \quad (12)$$

We represent the vortex part as

$$E_{\text{vort}}(r, \phi, z) = Z_E(z)R_E(r)e^{iQ\phi - i\Phi(\phi)}, \quad (13)$$

and analogously for the magnetic field. Along the z axis we get $Z_E(z) = e^{i\lambda z}$ as expected, and the eigenvalue λ is arbitrary for now, i.e., it is determined by the boundary conditions along the z axis. Upon inserting (13) into (4), the equation separates into the angular part and the radial part. The former reads

$$\Phi'' - i(\Phi')^2 + 2iQ\Phi' + i\lambda^2 = 0, \quad (14)$$

where l is the eigenvalue of the angular part. This is the crucial equation—the phase dynamics is nonlinear because μ is in general complex and the terms with $\nabla_{\perp}\mu$ contain nonlinear dependence on the phase. The equation is easily solved by first introducing $w \equiv \Phi'$ and then reducing it to quadratures. The outcome is

$$\Phi(\phi) = \cos(\sqrt{Q^2 + l^2}\phi + C_l). \quad (15)$$

In other words, we still stay with a winding solution but various winding numbers (equal to $\sqrt{Q^2 + l^2}$) are possible when multiple modes are excited. Clearly, only the solutions with integer windings are physical, otherwise they would not be single valued. The most general solution is thus a superposition of solutions $Z_E(\lambda, l; z)\Phi_E(\lambda, l; \phi)R_E(\lambda, l; r)$ with different l modes so as to result in a single-valued function. Now the radial part acquires the form

$$R_E'' + \frac{1}{r}R_E' + \left(\frac{\lambda}{r^2} - k^2 + \epsilon_{D0}\tilde{\omega}^2\right)R_E + \frac{\alpha\mu_c\tilde{\omega}^2}{E_c^2}R_E^3 = 0, \quad (16)$$

with $\tilde{\omega} \equiv \omega[1 - \omega_0^2/(\omega^2 + i\omega\gamma)]$. If we disregard the cubic term (amplitude-linearized approximation),² the well-known solution in terms of Bessel functions is obtained:

$$R_E(r) \approx c_E^{(1)}(\lambda, l)J_{Q_l}(ar) + c_E^{(2)}(\lambda, l)Y_{Q_l}(ar), \\ Q_l \equiv \sqrt{Q^2 + l^2}, \quad a \equiv \sqrt{\lambda - \epsilon_{D0}\mu_c\tilde{\omega}/\omega E_0^2 - k^2}. \quad (17)$$

²This is justified at least in some interval of z values, as the system is dissipative and loses power $\int(E^2 + H^2)$, so the amplitude progressively decreases along z .

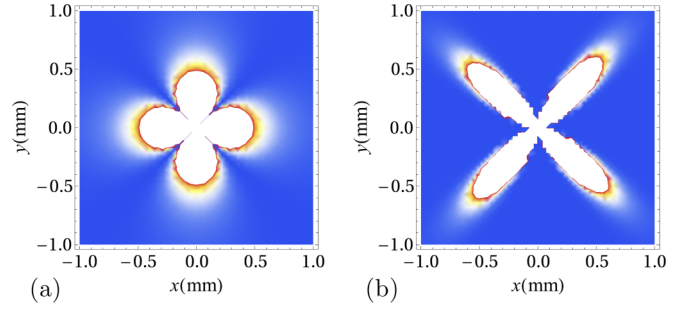


FIG. 5. Polygonal pattern $|E|^2$ for a vortex of charge $Q = 2$, for $k = 2$, $\epsilon_{D0} = 12.8$, and $\mu_c = 1.004$ (values of all parameters and constants in the main text), at radial slice $z = 1$, for a single vortex mode $l = 0$ (a), and for a linear combination of modes with $l = 0, 1$, and 2 decaying at infinity (b). The symmetry is $C_{2Q} = C_4$, which does not explain the C_Q and C_{3Q} regimes. Obviously, the crude picture of breaking the radial symmetry works but full explanation is lacking. It will come from the loop corrections.

Here, J and Y are the Bessel functions of first and second kind, respectively. Similar solutions $Z_H(z)$, $\Phi(\phi)$, and $R_H(r)$ are obtained for the magnetic field. The angular equation is identical for both fields: for this reason we have one solution Φ for both E and H . The eigenvalues λ and l and the values of the constants $c_{E,H}^{(1,2)}$ are determined by the boundary conditions. Obviously, (15) imposes the C_{2Q} symmetry, if $l = 0$. This simplest case is not necessarily the stable solution. We might have a sum over many l values. In principle, such sums may yield more complicated patterns, however we will see that when the physically reasonable boundary conditions are implemented (decay at infinity, single valuedness everywhere) one typically always has the robust C_{2Q} pattern. One important consequence of the fact that multiple l modes are possible is that due to nonlinear effects a new l mode can be created during the propagation along the z axis. We have already seen an example in Fig. 2(c). A quantitative analysis of this phenomenon requires a full nonlinear model and so can only be studied within the formalism of the next section.

This solution is not very satisfying but reproduces some of the features from the numerics, summarized at the start of the previous section: (1) the reduction of the full $O(2)$ symmetry down to a discrete symmetry C_n for some $n \in \mathbb{N}$, i.e., the polygonal form of the vortex, and (2) the value $n = 2Q$ is true in some but not in all situations. We show the solutions for a single angular mode from (15) and (17) in Fig. 5(a). In Fig. 5(b), we show a linear combination of angular modes with $l = 0, 1$, and 2 , with the coefficients $c_{E,H}^{(1,2)}$ in (17) chosen so that the total intensity still decays sufficiently fast at infinity. The symmetry is still C_{2Q} . Apparently, the regimes with the C_Q and C_{3Q} symmetries require loop corrections from nonlinear μ to be taken into account.

B. Loop corrections

The origin of the breaking of radial symmetry is the fact that a discrete set of modes in Fourier space is selected. This is best seen from the Fourier transform of the solutions (15) and (17). We will calculate the propagator $G(\mathbf{u})$ at

constant z , i.e., the Fourier transform $\mathbf{r} \mapsto \mathbf{u}$ of the solution with a Dirac delta source. This source imposes the boundary condition $R_E(0) \rightarrow \infty$, $\int dr r \cos \phi R_E(r) = 1$, giving $c_E^{(1)} =$

0 , $c_E^{(2)} = 2\pi/\Gamma(Q/2)$ in (17). Fourier transforming $(x, y) \mapsto (u_x, u_y)$ we get for a single mode (17), making use of the Bessel and Lommel integrals:

$$G_{E,H}(\mathbf{u}) = \frac{2\pi}{\Gamma(Q/2)} \frac{e^{iQ(\pi/2+\phi)}}{au} \left(\frac{\sin[(u-a)\Lambda]}{u-a} - \frac{\cos[(u+a)\Lambda - \pi Q]}{u+a} \right) + \frac{2\pi}{\Gamma(Q/2)} \frac{e^{-iQ(\pi/2+\phi)}}{au} \left(\frac{\cos[(u-a)\Lambda + \pi Q]}{u-a} - \frac{\cos[(u+a)\Lambda]}{u+a} \right). \quad (18)$$

Here, Λ is the ultraviolet (UV) small-length and high-momentum cutoff, i.e., the Fourier transform is performed by integrating $\int_{1/\Lambda}^{\infty} dr \int_0^{2\pi} d\phi$. The cutoff has a clear physical meaning: $1/\Lambda$ is the size of the vortex core (where the vortex ansatz stops working because the gradient of the field becomes too high). We clearly do not get anything new by just Fourier transforming. The goal is to move beyond the amplitude-linearized approximation of the previous section by considering the effects of nonconstant permittivity μ instead of constant (averaged) μ_c . This calculation is essentially elementary but might be tedious and boring for readers who are not fond of perturbative field theory. Most of the integrations are in Appendix C. Even the rest of this subsection can be skipped until the the equation (22), where we discuss the final result.

Putting μ from (3) in place of μ_c requires the solutions for ω_{NL} in terms of the magnetic field. The solutions are readily found from the Cardan formulas (we do not give them explicitly as they are cumbersome and not very illustrative). But the form of the H dependence of ω_{NL} is seen already from the Viète formula:

$$(\omega_{\text{NL}}^{(1)})^2 + (\omega_{\text{NL}}^{(2)})^2 + (\omega_{\text{NL}}^{(3)})^2 = \frac{1 + 2\Omega^2}{1 + |H|^2/\alpha E_c^2}, \quad (19)$$

so the solutions depend on $|H|^2$ only, with no higher powers of the magnetic field. Inserting this into \mathcal{L} , we get the nonlinear correction of the form

$$\delta\mathcal{L} = g_{2,0,0}|\nabla_{\perp}E|^2 + g_{0,2,0}|E|^2 + g_{0,2,2}|E|^2|H|^2 + g_{2,0,2}|\nabla_{\perp}E|^2|H|^2. \quad (20)$$

We thus have two quartic interaction terms and two quadratic terms. We do not intend to calculate the loop corrections in full detail; it is not worth the effort as we only want to capture the symmetry, i.e., the form of the angular dependence. First of all, the quadratic corrections $g_{2,0,0}$ and $g_{0,2,0}$ trivially renormalize the parameters in the bare propagator and do not change its functional form. Lowest-order non-trivial loop corrections to the self-energy come from $g_{0,2,2}$ and $g_{2,0,2}$. The electric field receives the correction $G_E^{-1} \mapsto (G_E + \Sigma_E^{(1)} + \Sigma_E^{(2)})^{-1}$ with

$$\Sigma_E^{(1)} = g_{0,2,2} \int d\mathbf{u}' G_H(\mathbf{u}') \approx g_{0,2,2} e^{3iQ/2} \sin(\pi Q) \ln \Lambda, \\ \Sigma_E^{(2)} = \frac{3}{2} g_{0,2,2} \int d\mathbf{u}' \int d\mathbf{u}'' G_H(\mathbf{u}') G_H(\mathbf{u}'') G_E(\mathbf{u} - \mathbf{u}' - \mathbf{u}'') \\ \approx \text{const} \times [a^{3/2} \cos(3Q\phi/2) - 2iQ^2 \ln a]. \quad (21)$$

We will write all equations for E , because this field receives interesting corrections from the gradient of μ [Eqs. (4) and (7)]. The magnetic field does not couple to the permeability ϵ in the same way in the original equation (5), and in the Lagrangian form (8) it does but ϵ does not contain such strong (nonpolynomial) nonlinearities as μ . One- and two-loop corrections appear not only in the self-energy but also in the vertex operators. However, the vertex corrections only have a weak momentum dependence and consequently the coordinate dependence (geometric patterns) of the solution is not significantly affected by them. For that reason we will not discuss them in detail.

The correction $\Sigma_E^{(1)}$ is the Hartree correction with a single vacuum bubble which is not very interesting: it merely introduces an additional mass term and does not influence the momentum dependence and thus the geometry of the patterns. As could be expected from power counting, it is logarithmically divergent in the UV cutoff Λ . Of course, this is not a problem in an effective theory; we have already explained the physical meaning of Λ . The watermelon diagram $\Sigma_{E,H}^{(2)}$ is crucial: it is momentum dependent. Its calculation is found in Appendix C. An informal way to estimate its effect is the following: the leading contribution comes from the region where $\mathbf{u} \approx \mathbf{u}' - \mathbf{u}''$ because this is a pole of the self-energy correction. Then we are left with angular integrals only, and they reduce to integrals of products of three rational functions [for the three propagators in (21)] of the half angle—this gives rise to $3\phi/2$ in the argument of the cosine. Now the dressed propagator $(G_{E,H}^{-1} + \Sigma)^{-1}$ needs to be Fourier transformed back to real space. We will only do this approximately (it is likely impossible to do exactly in closed form). The outcome is

$$E_{\text{vort}}(r, \phi, z) = \frac{e^{(i\lambda - 2Q^2 \ln a)z} \cos(Q\phi)}{\sqrt{kr}} \\ \times \left[c_E^{(1)}(\lambda, l) \left(1 + \frac{(2\pi)^{3/2} g_{0,2,2}}{\Gamma(Q/2)^3} \cos(3Q\phi/2) \right) + c_E^{(2)}(\lambda, l) \left(1 + \frac{(2\pi)^{3/2} g_{0,2,2}}{\Gamma(Q/2)^3} \sin(3Q\phi/2) \right) \right]. \quad (22)$$

No doubt the reader sees that the terms $\cos(3Q\phi/2)$ and $\sin(3Q\phi/2)$ give a pattern $|E_{\text{vort}}|^2$ with $3Q$ branches, in addition to the $2Q$ polygons obtained from the term $\cos(Q\phi)$. The interference between the two patterns might (1) break the symmetry completely and (2) lead to C_Q

symmetry if the relative phase between the leading term and the corrections is approximately $2\pi/Q$. Both cases are seen in numerical work: C_{3Q} appears in all left-handed materials [Figs. 2(a) and 2(c)], and elements of C_Q symmetry are present in almost all cases at long propagation distances z [Figs. 2(a)–2(c) and 3].

The self-energy has an imaginary part [equivalently, the solution (22) exhibits exponential decay in z], meaning that these configurations are not stable—they are only seen up to some propagation distance z . The exact order (along z) and stability of each of the patterns depend on the details of the permeability ϵ . One important and universal lesson is, however, that the decay rate [the real part of the exponent in (22)] is proportional to Q^2 , therefore the higher the value of $|Q|$ the faster it decays. This supports the general intuition that vortices with high winding numbers are not stable. But unlike the simplest case of the XY model or a superfluid where the stability only allows $Q = \pm 1$ we can in principle have arbitrarily high Q as we have seen also in the numerics; their lifetimes are smaller and smaller as Q grows, but still finite. The exponential decay itself is also confirmed by the numerics, as seen from Fig. 4.

C. Isotropy breaking: The look from the action

The basic mechanism leading to the symmetry breaking $O(2) \mapsto C_{3Q} \mapsto C_{2Q} \mapsto C_Q$ is seen already from the model Lagrangian (6). The symmetry breaking is essentially the consequence of the interplay of the nonlinear-sigma-model form of the kinetic term and the complex nonlinearity of the magnetic permeability μ . Therefore, we can take a static approximation of the z dynamics, ignoring the z dependence; clearly, in that framework we can only obtain the vocabulary of patterns, not the relative stability of C_Q , C_{2Q} , and C_{3Q} .³ The separation of variables remains a natural ansatz, and the vortex nature of the solution implies $E_{\text{vort}} = E_0(r)e^{i\Theta(\phi)}$ with $\oint d\phi \Theta(\phi) = 2\pi Q$ and analogously for the magnetic field. The Lagrangian (6) then becomes

$$\mathcal{L} = \frac{(E'_0)^2 + \frac{(\Theta')^2}{r^2} + k^2 E_0^2}{\mu} + \frac{(H'_0)^2 + \frac{(\Theta')^2}{r^2} + k^2 H_0^2}{\epsilon}. \quad (23)$$

The fact that μ contains $\omega_{\text{NL}}^2(|H|^2)$, which is in turn the solution of the cubic equation, introduces a branch cut in H because of the cubic roots. This is the simplest explanation of the origin of the C_{3Q} symmetry. More quantitatively, the story follows exactly the Landau-Ginzburg paradigm: while the initial Lagrangian only depends on $|E|^2$ and $|H|^2$ and thus preserves isotropy, the saddle-point solution is given by the equation

$$\frac{\epsilon(\nabla_{\perp}^2 - H)E - \epsilon' \nabla_{\perp} E \cdot \nabla_{\perp} H}{\epsilon^2} + \frac{\mu'}{\mu^2} |H|^{-1/3} = 0, \quad (24)$$

where we have used that $\mu = \mu(\omega_{\text{NL}}^2)$ and $\omega_{\text{NL}}^2 = \omega_{\text{NL}}^2(|H|^{2/3}, |H|^{4/3})$ (from the Cardan formulas). With the ansatz adopted above, the amplitude equation for $E_0(r)$ is the

nonlinear amplitude equation (16). The equation for the phase part Θ is more interesting. It reads

$$\frac{(\Theta')^2 \left(1 - \frac{\epsilon' E_0}{\epsilon H_0}\right) - k}{\epsilon} + \frac{2\mu'}{3\mu^2} |H|^{-1/3} = 0. \quad (25)$$

The cubic root carries a branch cut, and the last term really evaluates to $2\mu'/3\mu^2 \times H_0^{-1/3} e^{-i\Theta/3 + 2n\pi i/3}$ with $n = -1, 0$, and 1 . The solution Θ_0 which satisfies the phase winding condition is obtained in implicit form as

$$i(\Theta_0 + 2\pi n/3) = K_n \ln \left[\frac{k \left(1 - \frac{\epsilon' E_0}{\epsilon H_0}\right)}{E_0^2 + H_0^2} \sec^2 \left(\frac{Q}{2} \phi \right) \right], \quad (26)$$

where K_n is a constant determined by the amplitude solution and depending also on $n = -1, 0$, and 1 ; its exact value is hard to find analytically as we do not know the solution to the amplitude equation in the nonlinear regime. But that is not crucial for our general argument. The point is that the system can choose a solution with any of the values $n = -1, 0$, and 1 ; i.e., even though the equations of motion (and the Lagrangian) are isotropic, the solution is not. Each n branch behaves as $\approx 1/\cos^2(Q\phi/2)$, only they are rotated by $\pm 2\pi/3$ with respect to each other, and each of them has a C_Q symmetry. Put together, the three branches give C_{3Q} patterns. But all that holds if two of the cubic roots are complex. If all cubic roots are real, the phase remains single valued, and we only have C_Q symmetry, coming directly from (26) if we fix $n = 0$, i.e., if we only keep a single branch.⁴

What is the regime in which cubic roots are real and the symmetry is C_Q , as opposed to the complex roots and C_{3Q} patterns? The easiest way is to look at the cubic equation (3) for the magnetic permeability (and the nonlinear frequency ω_{NL}). For $\mu > 0$ (right-handed regime), the roots are all real and C_{3Q} patterns cannot occur. Indeed, the C_{3Q} phase is only present in Figs. 2(a) and 2(c), in left-handed media.

This approach is much more physical and elegant than the tour-de-force calculations of the previous two sections but it does not give explicit solutions for E and H ; it only classifies the symmetries of the solution. This is why we still needed the perturbative linear and two-loop analysis, to arrive at more quantitative results.

The saddle-point solution (26) is nonlinear, unlike the linearized solution found in the first subsection (15). It is not a vacuum in the usual field-theory sense, however, as it is not constant. We are dealing with dynamical criticality of the kind discussed in [21]. In the vicinity of this solution, the Lagrangian describes the fluctuations of amplitude δE and δH , and the fluctuations of phase $\delta \Phi$. Similar to the $O(3)$ -type spin models [23] and multibeam optical systems [32], and unlike simple XY-type models, the phase and amplitude fluctuations mix. By analyzing the fluctuation equations, it should be possible to understand analytically also the transition from the left-handed to the right-handed regime as the parameters are varied, i.e., what are the instabilities that drive it. We will not

³We could take the ansatz $e^{i\lambda z}$ instead; it would merely modify $k^2 \mapsto k^2 - \lambda$.

⁴We use the fact that a cubic equation has either one or all three solutions real.

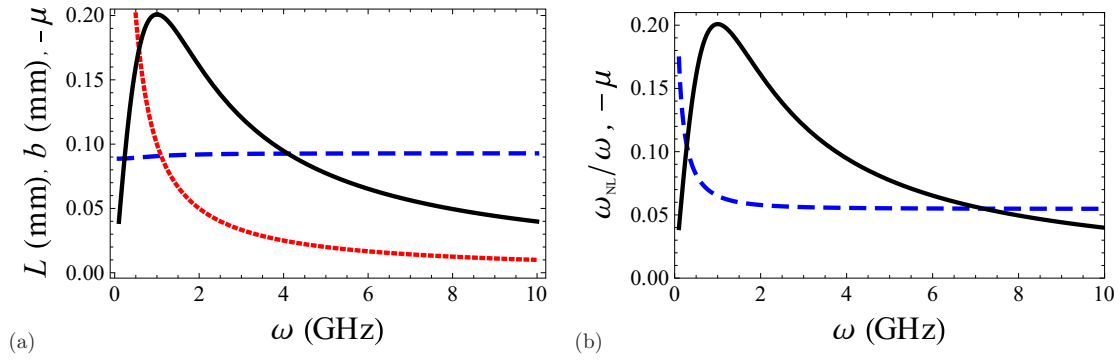


FIG. 6. (a) Frequency dependence of the typical propagation length scale for the dissipation of the vortex a^{2Q^2} (blue dashed line) and for the evolution of the symmetry-breaking C_n patterns (red dotted line). The symmetry breaking is detectable as long as the pattern evolution is faster than the dissipation, i.e., as long as the red curve is below the blue one. This is obviously the case for most of the frequency range. We also plot the frequency dependence of the negative permittivity $-\mu$ (black full line; because of the minus sign large positive values in the plot are really large negative values of μ). The left-handed regime is most prominent at intermediate frequencies, which are also inside the regime of the symmetry breaking. (b) Frequency dependence of the relative strength of nonlinear interactions $\omega_{\text{NL}}/\omega$ (blue dashed line) together with negative permittivity $-\mu$ as in (a) (black full line). Our calculations, based on a pair of nonlinear Schrödinger-like equations, are reliable as long as the nonlinearity is not too strong. This is again the case for all but very small frequencies, and again includes the left-handed regime.

attempt that here; it is a long subject that deserves separate work.

V. TOWARD EXPERIMENTAL VERIFICATION AND APPLICATIONS

We will now briefly discuss what an experimentalist can learn from our results and what to look for in practical work. Wave propagation through the metamaterial can be observed by measuring the transmission coefficients S_{ij} . From these coefficients, one can also reconstruct the electric-field intensity $|E|^2$, which can be directly compared to our intensity maps like Figs. 2 and 3 [33]. Another quantity which can be measured is the voltage waveform, which can be used to construct amplitude envelopes [34].

Therefore, the predicted symmetry breaking is in principle directly observable. But the question remains how widespread it will be for realistic values of the parameters. From a more applied viewpoint, this question is reversed: how to make a vortex transmission through a left-handed waveguide stable. In other words, how *not* to observe the symmetry breaking. It is true that the phenomenon disappears as soon as the vortex charge is zero, i.e., when the beam is not a vortex. However, the vortex patterns are likely important in applications. First, as a topologically protected object with conserved charge, a vortex is among the natural candidates for computational devices and information transmission (for the same reasons that solitons are also interesting in that regard: they are robust to noise, carry a discrete “quantum” number, i.e., charge, and are stable to small local perturbations). Second, in the presence of impurities in the sample, vortices can form in a nonlinear metamaterial from the initially nonvortexing beam [23].

Let us focus on the left-handed regime, which is the most interesting and the most relevant for applications. The first condition is therefore to be in the frequency regime with $\mu(\omega) < 0$. This can be checked directly from Eq. (2) as we did in Fig. 1(c). The second issue is that the symmetry breaking takes some finite time, i.e., some finite propagation length, which is of order b ; as can be seen from Fig. 3 and

directly from Eqs. (4) and (5), this is the length scale over which the patterns change. On the other hand, the one-loop calculation (22) shows that the intensity decays with the rate $\sim a^{-2Q^2}$. As long as this is less than the characteristic length b , one will likely not see the symmetry breaking but just eventual dissipation of the beam. Therefore, these two scales should be compared for some reasonable parameter values. We show this in Fig. 6(a) for $F = 0.4$, $\epsilon_{D0} = 12.8$, $\gamma = 1$ GHz, and $\omega_0 = 10$ GHz. Apparently, the length scale of the C_n pattern development (red dotted line) is nearly always shorter than the dissipation scale (blue dashed line), so we expect that the effect predicted in the paper is readily seen in experiment, at least for $Q = \pm 1$. For larger vortex charges, the dissipation grows quickly and high Q values are probably not easily observed. Conversely, if the goal is to keep a stable radially symmetric vortex pattern, one should remain at small frequencies, although for $\omega \ll \omega_0$ the material is not strongly left handed, as can be seen from the $-\mu(\omega)$ dependence, also given in the figure.

There is still one remaining issue. Our theoretical approach, based on a pair of nonlinear Schrödinger-like equations, inherently disregards some effects. It describes a quasi-monochromatic wave without wave mixing or dissipation due to higher harmonic generation [5]. Such phenomena become significant for strong nonlinearities, so we should compare the nonlinearities in ϵ and μ to the typical energy (frequency) scale of the vortex. In Eqs. (1) and (2) the approximate ratios of the nonlinear to linear terms are given by $|E|^2/\epsilon_{D0}$ and $\omega_{\text{NL}}/\omega_0 \sim (A/H)^{1/3}$. The first scale is frequency independent and solely depends on the beam intensity. The second scale depends on frequency and needs to be inspected more closely. In Fig. 6(b) we plot the nonlinearity ratio for the magnetic field for a range of frequencies ω , again together with the permittivity to make sure we are at the same time in the left-handed regime. The relative nonlinearity strength quickly saturates around a value $0.06 \ll 1$, so we are rather confident that our equations of motion still make sense.

Altogether, the conclusion is that the breaking of radial symmetry is observable by standard means (measuring the

transport coefficients and reconstructing the intensity map at the exit face of the metamaterial), as long as the frequency of the wave is not too low. This kind of instability kicks in at shorter propagation lengths [of order 0.1 mm in Fig. 6(a)] than the nonlinear diffraction effects studied for breathers in [35], suggesting that vortex signals are more fragile and less convenient for information transmission.

VI. DISCUSSION AND CONCLUSIONS

Our main result is contained already in the title—left-handedness and nonlinearity together create the breaking of the $O(2)$ symmetry down to a discrete group, with the pattern vocabulary consisting of the C_{3Q} , C_{2Q} , and C_Q patterns. In the right-handed system with the same nonlinearity the isotropy is broken again, but the pattern vocabulary only has C_{2Q} and C_Q stages. How exactly the patterns evolve into each other and through which instabilities is not universal, and it depends on the exact form of ϵ and μ . In our model, the ϵ dependence is mainly encapsulated in the dissipation γ : the left-handed nondissipative case is usually dominated by C_Q after a much shorter C_{3Q} phase, whereas the dissipative left-handed metamaterials most prominently show C_{3Q} patterns. For the right-handed materials, nondissipative and dissipative dynamics show mainly C_{2Q} and C_Q patterns, respectively.

A detailed account of the pattern dynamics was only possible through numerical work. But the vocabulary itself—the existence of symmetries C_{3Q} , C_{2Q} , and C_Q —we were able to understand analytically. The dynamic Landau-Ginzburg picture reveals this as a consequence of the cubic root nonlinearity in the magnetic permittivity, and the fact that the cubic equation has either two complex roots in the left-handed regime or all three real roots in the right-handed regime, and the presence or absence of dissipation in the electric permeability. In the framework of our field theory model, the second derivative of the free energy (on-shell Lagrangian, Landau-Ginzburg functional) likely has a jump when the symmetry changes. This is a strong encouragement that the phenomena we observe here, and in general the walk through the pattern vocabulary, can be understood from the viewpoint of order and disorder transitions.

Similar phenomena were studied also in [15,18] and above all [10], where C_{3Q} necklaces were found, within a model of left-handed metamaterials given in [15] and similar to ours. Clearly, we have not exhausted this subject; more research is still needed to fully understand the transition between different symmetries and their instabilities. Vortices in metamaterials seem to be a promising arena, as in a metamaterial the nonlinearity and the frequency band where the material is left-handed can to some extent be tuned at will. Therefore, the phase diagram of collective vortex interactions can also be studied, and is an obvious topic for future work.

ACKNOWLEDGMENTS

This work has made use of the Sci-Hub service. We are grateful to Milan Petrović and Mariya Medvedyeva for helpful remarks. We also thank the referees for some important and stimulating questions. Work at the Institute of Physics is funded by Ministry of Education, Science, and Technological Development, under Grant No. OI171017.

APPENDIX A: DERIVATION OF THE EQUATIONS OF MOTION FROM THE MAXWELL EQUATIONS

Start from the definitions $\hat{D} = \epsilon \hat{E}$ and $\hat{B} = \mu \hat{H}$ and the Maxwell equations in the absence of external charges and currents ($\rho = \hat{j} = 0$):

$$\begin{aligned} \nabla \cdot \hat{D} = \rho = 0, \quad \nabla \cdot \hat{B} = 0, \quad \nabla \times \hat{E} = -\partial_t \hat{B}, \\ \nabla \times \hat{H} = 4\pi \hat{j} + \partial_t \hat{D} = \partial_t \hat{D}. \end{aligned} \quad (\text{A1})$$

We make the following assumptions.

(1) We assume small gradients of the permittivities ϵ and μ , so their second and higher derivatives are disregarded. Since $\omega \propto k$, it means that mixed derivatives of the form $\partial_t \nabla \epsilon$ are also disregarded. In other words, the characteristic length scale l along the z axis on which ϵ and μ change is assumed to be large compared to the characteristic scale b of the changes in E and H .

(2) We assume that the time dependence is harmonic so $\partial_t = -i\omega$.

Acting on the last equation by $\nabla \times$ and making use of the identity $\nabla \times \nabla \times \hat{H} = -\nabla^2 \hat{H} + \nabla(\nabla \cdot \hat{H})$, one gets for the left-hand side

$$\begin{aligned} \nabla \times \nabla \times \hat{H} &= -\nabla^2 \hat{H} + \nabla \left(\nabla \cdot \frac{\hat{B}}{\mu} \right) \\ &= -\nabla^2 \hat{H} + \nabla \left(\frac{1}{\mu} \nabla \cdot \hat{B} \right) - \nabla \left(\frac{\nabla \mu}{\mu^2} \cdot \hat{B} \right) \\ &= -\nabla^2 \hat{H} + \nabla \left(\frac{1}{\mu} \nabla \cdot \hat{B} \right) - \nabla \cdot \left(\frac{\nabla \mu}{\mu^2} \right) \hat{B} \\ &\quad - \frac{\nabla \mu}{\mu^2} \nabla \cdot \hat{B} = -\nabla^2 \hat{H} + 0 + O(1/l^2) + 0 \\ &= -\nabla^2 \hat{H}, \end{aligned} \quad (\text{A2})$$

where we used $\nabla \cdot \hat{B} = 0$ and disregarded the second derivative of μ . The right-hand side yields

$$\begin{aligned} \nabla \times \nabla \times \hat{H} &= \nabla \times (\partial_t \hat{D}) = -i\omega \nabla \times \hat{D} = -i\omega \nabla \times (\epsilon \hat{E}) \\ &= -i\omega (\nabla \epsilon) \hat{E} - i\omega \epsilon \nabla \times \hat{E} \\ &= -i\omega (\nabla \epsilon) \hat{E} - \omega^2 \epsilon \mu \hat{H} = O(1/l^2) + \omega^2 \epsilon \mu \hat{H}, \end{aligned} \quad (\text{A3})$$

so we obtain

$$\nabla^2 \hat{H} + \omega^2 \epsilon \mu \hat{H} = 0. \quad (\text{A4})$$

For the \hat{E} field we start from the third Maxwell equation, act by $\nabla \times$, and find for the left-hand side

$$\begin{aligned} \nabla \times \nabla \times \hat{E} &= -\nabla^2 \hat{E} + \nabla(\nabla \cdot \hat{E}) = -\nabla^2 \hat{E} - \nabla \left(\nabla \cdot \frac{\hat{D}}{\epsilon} \right) \\ &= \nabla^2 \hat{E} - \nabla \left(\frac{1}{\epsilon} \nabla \cdot \hat{D} \right) + \nabla \left(\frac{\nabla \epsilon}{\epsilon^2} \right) \epsilon \hat{E} \\ &\quad + \frac{\nabla \epsilon}{\epsilon^2} \nabla \cdot \hat{D} = -\nabla^2 \hat{E} + 0 + O(1/l^2) + 0 \\ &= -\nabla^2 \hat{E}, \end{aligned} \quad (\text{A5})$$

and for the right-hand side we get

$$\begin{aligned}\nabla \times \nabla \times \hat{E} &= -\partial_t(\nabla \times \hat{B}) = -\partial_t(\nabla \times \hat{B}) \\ &= -\partial_t[\nabla \times (\mu \hat{H})] = -\partial_t[(\nabla \mu) \hat{H} + \mu \nabla \times \hat{H}] \\ &= -(\partial_t \nabla \mu) \hat{H} - \nabla \mu \cdot \partial_t \hat{H} - \partial_t(\mu \partial_t \hat{D}) \\ &= O(1/l^2) - \frac{\nabla \mu}{\mu} \nabla \hat{E} + \omega^2 \epsilon \mu \hat{E},\end{aligned}\quad (\text{A6})$$

so

$$\nabla^2 \hat{E} + \omega^2 \epsilon \mu \hat{E} - \frac{\nabla \mu}{\mu} \nabla \hat{E} = 0. \quad (\text{A7})$$

For our geometry we take the paraxial beam approximation, with the ansatz $\hat{E} = E(x, y)e^{i(kz - \omega t)}$, $\hat{H} = H(x, y)e^{i(kz - \omega t)}$, so the nabla acts as

$$\nabla \hat{E} = (\nabla_{\perp} E, \partial_z E + ikE)e^{i(kz - \omega t)}, \quad (\text{A8})$$

and the Laplacian operator acts as

$$\nabla^2 \hat{E} = (\nabla_{\perp}^2 E + 2ik\partial_z E - k^2 E)e^{i(kz - \omega t)}, \quad (\text{A9})$$

and analogously for the magnetic field. Now to write the equations motion in the final form we rescale $E \rightarrow E \times 2kb$, $H \rightarrow H \times 2kb$, and $z \mapsto z \times 2kb$, where b is some characteristic length scale along the z axis, and divide the equations by bk^2 to obtain the equations (4) and (5), reprinted here for convenience:

$$\begin{aligned}-\frac{i}{b}\partial_z E &= \nabla_{\perp}^2 E + [\omega^2 \epsilon (|E|^2) \mu (|H|^2) - k^2] E \\ &\quad - \frac{\nabla_{\perp} \mu (|H|^2)}{\mu (|H|^2)} \nabla_{\perp} E - i \frac{\partial_z \mu (|H|^2)}{2\mu (|H|^2)} E,\end{aligned}\quad (\text{A10})$$

$$-\frac{i}{b}\partial_z H = \nabla_{\perp}^2 H + [\omega^2 \epsilon (|E|^2) \mu (|H|^2) - k^2] H. \quad (\text{A11})$$

For comparison to the equations given in [4,7,12], one needs (1) to rescale $H \mapsto \omega^2/c^2 H$ to get the term $-\gamma^2 H = -k^2/\omega^2$ in (A11) and (2) to absorb the factor $-k^2$ in (A10) in the definition of ϵ_{D0} . This is possible as ϵ and μ have a constant term (equal ϵ_{D0} and 1, respectively) so the product $\epsilon\mu$ also has a constant term proportional to ϵ_{D0} , and the contribution $k^2 E$ can be absorbed as $\epsilon_{D0} \mapsto \epsilon_{D0} - k^2$. We thus arrive at a system identical to that from [4], except for the extra terms for the propagation along the z axis.

APPENDIX B: CONFIGURATIONS WITH NO VORTICITY IN THE MAGNETIC FIELD

Here we show that our results stay valid also when only the electric field has vortex patterns whereas the magnetic field starts analytic everywhere. As we discuss in the main text, this situation is experimentally more relevant than the one assumed in most calculations in the paper (that both the

electric and the magnetic field have a vortex as they enter the material). The electric field is typically a few orders of magnitude more intense than the magnetic field, as seen in [4]. Therefore, one typically controls the electric field directly, imposing a given boundary condition at the front end of the material. Despite this fact, the magnetic field remains very important: the coupled equations of motion (4) and (5) require both E and H to be nonzero. Indeed, as explained in [4], the left handedness comes as a consequence of the hysteresis-type dependence of the magnetic permittivity on H . So while it is crucial that E and H are both nonzero, it is also true that the results should remain valid for $|H| \ll |E|$, and for the boundary condition that only has a vortex in E at the front of the metamaterial, not for H . With such boundary conditions and the same parameter values as before, Fig. 7 repeats the calculations of Fig. 3. Obviously, the symmetries remain the same and the similarity of the results for the two cases is striking. Obviously, the $|E|^2$ map is insensitive to the details of the initial magnetic-field pattern, as one expects from experiments and common wisdom in nonlinear optics. We are thus content that the numerically simplifying assumption of identical $z = 0$ boundary conditions for E and H does not put into question the findings of our paper.

APPENDIX C: THE CALCULATION OF THE SELF-ENERGY DIAGRAMS

We discuss here in some more detail the equations (21) from the main text. First we give the expressions for the couplings $g_{2,0,0}$, $g_{0,2,0}$, $g_{2,0,2}$, and $g_{0,2,2}$, which come from the expansion over the magnetic field H of the nonlinear dependence $\mu(H)$ in (20):

$$g_{2,0,0} = \frac{\alpha E_c^4 \omega_0^2 - (\omega - i\Gamma)\omega\alpha E_c^8}{H_0 + \alpha E_c^4 [\omega_0^2 - (\omega - i\Gamma)\omega\alpha E_c^2]}, \quad (\text{C1})$$

$$g_{0,2,0} = (k^2 - \lambda^2)g_{2,0,0}, \quad (\text{C2})$$

$$g_{2,0,2} = 2\alpha E_c^2 H_0 \frac{\omega_0^2 - (\omega - i\Gamma)\omega\alpha E_c^4}{\{H_0 + \alpha E_c^4 [\omega_0^2 - (\omega - i\Gamma)\omega\alpha E_c^2]\}^2}, \quad (\text{C3})$$

$$g_{0,2,2} = (k^2 - \lambda^2)g_{2,0,2}. \quad (\text{C4})$$

For simplicity, we will treat the case when $\lambda = k$ and thus $g_{0,2,0} = g_{0,2,2} = 0$. This simplifies the calculations substantially while it does not change the symmetry of the solution. It is possible to evaluate the diagram $\Sigma^{(1)}$ exactly in terms of sine and cosine integrals Si and Ci. The angular integration is straightforward; the integration over u results in four combinations of the trigonometric integrals, for the four terms in (18). Three of the four integrals are finite and therefore they just shift the mass term. The third term of the propagator is logarithmically divergent:

$$\Sigma_3^{(1)} = \frac{4\pi \sin \pi Q}{Q\Gamma(Q/2)} e^{-3i\pi Q/2} \frac{1}{a^2} \{\gamma_E + \ln \Lambda + (-1)^Q a [\cos(a\Lambda) \text{Ci}(a\Lambda) + \sin(a\Lambda) \text{Si}(a\Lambda)]\}. \quad (\text{C5})$$

To judge the effect of this term, we should extract the mass squared r_m of the bare propagator, writing it out for small u :

$$G(\mathbf{u} \rightarrow 0) = \frac{2\pi}{\Gamma(Q/2)} \frac{1}{u(u^2 - a^2)} \{e^{iQ(\pi/2+\phi)} [\cos(a\Lambda - \pi Q) - \sin(a\Lambda)] + e^{-iQ(\pi/2+\phi)} [\cos(a\Lambda + \pi Q) - \cos(a\Lambda)]\}. \quad (\text{C6})$$

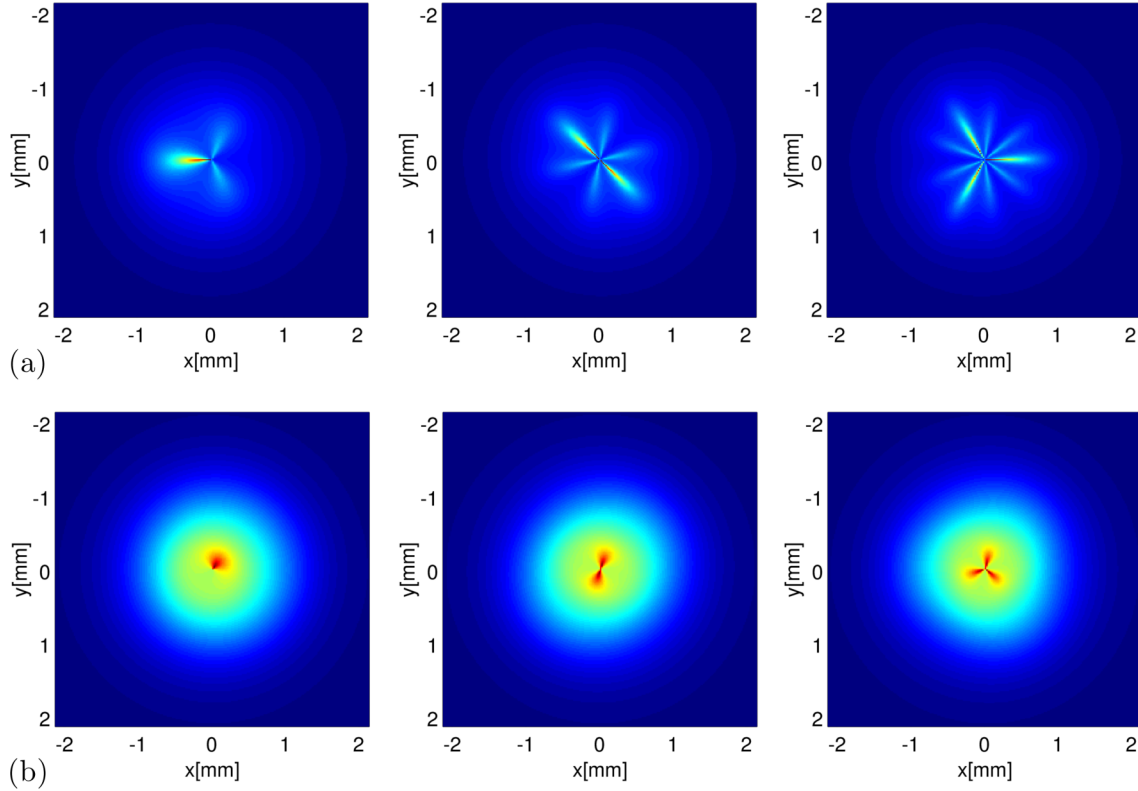


FIG. 7. The patterns for $Q = 1, 2,$ and 3 vortices (left to right), in a dissipative (a) and lossless (b) left-handed metamaterial. All parameters are the same as in Fig. 3 but the boundary condition at $z = 0$ is now a vortex for the electric field E and a homogenous background for H . The symmetries and the whole qualitative picture are the same as before, confirming that the predictions of the paper do not require preparing a vortex in magnetic field at the entry.

Since $G^{-1}(\mathbf{u} \rightarrow 0) \propto u = 0$, the bare propagator is massless. The one-loop correction $\Sigma^{(1)}$ therefore gives a cutoff-dependent mass $r_M \sim \ln \Lambda$, which could be absorbed in the overall normalization of the propagator. As we declared in the main text, the one-loop self-energy does not do much.

The crucial diagram $\Sigma^{(2)}$, the popular watermelon diagram, cannot be calculated exactly. It can be evaluated in the regime of small external momentum \mathbf{u} , i.e., when $u < u', u''$; more precisely, we can look at the regime when $u < u_0 < u', u''$ for some (arbitrary) scale u_0 and expand in a series in u/u_0 . Let us denote such an entity by $\Sigma^{(2)}(\mathbf{u}; u_0)$: it contains enough information for our purposes: we are interested mainly in angular integrations which determine the symmetry, and these can be done exactly as they separate from the integrations over the module u in the small- u limit. For $\mathbf{u} = 0$ the watermelon diagram reads (with $\int \equiv \int_0^{2\pi} d\phi' \int_0^{2\pi} d\phi'' \int du' \int du''$)

$$\Sigma^{(2)} \approx \int \frac{G(\mathbf{u}')G(\mathbf{u}'')}{v} \{e^{iQ[\pi/2+(\phi-\phi'-\phi'')]}\} [\cos(a\Lambda - \pi Q) - \sin(a\Lambda)] + e^{-iQ[\pi/2+(\phi-\phi'-\phi'')]}\} [\cos(a\Lambda + \pi Q) - \cos(a\Lambda)],$$

$$v \equiv \sqrt{(u')^2 + (u'')^2 - 2u'u'' \cos(\phi - \phi' - \phi'')}. \quad (\text{C7})$$

One angular integration is performed by taking $\phi' \mapsto \phi' + \phi''$, which makes the ϕ'' integral completely trivial, and the ϕ' integral is evaluated in terms of the elliptic integrals E and K . The outcome is finite, hence it is observable (not only at the cutoff scale) and reads

$$\Sigma^{(2)}(0; u_0) = \left(\frac{2\pi}{a\Gamma(Q/2)}\right)^3 e^{3iQ/2} \cos(3Q\phi/2)^2 \int du' \int du'' \frac{[(u')^2 - (u'')^2](u' + u'')E\left(-\frac{4u'u''}{(u'+u'')^2}\right)}{(u')^2(u'')^2[(u')^2 - a^2][(u'')^2 - a^2][(u')^2 - (u'')^2]}$$

$$= \frac{1}{4\pi} \left(\frac{2\pi}{a\Gamma(Q/2)}\right)^3 e^{3iQ/2} \cos(3Q\phi/2)^2 (a^{3/2} - 1/\Lambda^{3/2}) + O(1/\Lambda^2). \quad (\text{C8})$$

In particular, this means that a nontrivial mass term is acquired, of the order $a^{3/2}$. This mass is anisotropic, and the factor $\cos(3Q\pi/2)^2$ is all we need for the $3Q$ polygon. The leading correction in u/u_0 is in fact inessential for the symmetry, but it is

important as it contains a nonzero imaginary part, introducing a finite lifetime for such patterns. It reads

$$\begin{aligned}\Sigma^{(2)}(u; u_0) &= \int \frac{G(\mathbf{u}')G(\mathbf{u}'')}{w} \{e^{iQ[\pi/2+(\phi'-\phi'')]}[\cos(a\Lambda - \pi Q) - \sin(a\Lambda)] + e^{-iQ[\pi/2+(\phi'-\phi'')]}[\cos(a\Lambda + \pi Q) - \cos(a\Lambda)]\} \\ &= \frac{1}{4\pi} \left(\frac{2\pi}{a\Gamma(Q/2)} \right)^3 e^{3iQ/2} \left(\frac{2ia^{3/2}}{\pi} \sin(3Q\phi/2) + \frac{2\Lambda^{3/2}}{\pi} \cos(3Q\phi/2) \right), \\ w &\equiv \sqrt{(u')^2 + (u'')^2 - 2u'u'' \cos(\phi' - \phi'') - 2u[u' \cos(\phi - \phi') + u'' \cos(\phi - \phi'')]}.\end{aligned}\quad (\text{C9})$$

At leading order, this tedious expression behaves like $1/r^3$, falling off much quicker than the bare propagator (18), which goes as $1/\sqrt{r}$ (most obvious from the Bessel-function form of the real-space solution), suggesting that the shape of the vortex, which is mainly determined by long-distance behavior, is not much influenced by the finite- u correction to $\Sigma^{(2)}$.

-
- [1] V. G. Veselago, The electrodynamics of substances with simultaneously negative values of epsilon and mu, *Usp. Fiz. Nauk* **92**, 517 (1967) (in Russian).
- [2] R. A. Shelby, D. R. Smith, and S. Schultz, Experimental verification of a negative index of refraction, *Science* **292**, 77 (2001).
- [3] T. J. Yen, W. J. Padilla, N. Fang, D. C. Vier, D. R. Smith, J. B. Pendry, D. N. Basov and D. Zhang Terahertz magnetic response from artificial materials, *Science* **303**, 1494 (2004).
- [4] A. A. Zharov, I. V. Shadrivov and Y. S. Kivshar, Nonlinear Properties of Left-Handed Metamaterials, *Phys. Rev. Lett.* **91**, 037401 (2003).
- [5] M. Lapine, I. V. Shadrivov and Y. S. Kivshar, Colloquium: Nonlinear metamaterials, *Rev. Mod. Phys.* **86**, 1093 (2014); A. Baev, P. N. Prasad, H. Agren, M. Samoć and M. Wegener, Metaphotonics: An emerging field with opportunities and challenges, *Phys. Rep.* **594**, 1 (2015).
- [6] I. V. Shadrivov, N. A. Zharova, A. A. Zharov and Y. S. Kivshar, Nonlinear transmission and spatiotemporal solitons in metamaterials with negative refraction, *Optics Express* **13**, 1291 (2005).
- [7] I. V. Shadrivov and Y. S. Kivshar, Spatial solitons in nonlinear left-handed metamaterials, *J. Opt. A: Pure Appl. Opt.* **7**, S68 (2005).
- [8] I. V. Shadrivov, A. A. Sukhorukov, Y. S. Kivshar, A. A. Zharov, A. D. Boardman and P. Egan, Nonlinear surface waves in left-handed materials, *Phys. Rev. E* **69**, 016617 (2004).
- [9] S. Wen, Y. Wang, W. Su, Y. Xiang, X. Fu and D. Fan, Modulation instability in nonlinear negative-index material, *Phys. Rev. E* **73**, 036617 (2006).
- [10] S. Z. Silahlı, W. Walasik and N. M. Litchinitser, Modulation instability of structured-light beams in negative-index metamaterials, *J. Opt.* **18**, 054010 (2016).
- [11] N. L. Tsitsas, N. Rompotis, I. Kourakis, P. G. Kevrekidis and D. J. Frantzeskakis, Higher-order effects and ultrashort solitons in left-handed metamaterials, *Phys. Rev. E* **79**, 037601 (2009).
- [12] I. V. Shadrivov, A. A. Sukhorukov and Y. S. Kivshar, Guided modes in negative-refractive-index metamaterials, *Phys. Rev. E* **67**, 057602 (2003).
- [13] A. Namdar, I. V. Shadrivov and Y. S. Kivshar, Backward Tamm states in left-handed metamaterials, *App. Phys. Lett.* **89**, 114104 (2006).
- [14] H. Liu, J. Lei, G. Jiang, X. Guan L. Ji, and Z. Ma, Observation of tunable nonlinear effects in an analogue of superconducting composite right/left hand filter, *Sci. Rep.* **5**, 14846 (2015).
- [15] L. Fan, Z. Chen, Y.-C. Deng, J. Ding, H. Ge, S.-Y. Zhang, Y.-T. Yang and H. Zhang, Nonlinear effects in a metamaterial with double negativity, *App. Phys. Lett.* **105**, 041904 (2014).
- [16] A. O. Korotkevich, K. E. Rasmussen, G. Kovačič, V. Roytburd, A. I. Maimistov and I. R. Gabitov, Optical pulse dynamics in active metamaterials with positive and negative refractive index, *J. Opt. Soc. Am. B* **30**, 1077 (2013).
- [17] M. Saha, A. K. Sarma, Modulation instability in nonlinear metamaterials induced by cubic-quintic nonlinearities and higher order dispersive effects, *Opt. Commun.* **291**, 321 (2005).
- [18] Y. Shen, P. G. Kevrekidis, G. P. Veldes, D. J. Frantzeskakis, D. DiMarzio, X. Lan, and V. Radisic, From solitons to rogue waves in nonlinear left-handed metamaterials, *Phys. Rev. E* **95**, 032223 (2017).
- [19] M. Liu, D. A. Powell, I. V. Shadrivov, M. Lapine and Y. S. Kivshar, Spontaneous chiral symmetry breaking in metamaterials, *Nat. Comm.* **5**, 4441 (2014).
- [20] M. A. Gorchach, D. A. Dobrykh, A. P. Slobozhanyuk, P. A. Belov and M. Lapine, Nonlinear symmetry breaking in photometamaterials, *Phys. Rev. B* **97**, 115119 (2018).
- [21] M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, *Rev. Mod. Phys.* **65**, 851 (1993).
- [22] M. I. Rabinovich, A. B. Ezersky, and P. D. Weidman, *The Dynamics of Patterns* (World Scientific, Singapore, 2000).
- [23] L. I. Pismen, *Vortices in Nonlinear Fields* (Oxford University, London, 1999).
- [24] H. Kleinert, *Gauge Fields in Condensed Matter Physics* (World Scientific, Singapore, 1989).
- [25] P. W. Anderson, Two new vortex liquids, *Nature Physics* **3**, 160 (2007).
- [26] A. L. Fetter, Rotating trapped Bose-Einstein condensates, *Rev. Mod. Phys.* **81**, 647 (2009).
- [27] J. D. Sau and S. Sachdev, Mean-field theory of competing orders in metals with antiferromagnetic exchange interactions, *Phys. Rev. B* **89**, 075129 (2014).
- [28] V. L. Berezinsky, Razrushenie dalnego poryadka v odnomernykh i dvumernykh sistemakh s neprerivnoy gruppy simmetrii I. Klassicheskije sistemy, *JETF* **59**, 907 (1970) (in Russian); J. Kosterlitz and D. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, *J. Phys. C* **6**, 1181 (1973).
- [29] R. W. Boyd, *Nonlinear Optics* (Academic, New York, 2008).
- [30] J. L. Cieslinski and T. Nikiciuk, A direct approach to the construction of standard and non-standard Lagrangians for dissipative-like dynamical systems with variable coefficients, *J. Phys. A* **43**, 175205 (2010).

- [31] T. Shah, R. Chattopadhyay, K. Vaidya and S. Chakraborty, Conservative perturbation theory for nonconservative systems, *Phys. Rev. E* **92**, 062927 (2015).
- [32] M. Čubrović and M. S. Petrović, Quantum criticality in photorefractive optics: Vortices in laser beams and antiferromagnets, *Phys. Rev. A* **96**, 053824 (2017).
- [33] P. Alitalo, S. Maslovski and S. Tretyakov, Experimental verification of the key properties of a three-dimensional isotropic transmission-line superlens, *J. Appl. Phys.* **99**, 124910 (2006).
- [34] A. B. Kozyrev and D. W. van der Weide, Trains of envelope solitons in nonlinear left-handed transmission line media, *Appl. Phys. Lett.* **91**, 254111 (2007).
- [35] A. D. Boardman, N. King, R.-C. Mitchell-Thomas, V. N. Malnev, and Y. G. Rappoport, Gain control and diffraction-managed solitons in metamaterials, *Metamaterials* **2**, 145 (2008).

Quantum criticality in photorefractive optics: Vortices in laser beams and antiferromagnets

Mihailo Čubrović*

Scientific Computing Laboratory, Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

Milan S. Petrović

*Institute of Physics, P.O. Box 57, 11001 Belgrade, Serbia
and Texas A&M University at Qatar, P.O. Box 23874, Doha, Qatar
(Received 5 December 2016; published 9 November 2017)*

We study vortex patterns in a prototype nonlinear optical system: counterpropagating laser beams in a photorefractive crystal, with or without the background photonic lattice. The vortices are effectively planar and have two “flavors” because there are two opposite directions of beam propagation. In a certain parameter range, the vortices form stable equilibrium configurations which we study using the methods of statistical field theory and generalize the Berezinsky-Kosterlitz-Thouless transition of the XY model to the “two-flavor” case. In addition to the familiar conductor and insulator phases, we also have the perfect conductor (vortex proliferation in both beams or “flavors”) and the frustrated insulator (energy costs of vortex proliferation and vortex annihilation balance each other). In the presence of disorder in the background lattice, a phase appears which shows long-range correlations and absence of long-range order, thus being analogous to glasses. An important benefit of this approach is that qualitative behavior of patterns can be known without intensive numerical work over large areas of the parameter space. The observed phases are analogous to those in magnetic systems, and make (classical) photorefractive optics a fruitful testing ground for (quantum) condensed matter systems. As an example, we map our system to a doped $O(3)$ antiferromagnet with \mathbb{Z}_2 defects, which has the same structure of the phase diagram.

DOI: [10.1103/PhysRevA.96.053824](https://doi.org/10.1103/PhysRevA.96.053824)**I. INTRODUCTION**

Nonlinear and pattern-forming systems [1–3] have numerous analogies with strongly correlated systems encountered in condensed matter physics [4,5], and on the methodological level they are both united through the language of field theory, which has become the standard language to describe strongly correlated electrons [6,7] as well as nonlinear dynamical systems [8]. In the field of pattern formation, some connections to condensed matter systems have been observed; see, e.g., Ref. [4]. More recently, extensive field-theoretical studies of laser systems were performed, e.g., Refs. [9–12], and also compared to experiment [13]. However, this topic is far from exhausted and we feel many analogies between quantum many-body systems and pattern-formation dynamics remain unexplored and unexploited. In particular, nonlinear optical systems and photonic lattices are flexible and relatively cheap to build [3] and they can be used to “simulate” a broad spectrum of phenomena concerning band structure, spin ordering, and conduction in strongly correlated electron systems; some of the work in this direction can be found in Refs. [14,15].

Our goal is to broaden the connections between the strongly correlated systems and nonlinear optics and to put to work the mighty apparatus of field theory to study the patterns in a nonlinear optical system from the viewpoint of phase transition theory: Pattern dynamics in certain cases shows critical behavior which is analogous to phenomena seen in magnetic systems. To that end, we use the formalism of perturbative field theory and renormalization group analysis but we also perform numerical simulations from the first principles, i.e., directly integrating the equations of motion to provide an *independent*

check of our main conclusions. We also establish a connection to an $O(3)$ antiferromagnetic model which is encountered in the study of strongly correlated electron systems. The analogy is not just qualitative: We construct the phase diagrams of both systems and find they have the same structure. Introducing disorder into the system further enriches the physics, and it is physically motivated: In optics, disorder is rooted in the imperfections of the photonic lattice, and in magnetic systems it comes from the quenched spin impurities which are regularly found in realistic samples. It turns out that in both cases a glassy phase arises. This is another important research topic and it is again appealing to realize glasses in photonic lattice systems, where the parameters are easy to tune.

A. On topology and vortices

The key phenomenon which governs the phenomenology of the systems studied is the existence of topologically nontrivial solutions or *topological solitons* [16]: These are the solutions which map the physical boundary of the system to the whole configuration space of the field, so one explores all field configurations by “going around the system.” For example, in a two-dimensional system (in the x - y plane) with $U(1)$ phase symmetry, the configuration space is a circle (the phase lies between 0 and 2π) and the boundary of the physical space (i.e., the two-dimensional plane) is again a circle, the “boundary” of the plane at infinity. The topological soliton is a pattern of the $U(1)$ field which spans the whole phase circle (its phase goes from 0 to 2π), as one moves around the far-away circle in the x - y plane. Of course, this is the vortex—the most famous and best studied topological configuration. Similar logic leads to the classification of topological defects of other, more complicated symmetry groups. A potential source of confusion is that in nonlinear dynamics and theory of partial

*mcubrovic@gmail.com

differential equations, the “integrable” solutions, i.e., linearly (often also nonlinearly) stable solutions which can be obtained by inverse scattering or similar methods and which propagate through each other without interacting, are also called solitons, or more precisely *dynamical solitons*. In optics, they are often called spatial solitons. Dynamical solitons in nonlinear optics are a celebrated and well-studied topic [17–22]; they show an amazing variety of patterns and phenomena like localization, Floquet states [14], etc. But in general they do not have a topological charge. In contrast, topological solitons carry a topological charge (winding number for vortices) and their stability is rooted in topological protection (conservation of topological charge).

The phenomenon of vortices is perhaps best known in three spatial dimensions. The phase of the wave function can wind, forming a vortex line. These vortices are stable when the phase symmetry is broken by magnetic field. Famously, vortices may coexist with the superconducting order ($U(1)$ symmetry breaking) in type-II superconductors or exist only in the normal phase, upon destroying the superconductivity (type I). The primary example in two spatial dimensions is the vortex unbinding phase transition of infinite order found by Berezinsky *et al.* for the planar XY model [23]. The formal difference between the two- and three-dimensional vortices is that the latter gives rise to an emergent gauge field; this does not happen in the XY -like system in two dimensions [24]. While the nonlinear optical system we study is three-dimensional, its geometry and relaxational dynamics make it natural to treat it as a $(2 + 1)$ -dimensional system (the x and y coordinates are spatial dimensions, the z direction has the formal role of time, and physical time t has the role of a parameter). We therefore have a similar situation to the XY model: pointlike vortices in the plane (and no gauge field).

Vortex matter is known to emerge in liquid helium [25], Bose-Einstein condensates [26], and magnetic systems [27]. The basic mechanisms of vortex dynamics are thus well known. However, unusual physics can arise if the system has multiple components and each of them can form vortices which mutually interact. This is precisely our situation: We have a system of two laser beams propagating in opposite directions, and we will compare it to a two-component antiferromagnet. So far, such situations have been explored in multicomponent superconductors [28] which have attracted some attention, as they can be realized in magnesium diboride [29]. But these are again bulk systems, not planar. Vortices in planar multicomponent systems have not been very popular, an important exception being the two-component Bose-Einstein condensates of Ref. [30], which were found to exhibit complex vortex dynamics; in these systems, contrary to our case, the two components have an explicit attractive interaction, unlike our case where they interact indirectly, by coupling to the total light intensity (of both components).

B. The object of our study

In this paper, we study phases and critical behavior of topological configurations (vortices and vortex lattices) in a specific and experimentally realizable nonlinear optical system: laser beams counterpropagating (CP) through a photorefractive (PR) crystal. This means we have an elongated PR crystal

(with one longitudinal and two transverse dimensions) and two laser beams shone onto each end. We thus effectively have two fields, one forward propagating and one backward propagating. The optical response of the crystal depends nonlinearly on the *total* intensity of both beams, which means the beams effectively interact with each other. This system has been thoroughly investigated for phenomena such as dynamical solitons [17,31,32], vortex stability on the photonic lattice [18–20,33–36], and global rotation [37]. We will see that the CP beams are an analog of the two-component planar antiferromagnet, which can further be related to some realistic strongly correlated materials [38–40]. The two beams are now equivalent to two sublattices which interact through a lattice deformation or external field. The PR crystal is elongated and the axial propagation direction has the formal role of time, which has a finite span, the length of the crystal. For the antiferromagnet, the third axis is the usual imaginary time compactified to the radius $1/T$, i.e., inverse temperature. Both systems contain vortices as topological defects, i.e., solutions with integer topological charge. In the PR optical system, vortices arise as a consequence of the $U(1)$ symmetry of the electromagnetic field. In the antiferromagnets we consider, the $O(3)$ symmetry of the antiferromagnet gives rise to \mathbb{Z}_2 -charged defects, which exhibit the same interactions as the vortices. The optical system is not subject to noise (i.e., it lives at zero temperature), and thus the criticality we talk about is obviously not the same as thermodynamic phase transitions. Phase transitions happen upon varying the parameters, not temperature, so they may be described as quantum critical phenomena in the broad sense taken in Ref. [38]—any critical behavior controlled not by thermal fluctuations but by parameter dependence.

In the PR counterpropagating beam system, our focus are the vortices but in order to study them we need to do some preparational work. We first recast the system in Lagrangian and then in Hamiltonian form so it can be studied as a field theory, which depends parametrically on the time t . Then we consider the time dynamics of the system and show that in a broad parameter range the patterns relax to a static configuration which can be studied within *equilibrium* field theory. Along the way, we also study the stability of topologically trivial (vortex-free) configurations and then consider the phases of the static vortex configurations. The analytical insight we obtain also allows us to avoid overextensive numerics—analytical construction of the phase diagram tells us which patterns can in principle be expected in different corners of the parameter space. By “blind” numerical approach, this result could only be found through many runs of the numerics.

In the antiferromagnetic spin system, the nontopological excitations are simple: They are spin waves, perturbed away from the noninteracting solution by the quartic terms in the potential. There are no dynamical solitons. But we will see that topological excitations lead to a phase diagram which, after reasonable approximations, can be *exactly* mapped to the phase diagram of the photorefractive crystal. The reason is that both can be reduced to an effective Hamiltonian for a *two-component* vortex system; i.e., every vortex has two charges or two “flavors.” In the photorefractive crystal it happens naturally, as there are two beams, forward and

backward propagating. In the Heisenberg antiferromagnet it is less obvious and is a crucial consequence of the collinearity of the spin pattern. We will focus on common properties of the two systems and map the phase diagrams onto each other. In the antiferromagnetic system, different phases are separated by quantum phase transitions—phase transitions driven by the quantum fluctuations instead of temperature.

On disorder

It is known that impurities pin the vortices and stabilize them. This leads to frozen dynamics even though no symmetry is broken, the phenomenon usually associated with glasses. In simple systems such as the Ising model with disorder, one generically has two phases: The disordered (paramagnetic) phase remains and the ordered (magnetic) phase is replaced by a regime with algebraic correlations and no true order. In many cases, such phases are called glasses. The exact definition of a glass is lacking; normally, they show (i) long-range correlations, (ii) absence of long-range order, i.e., of a nonzero macroscopic order parameter, and (iii) “frozen dynamics,” i.e., free energy landscape with numerous local minima in which the system can spend a long time [41,42]. While the most popular example are probably spin glasses in Ising-like models such as Sherrington-Kirkpatrick and Edwards-Anderson models, glasses are also known to appear in the XY model with disorder in two dimensions, the Cardy-Ostlund model, which postulates both random couplings and a random magnetic field [43–45]. Our model is essentially a two-flavor generalization of the XY model, although in order to solve it we need to simplify it. According to Refs. [43–45], the details differ depending on how the disorder is implemented, but the two-phase system (paramagnetic, i.e., disordered, and glass) is ubiquitous. In the two-component version, the phase diagram becomes richer, and on top of the glassy phase and the insulator (disordered) phase we find a few other phases. In nonlinear optics, the topic of random lasers has attracted considerable attention [9–12,46]. Here one has a complex version of the XY model, with the additional complication that not only phase but also amplitude is free to vary, but only with random couplings (no random field). On top of the glassy and the disordered phase, one or two additional phases appear.

In the presence of disorder, the relation to magnetic systems in condensed matter physics is very inspiring, since a number of complex materials show different ordering mechanisms (spin and charge density waves, superconductivity, etc.) in parallel with significant influence of disorder. Just as in the disorder-free case, we are particularly interested in possible spin-glass phenomena in doped insulating $O(3)$ antiferromagnets [39,40,47–49] and in the last section we will discuss also the spin-glass phase in such systems.

C. The plan of the paper

The structure of the paper is as follows. In the next section, we describe the dynamical system which lies at the core of this paper: counterpropagating laser beams in a photorefractive crystal. We give the equations of motion and repackage them in the Lagrangian form. In Sec. III, we study the vortex dynamics: We construct the vortex Hamiltonian and classify the order parameters. Then we study the renormalization group (RG)

flow and obtain the phase diagram. Finally, we discuss the important question of how to recognize the various phases in experiment: What do the light intensity patterns look like and how do they depend on the tunable parameters? Section IV brings the same study for the system with disorder. After describing the disordered system, we perform the replica trick for the disordered vortex Hamiltonian and solve the saddle-point equations to identify the phases and order parameters, again refining the results with RG calculations. The fifth section takes a look at a doped collinear antiferromagnet, a model encountered in the description of many strongly coupled materials, and shows how the dynamics of topological solitons is again described by a two-flavor vortex Hamiltonian. We discuss the relation between the phase diagrams of the two systems and the possibilities of modeling the condensed matter systems experimentally by the means of photorefractive optics. The last section sums up the conclusions. In Appendix A, we describe the numerical algorithm we use to check the analytical results for the phase diagram. In Appendix B, we show in detail that the CP beams are capable of reaching equilibrium (i.e., stop changing in time)—if they would not, the application of equilibrium field theory would not be justified. Appendix C discusses the stability of nonvortex configurations—although somewhat peripheral to the main topic of the paper, it is useful to better understand the geometry of patterns. In Appendix D, we give the (routine) algebra that yields the vortex interaction Hamiltonian from the microscopic equations. Appendix E contains an improved mean-field theory for the clean system, which we do not use much throughout the paper but we include it for completeness (we prefer either the simplest single-vortex mean-field reasoning or the full RG analysis, which are described in the main text). Appendix F discusses an important technicality concerning the CP geometry, i.e., the specific boundary conditions of the CP beam system where the boundary conditions for one beam are given at the front face and for the other at the back face of the crystal. Appendix G contains some details on mean-field and RG calculations of the phase diagram for the dirty system: The dirty case includes some tedious algebra we feel appropriate to leave out from the main text.

II. THE MODEL OF COUNTERPROPAGATING BEAMS IN THE PHOTOREFRACTIVE CRYSTAL

We consider a photorefractive crystal of length L irradiated by two laser beams. The beams are paraxial and propagate head on from the opposite faces of the crystal in the z direction. Photorefractive crystals induce self-focusing of the beams—the vacuum (linear) wave equation is modified by the addition of a frictionlike term, so the diffusion of the light intensity (the broadening of the beam) is balanced out by the convergence of the beam onto an “attractor region.” The net result is the balance between the dissipative and scattering effects, allowing for stable patterns to form. The physical ground for this is the redistribution of the charges in the crystal due to the Kerr effect. The nonlinearity, i.e., the response of the crystal to the laser light, is contained in the change of the refraction index which is determined by the local charge density. A sketch of the system is given in Fig. 1. Before entering the crystal, the laser beams can be given any desirable pattern of both intensity

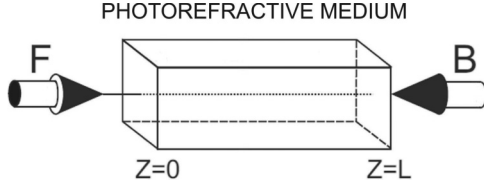


FIG. 1. Sketch of the experimental setup for the study of the CP beams in the PR crystal. The crystal has the shape of a parallelepiped, and the beams propagate along the longitudinal, z axis: the forward (F) beam from $z = 0$ to $z = L$, and the backward (B) beam the other way round. The intensity patterns are observed at the transverse faces of the crystal, at $z = 0$ and $z = L$.

and phase. In particular, one can create vortices (winding of the phase) making use of the phase masks [3] or other, more modern ways.

Assuming the electromagnetic field of the form $\mathbf{E} = e^{i\omega t + i\mathbf{q}\cdot\mathbf{r}}(F e^{ikz} + B e^{-ikz})$, we can write equations for the so-called envelopes F and B of the forward- and backward-propagating beams along the z axis (the frequency, transverse, and longitudinal momentum are denoted respectively by ω, \mathbf{q}, k). The wave equations for F and B are now

$$\begin{aligned} \pm i \partial_z \Psi_{\pm}(z; x, y; t) + \Delta \Psi_{\pm}(z; x, y; t) \\ = \Gamma E(z; x, y; t) \Psi_{\pm}(z; x, y; t), \end{aligned} \quad (1)$$

where the plus and minus signs on the left-hand side stand for the forward- and backward-propagating component of the beam amplitude doublet $\Psi \equiv (\Psi_+, \Psi_-) \equiv (F, B)$, and Γ is the dimensionless PR coupling constant. The two beams (flavors of the field Ψ) will from now on be denoted either by F/B or more often by Ψ_{\pm} . We will use α as the general flavor index for summation, e.g., $\Psi_{1\alpha} \Psi_{2\alpha} = \Psi_{1+} \Psi_{2+} + \Psi_{1-} \Psi_{2-}$. The charge field E on the right-hand side of the equation is the electric field sourced by the charges in the crystal (i.e., it does not include the external electric field of the beams). Its evolution is well represented by a relaxation-type equation [17]:

$$\begin{aligned} \frac{\tau}{1 + I(z; x, y; t)} \partial_t E(z; x, y; t) + E(z; x, y; t) \\ = - \frac{I(z; x, y; t)}{1 + I(z; x, y; t)}. \end{aligned} \quad (2)$$

Here, $I \equiv I_{\Psi} + I_x$ is the total light intensity at a given point, $I_{\Psi} \equiv |F|^2 + |B|^2$ is the beam intensity, and I_x the intensity of the fixed background. The meaning of I_x is that the crystal is all the time irradiated by some constant light source, independent of the counterpropagating beams with envelopes F, B . We will usually take a periodic lattice as the background, allowing also for the defects (missing cells) in the lattice when studying the effects of disorder. The relaxation time is τ . The time derivative $\partial_t E$ is divided by $1 + I$, meaning that the polarizability of the crystal depends on the total light intensity: Strongly irradiated regions react faster. In the numerical calculations, we solve Eqs. (1) and (2) with no further assumptions, as explained in Appendix A. For analytical results, we will need to transform them further, assuming a vortex pattern.

The equation for the charge field has no microscopic basis; it is completely phenomenological, but it excellently represents

the experimental results [3]. Notice that the derivative $\partial_t E$ in (2) is strictly negative (since intensity is non-negative): It thus has the form of a relaxation equation, and one expects that a class of solutions exists where $\partial_t E(t \rightarrow \infty) \rightarrow 0$, i.e., the system relaxes to a time-independent configuration. We show this in Appendix B; in the main text we will not discuss this issue but will simply take the findings of Appendix B for granted. Notice that there are also parameter values for which no equilibrium is reached [37,50,51].

For slow time evolution (in the absence of pulses), we can Laplace transform the equation (2) in time [$E(t) \mapsto E(u) = \int_0^{\infty} dt e^{-ut} E(t)$] to get the algebraic relation

$$\begin{aligned} E(z; x, y; u) &= - \frac{\Psi^{\dagger} \Psi + I_x - \tau E_0}{1 + \tau u + I_x + \Psi^{\dagger} \Psi} \\ &= -1 + \frac{1 + \tau u + \tau E_0}{1 + \tau u + I_x + \Psi^{\dagger} \Psi}. \end{aligned} \quad (3)$$

The original system (1) can now be described by the Lagrangian:

$$\begin{aligned} \mathcal{L} &= i \Psi^{\dagger} \sigma_3 \partial_z \Psi - |\nabla \Psi|^2 + \Gamma \Psi^{\dagger} \Psi \\ &\quad - \Gamma(1 + \tau E_0 + \tau u) \ln(1 + \tau u + I_x + \Psi^{\dagger} \Psi), \end{aligned} \quad (4)$$

where σ_3 is the Pauli matrix $\sigma_3 = \text{diag}(1, -1)$. One can introduce the effective potential

$$V_{\text{eff}}(\Psi^{\dagger}, \Psi) = -\Gamma \ln \frac{e^{\Psi^{\dagger} \Psi}}{(1 + \tau u + I_x + \Psi^{\dagger} \Psi)^{1 + \tau(E_0 + u)}}, \quad (5)$$

so we can write the Lagrangian as $\mathcal{L} = i \Psi^{\dagger} \sigma_3 \partial_z \Psi - |\nabla \Psi|^2 - V_{\text{eff}}(\Psi^{\dagger}, \Psi)$. This is the Lagrangian of a nonrelativistic field theory (a nonlinear Schrödinger field equation) in $2 + 1$ dimensions (x, y, z), where the role of time is played by the longitudinal distance z and the physical time t (or u upon the Laplace transform) is a parameter. The span of the z coordinate $0 < z < L$ will influence the behavior of the system, while the dimensions of the transverse plane are not important for the effects we consider.

Our main story is now the nature and interactions of the topologically nontrivial excitations in the system (4). A task which is in a sense more basic, the analysis of the topologically trivial vacua of (4) and perturbative calculation of their stability, is not of our primary interest now, in part because this was largely accomplished by other methods in Refs. [31,32]. We nevertheless give a quick account in Appendix C; first, because some conclusions about the geometry of the patterns can be carried over to vortices, and second, to give another example of applying the field-theoretical formalism whose power we wish to demonstrate and popularize in this paper.

III. VORTICES AND MEAN FIELD THEORY OF VORTEX INTERACTIONS

A. The classification of topological solutions and the vortex Hamiltonian

Now we discuss the possible topological solitons in our system. Remember once again that they differ from dynamical

solitons such as those studied in Ref. [17] and references therein. In order to classify the topologically nontrivial solutions, consider first the symmetries of the Lagrangian (4). It describes a doublet of two-dimensional (2D) complex fields which interact solely through the phase-invariant total intensity $I = \Psi^\dagger \Psi$ (and the spatial derivative term $|\nabla \Psi|^2$), while in the kinetic term $\Psi^\dagger \sigma_3 \partial_z \Psi$ the two components have opposite signs of the “time” derivative, so this term cannot be reduced to a functional of I . The intensity I has the symmetry group $SU(2)$ (the isometry group of the three-dimensional sphere in Euclidean space) and the kinetic term has the group $SU(1,1)$ (the transformations which leave the combination $|F|^2 - |B|^2$ invariant, i.e., the isometry of the hyperboloid). The intersection of these two is the product $U(1)_F \otimes U(1)_B$: The forward- and backward-propagating doublet (F, B) has phases $\theta_{F,B}$ which can be transformed independently, as $\theta_{F,B} \mapsto \theta_{F,B} + \delta\theta_{F,B}$.

The classification of possible topological solitons is straightforward from the above discussion [52]. They can be characterized in terms of homotopy groups. We remind readers that the homotopy group π_n of the group G is the group of transformations which map the group manifold of G onto the n -dimensional sphere S_n . In D -dimensional space, the group π_{D-1} therefore classifies what a field configuration looks like from far away (from infinity): It classifies the mappings from the manifold of the internal symmetry group of the system to the spherical “boundary shell” in physical space at infinity. Since the beams in our PR crystal effectively see a two-dimensional space (we regard z as time), we need the first homotopy group π_1 to classify the topological solitons. Since $\pi_1(U(1)) = \pi_1(S_1) = \mathbb{Z}$ and $\pi_1(\mathbb{G} \otimes \mathbb{G}) = \pi_1(\mathbb{G}) \otimes \pi_1(\mathbb{G})$ for any group \mathbb{G} , the topological solutions are flavored vortices, and the topological charge is the pair of integers $\{Q_F, Q_B\}$.

Let us now derive the effective interaction Hamiltonian for the vortices and study the phase diagram. In principle, this story is well known: For a vortex at \mathbf{r}_0 , in the polar coordinates (r, ϕ) , we write $\Psi(\mathbf{r}) = \psi \exp(i\theta(\mathbf{r}))$ for $|\mathbf{r} - \mathbf{r}_0|/|\mathbf{r}_0| \ll 1$, and a vortex of charge Q has $\theta(\phi) = Q\phi/2\pi$. In general the phase has a regular and a singular part, $\nabla \Psi = \psi(\nabla \delta\theta + \nabla \times \zeta \mathbf{e}_z)$, where finally $\zeta = Q \ln |\mathbf{r} - \mathbf{r}_0|$. The difference in the CP beam system lies in the existence of two beam fields (flavors) and the nonconstant amplitude field $\psi_{\pm}(r)$, so the vortex looks like

$$\Psi_{0\pm}(\mathbf{r}) = \psi_{0\pm}(r) e^{i\delta\theta_{\pm}(\phi) + i\theta_{0\pm}(\phi)}. \quad (6)$$

When we insert this solution into the equations of motion (or, equivalently, the Lagrangian), it is just a matter of algebra to obtain the vortex Hamiltonian, analogous to the well-known one but with two components (flavors) and their interaction. We refer the reader to the Appendix D for the full derivation. The outcome is perhaps expected: We get the straightforward generalization of the familiar Coulomb gas picture for the XY model where all interactions of different flavors, F - F , B - B , and F - B , are allowed. In order to write the Hamiltonian (and further manipulations with it) in a concise way, it is handy to introduce shorthand notation $\vec{Q} \equiv (Q_+, Q_-)$, $\vec{Q}_1 \cdot \vec{Q}_2 \equiv Q_{1+}Q_{2+} + Q_{1-}Q_{2-}$, and $\vec{Q}_1 \times \vec{Q}_2 \equiv Q_{1+}Q_{2-} - Q_{1-}Q_{2+}$. For the self-interaction within a vortex \vec{Q}_1 , we have $\vec{Q}_1 \cdot \vec{Q}_1 =$

$Q_{1+}^2 + Q_{1-}^2$ but $\vec{Q}_1 \times \vec{Q}_1 \equiv Q_{1+}Q_{1-}$ (i.e., there is a factor of 2 mismatch with the case of two different vortices). Now for vortices at locations $\mathbf{r}_i, i = 1, \dots, N$ with charges $\{Q_{i+}, Q_{i-}\}$ we get

$$\begin{aligned} \mathcal{H}_{\text{vort}} = & \sum_{i < j} (g \vec{Q}_i \cdot \vec{Q}_j + g' \vec{Q}_i \times \vec{Q}_j) \ln r_{ij} \\ & + \sum_i (g_0 \vec{Q}_i \cdot \vec{Q}_i + g_1 \vec{Q}_i \times \vec{Q}_i). \end{aligned} \quad (7)$$

The meaning of the Hamiltonian (7) is obvious. The first term is the Coulomb interaction of vortices; notice that only like-flavored charges interact through this term (because the kinetic term $|\nabla \Psi|^2$ is homogenous quadratic). The second term is the forward-backward interaction, also with Coulomb-like (logarithmic) radial dependence. This interaction comes from the mixing of the F and B modes in the fourth term in Eq. (D2), and it is generated, as we commented in Appendix D, when the amplitude fluctuations $\delta\psi_\alpha(r)$, which couple linearly to the phase fluctuations, are integrated out. In a system without amplitude fluctuations, i.e., classical spin system, this term would not be generated. The third and fourth terms constitute the energy of the vortex core. The self-interaction constants g_0, g_1 are of course dependent on the vortex core size and behave roughly as $g \ln a/\epsilon, g' \ln a/\epsilon$, where ϵ is the UV cutoff. The final results will not depend on ϵ , as expected, since g_0, g_1 can be absorbed in the fugacity y (see the next subsection). Expressions for the coupling constants in terms of original parameters are given in (D11).

In three space dimensions, vortices necessitate the introduction of a gauge field [24] which, in multicomponent systems, also acquires the additional flavor index [28,53]. In our case, there is no emergent gauge field and the whole calculation is a rather basic exercise at the textbook level but the results are still interesting in the context of nonlinear optics and analogies to magnetic systems: They imply that the *phase* structure (vortex dynamics) can be spotted by looking at the *intensity* patterns (light intensity I or local magnetization \mathcal{M} ; see the penultimate section).

B. The phase diagram

1. The mean-field theory for vortices

The phases of the system can be classified at the mean field level, following, e.g., Refs. [24,41]. In order to do that, one should construct the partition function, assuming that well-defined time-independent configuration space exists. We have already mentioned the question of equilibration and address it in detail in Appendix B. Knowing that the system reaches equilibrium (in some part of the parameter space), we can count the ways in which a system of vortices can be placed in the crystal—this is by definition the partition function \mathcal{Z} . First, the number of vortices N can be anything from 0 to infinity; second, the vortex charges can be arbitrary; and finally, the number of ways to place each vortex in the crystal is simply the total surface section of the crystal divided by the size of the vortex. Then, each vortex carries a Gibbs weight proportional to the energy, i.e., the vortex Hamiltonian (7) for a single

vortex.¹ Let us focus first on a single vortex. If the vortex core has linear dimension a and the crystal cross section linear dimension Λ , the vortex can be placed in any of the $(\Lambda/a)^2$ cells (and in the mean-field approach we suppose the vortex survives all the way along the crystal, from $z = 0$ to $z = L$, so there is no additional freedom of placing it along some subinterval of z). This gives

$$\mathcal{Z} = \sum_{Q_+, Q_-} \left(\frac{\Lambda}{a}\right)^2 e^{-L\mathcal{H}_1} = \sum_{Q_+, Q_-} e^{2\ln\frac{\Lambda}{a} - L(g\vec{Q} \cdot \vec{Q}_+ + g'\vec{Q} \times \vec{Q}_-) \ln\frac{\Lambda}{a}}. \quad (8)$$

Remember that \mathcal{H} is energy density along the z axis, so it appears multiplied by L . The factor $\ln(\Lambda/a)$ in the second term of the exponent comes from the Coulomb potential of a single vortex (in a plane of size Λ). The exponent can be written as $-L\mathcal{F}^{(1)}$, with $\mathcal{F}^{(1)} = \mathcal{H}_1 - (1/L)S_1$, recovering the relation between the free energy $\mathcal{F}^{(1)}$ and entropy $S^{(1)}$ of a single vortex. The entropy comes from the number of ways to place a vortex of core size a in the plane of size $\Lambda \gg a$: $S \sim \ln(\Lambda/a)^2$. Suppose for now that elementary excitations have $|Q_{\pm}| \leq 1$, as higher values increase the energy but not the entropy, so they are unlikely (when only a single vortex is present). Now we can consider the case of single-charge vortices with possible charges $(1, 0), (-1, 0), (0, 1), (0, -1)$, and the case of two-charge vortices where F and B charge may be of the same sign or opposite signs, $(1, 1), (-1, -1), (1, -1), (-1, 1)$:

$$\mathcal{F}_0^{(1)} = \left(g - \frac{2}{L}\right) \ln\frac{\Lambda}{a}, \quad \vec{Q} = (\pm 1, 0) \text{ or } \vec{Q} = (0, \pm 1), \quad (9)$$

$$\mathcal{F}_1^{(1)} = \left(2g - g' - \frac{2}{L}\right) \ln\frac{\Lambda}{a}, \quad (Q_+, Q_-) = (\pm 1, \mp 1), \quad (10)$$

$$\mathcal{F}_2^{(1)} = \left(2g + g' - \frac{2}{L}\right) \ln\frac{\Lambda}{a}, \quad (Q_+, Q_-) = (\pm 1, \pm 1). \quad (11)$$

Now we identify four regimes, assuming that $g, g' > 0$:²

(1) For $L > 2/g$, a vortex always has positive free energy so vortices are unstable like in the low-temperature phase of the textbook Berezinsky-Kosterlitz-Thouless (BKT) system. This is the vortex-free phase where the phase $U(1)_F \otimes U(1)_B$ does not wind. This phase we logically call *vortex insulator* in analogy with the single-flavor case.

(2) For $2/g > L > 1/(g - g'/2)$, a double-flavor vortex always has positive free energy but single-flavor vortices are stable; in other words, there is proliferation of vortices of the form $\vec{Q} = (Q_+, 0)$ or $\vec{Q} = (0, Q_-)$. This phase is like the conductor phase in a single-component XY model, and the

topological excitations exist for the reduced symmetry group, i.e., for a single $U(1)$. We thus call it *vortex conductor*; it is populated mainly by single-flavor vortices $(Q, 0), (0, Q)$.

(3) For $1/(g - g'/2) > L > 1/(g + g'/2)$, double-vortex formation is only optimal if the vortex has $Q_+ + Q_- = 0$, which corresponds to the topological excitations of the diagonal $U(1)_d$ symmetry subgroup, the reduction of the total phase symmetry to the special case $(\theta_F, \theta_B) \mapsto (\theta_F + \delta\theta, \theta_B - \delta\theta)$. In other words, vortices of the form $(Q_+, -Q_-)$ proliferate. Here, higher charge vortices may be more energetically favorable than unit-charge ones, contrary to the initial simplistic assumption, the reason being that the vortex core energy proportional to gQ_+^2 may be more than balanced out by the intravortex interaction proportional to $-g'Q_+^2$ (depending on the ratio of g and g'). This further means that there may be multiple ground states of equal energy (frustration). We thus call this case *frustrated vortex insulator* (FI); it is populated primarily with vortices of charge $(Q, -Q)$.

(4) For $1/(g + g'/2) > L$ vortex formation always reduces the free energy, no matter what the relation between Q_+ and Q_- is, and each phase can wind separately: $(\theta_F, \theta_B) \mapsto (\theta_F + \delta\theta_F, \theta_B + \delta\theta_B)$. Vortices of both flavors proliferate freely at no energy cost and for that reason we call this phase *vortex perfect conductor* (PC). We deliberately avoid the term superconductor to avoid the (wrong) association of this phase with the vortex lines and type I or type II superconductors familiar from the three-dimensional (3D) vortex systems: Remember there is no emergent gauge field for the vortices in two spatial dimensions, and we only have perfect conductivity in the sense of zero resistance for transporting the (topological) charge, but no superconductivity in the sense of breaking a gauge symmetry.

A more systematic mean-field calculation will give the phase diagram also for an arbitrary number of vortices. This is not so interesting as it already does not require much less work than the RG analysis, which is more rigorous and more accurate for this problem. For completeness, we give the multivortex mean-field calculation in Appendix E.

One might worry that the our whole approach misses the CP geometry of the problem, i.e., the fact that the Ψ_+ field has a source at $z = 0$ and the Ψ_- field at $z = L$. In Appendix F, we show that nothing is missed at the level of approximations taken in this paper, i.e., mean-field theory in this subsection and the lowest-order perturbative RG in the next one. Roughly speaking, it is because the sources are irrelevant in the RG sense—the bulk configuration dominates over the boundary terms. The appendix states this in much more precise language.

2. RG analysis

We have classified the symmetries and thus the phases of our system at the mean-field level. To describe quantitatively the borders between the phases and the phase diagram, we will perform the renormalization group (RG) analysis. Here we follow closely the calculation for conventional vortex systems [24]. We consider the fluctuation of the partition function $\delta\mathcal{Z}$ upon the formation of a virtual vortex pair at positions $\mathbf{r}_1, \mathbf{r}_2$ with charges $\vec{q}, -\vec{q}$ (with $\mathbf{r}_1 + \mathbf{r}_2 = 2\mathbf{r}$ and $\mathbf{r}_1 - \mathbf{r}_2 = \mathbf{r}_{12}$), in the background of a vortex pair at positions

¹Again, this is not generally true for out-of-equilibrium configurations but if the system reaches equilibrium, i.e., stable fixed point, this follows by usual statistical mechanics reasoning.

²One specificity of multicomponent vortices is that the coupling constants may be negative, as can be seen from (D11). In that case, the ordering of the four regimes (how they follow each other upon dialing L) changes but the overall structure remains.

$\mathbf{R}_1, \mathbf{R}_2$ (with $\mathbf{R}_1 + \mathbf{R}_2 = 2\mathbf{R}$ and $\mathbf{R}_1 - \mathbf{R}_2 = \mathbf{R}_{12}$) with charges \bar{Q}_1, \bar{Q}_2 . This is a straightforward but lengthy calculation and we state just the main steps. First, it is easy to show that the creation of single-charge vortices is irrelevant for the RG flow so we disregard it. Also, we can replace the core self-interaction constants $g_{0,1}$ with the fugacity parameter defined as $y \equiv \exp[-\beta(g_0 + g_1) \ln \epsilon]$. Here we introduce the notation $\beta \equiv L$ in analogy with the inverse temperature β in standard statistical mechanics, in order to facilitate the comparison with the literature on vortices in spin systems, and also with antiferromagnetic systems in Sec. V.³

Now from the vortex Hamiltonian $\mathcal{H}_{\text{vort}}$ the fluctuation equals (at the quadratic order in y and r)

$$\begin{aligned} \frac{\delta \mathcal{Z}}{\mathcal{Z}} = & 1 + \frac{y^4}{4} \sum_{q_{\pm}} \int dr_{12} r_{12}^3 e^{g\bar{q} \cdot \bar{q} + g'\bar{q} \times \bar{q}} \\ & \times \left[\int dr r^2 (g\bar{Q}_1 \cdot \bar{q} + g'\bar{Q}_1 \times \bar{q}) \right. \\ & \times \nabla \ln |\mathbf{R}_1 - \mathbf{r}| + (g\bar{Q}_2 \cdot \bar{q} + g'\bar{Q}_2 \times \bar{q}) \\ & \left. \times \nabla \ln |\mathbf{R}_2 - \mathbf{r}| \right]^2. \end{aligned} \quad (12)$$

Notice that ∇ is taken with respect to \mathbf{r} . The above result is obtained by expanding the Coulomb potential in r_{12} (the separation between the virtual vortices being small because of their mutual interaction) and then expanding the whole partition function (i.e., the exponent in it) in y around the equilibrium value \mathcal{Z} . The term depending on the separation r_{12} is the mutual interaction energy of the virtual charges, and the subsequent term proportional to r^2 is the interaction of the virtual vortices with the external ones (the term linear in r cancels out due to isotropy). Then by partial integration and summation over $q_{\pm} \in \{1, -1\}$ we find

$$\begin{aligned} \frac{\delta \mathcal{Z}}{\mathcal{Z}} = & 1 + y^4 [8\pi g^2 \bar{Q}_1 \cdot \bar{Q}_2 + 8\pi (g')^2 \bar{Q}_1 \cdot \bar{Q}_2 \\ & + 16\pi g g' \bar{Q}_1 \times \bar{Q}_2] I_3 \ln R_{12} \\ & + y^4 [4\pi g(g + g')(\bar{Q}_1 \times \bar{Q}_1 + \bar{Q}_2 \times \bar{Q}_2) \\ & \times I_1 + 8(g')^2 I_1] \ln \epsilon, \end{aligned} \quad (13)$$

with $I_n = \int_{\epsilon a}^{\Lambda a} dr r^{n+g+g'}$. Now, by taking into account the definition of the fugacity y , rescaling $\Lambda \mapsto \Lambda(1 + \ell)$, performing the spatial integrals, and expanding over ℓ , we can equate the bare quantities g, g', y in (7) with their corrected values in $\mathcal{Z} + \delta \mathcal{Z}$ to obtain the RG flow equations:

$$\begin{aligned} \frac{\partial g}{\partial \ell} = & -16\pi(g^2 + g'^2)y^4, & \frac{\partial g'}{\partial \ell} = & -2\pi g g' y^4, \\ \frac{\partial y}{\partial \ell} = & 2\pi(1 - g - g')y. \end{aligned} \quad (14)$$

³Of course, the physical meaning of β in our system is very different: We have no thermodynamic temperature or thermal noise, and the third law of thermodynamics is not satisfied for the ‘‘temperature’’ $1/\beta = 1/L$. We merely use the β notation to emphasize the similarity between free energies of different systems, not as a complete physical analogy.

Now let us consider the fixed points of the flow equations. If one puts $g' = 0$, they look very much like the textbook XY model RG flow, except that the fugacity enters as y^4 instead of y^2 (simply because every vortex contributes two charges). They yield the same phases as the mean-field approach as it has to be, but now we can numerically integrate the flow equations to find exact phase borders. The fugacity y can flow to zero (meaning that the vortex creation is suppressed and the vortices tend to bind) or to infinity, meaning that vortices can exist at finite density. At $y = 0$, there is a fixed line $g + g' = 1$. This line is attracting for the half-plane $g + g' > 1$; otherwise, it is repelling. There are three more attraction regions when $g + g' < 1$. First, there is the point $y \rightarrow \infty, g = g' = 0$ which has no analog in single-component vortex systems. Then, there are two regions when $g \rightarrow \infty$ and $g' \rightarrow \pm\infty$ (and again $y \rightarrow \infty$). Of course, the large g, g' regime is strongly interacting and the perturbation theory eventually breaks down, so in reality the coupling constants grow to some finite values g_*, g'_* and g_{**}, g'_{**} rather than to infinities. The situation is now the following:

(1) The attraction region of the fixed line is the vortex insulator phase: The creation rate of the vortices is suppressed to zero.

(2) The zero-coupling fixed point attracts the trajectories in the vortex perfect conductor phase: Only the fugacity controls the vortices and arbitrary charge configurations can form. Numerical integration shows that this point also has a finite extent in the parameter space.

(3) In the attraction region of the fixed point with $g_* < 0$ and $g'_* > 0$ (formally they flow to $-\infty$ and $+\infty$, respectively), same-sign F and B charges attract each other and those with the opposite sign which repel each other. This is the frustrated insulator.

(4) The fixed point with $g_{**}, g'_{**} < 0$ (formally both flow to $-\infty$) corresponds to the conductor phase.

The RG flows in the g - g' plane are given in Fig. 2. Full RG calculation is given in Fig. 2(b); for comparison, we include also the mean-field phase diagram (following from the previous subsection and Appendix E) in Fig. 2(a). In the half-plane $g + g' > 1$ every point evolves toward a different, finite point (g, g') in the same half-plane. In the other half-plane we see the regions of points moving toward the origin or toward one of the two directions at infinity. The PC phase (the attraction region of the point $(0, 0)$) could not be obtained from the mean field calculation (i.e., it corresponds to the single point at the origin at the mean field level).

It may be surprising that the coupling constants can be negative, with like charges repelling and opposite charges attracting each other. However, this is perfectly allowed in our system. In the usual XY model, the stiffness is proportional to the kinetic energy coefficient and thus has to be positive. Here, the coupling between the fluctuations of F and B beams introduces other contributions to g, g' and the resulting expressions (D11) give bare values of g, g' that can be negative, and the stability analysis of the RG flow clearly shows that for nonzero g' , the flow can go toward negative values even if starting from a positive value in some parameter range. If we fix $g' = 0$, the flow equations reproduce the ones from the single-component XY model, and the phase diagram is reduced to just the $g' = 0$ line. If we additionally suppose that

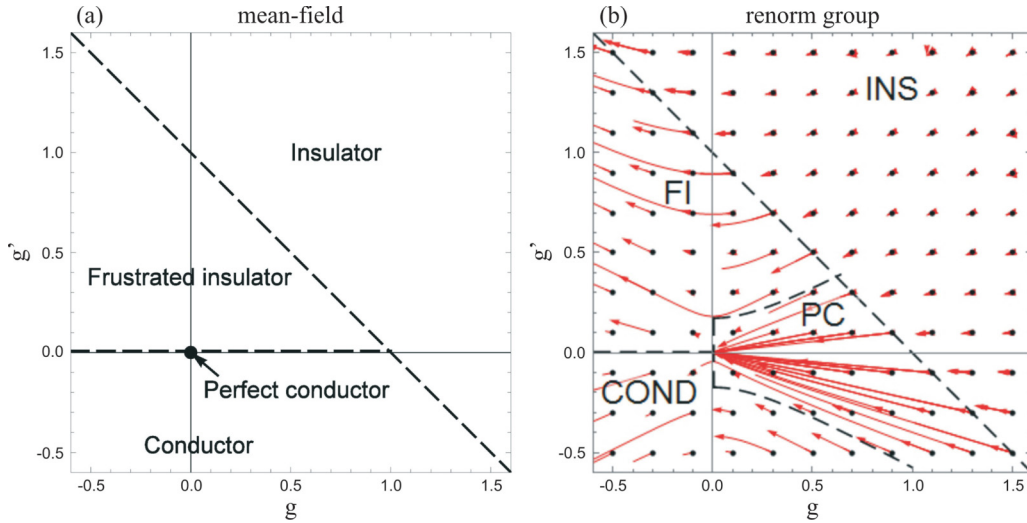


FIG. 2. Phase diagram for the clean system in the g - g' plane, at the mean-field level (a) and with RG flows (b). We show the flows for a grid of initial points, denoted by black dots; red lines are the flows. Four phases exist, whose boundaries are delineated by black dashed lines. In the mean-field calculation (a) all phase boundaries are analytical. In the RG calculation, the straight line $g + g' = 1$ is obtained analytically whereas the other phase boundaries can only be found by numerical integration of the flow equations (14). The flows going to infinity are the artifacts of the perturbative RG; they probably correspond to finite values which are beyond the scope of our analytical approach. Notice how the flows in the $g + g' > 1$ phase all terminate at different values.

the bare value of g is non-negative, than we are on the positive $g' = 0$ semiaxis in the phase diagram—here we see only two phases, insulator (no vortices, $g \rightarrow \text{const.}$) and perfect conductor ($g \rightarrow 0$). However, for g' fixed to zero (that is, with a single flavor only), the perfect conductor reduces to the usual conductor phase of the single-component XY model—in other words, we reproduce the expected behavior.

Physically, it is preferable to give the phase diagram in terms of the quantities Γ, τ, I, I_x, L that appear in the initial equations of motion (1) and (2): The light intensities can be directly measured and controlled, whereas the relaxation time and the coupling cannot, but at least they have a clear physical interpretation. The relations between these and the effective Hamiltonian quantities y, g, g' are found upon integrating out the intensity fluctuations to obtain (7) and the explicit relations are stated in (D11). Making use of these we can easily plot the phase diagram in terms of the physical quantities for comparison with experiment. However, for the qualitative understanding we want to develop here, it is much more convenient to use g, g' as the phase structure is much simpler.

As an example, we plot the Γ - g' diagram in Fig. 3 (we have kept g' to keep the picture more informative; the Γ - L and Γ - I diagrams contain multiple disconnected regions for each phase). The noninteracting fixed point $g = g' = 0$ is now mapped to $\Gamma = 0$. The tricritical point where the PC, the FI, and the conductor phases meet is at $R = 1$. Therefore, the rule of thumb is that low couplings Γ produce stable vortices with conserved charges—the perfect vortex conductor. Increasing the coupling pumps the instability up, and the kind of instability (and the resulting phase) is determined by the relative strength of the photonic lattice compared to the propagating beams. Obviously, such considerations are only a rule of thumb and detailed structure of the diagram is more complex. This is one of the main motives of this study—blind numerical search for patterns without the theoretical approach

adopted here would require many runs of the numerics for a good understanding of different phases.

C. Geometry of patterns

Now we discuss what the intensity pattern $I(\mathbf{r})$ looks like in various phases, for various boundary conditions. This is very important as this is the only thing which can be easily measured in experiment—phases θ_α are not directly observable, while the intensity distribution is the direct outcome of the imaging of the crystal [31]. We shall consider three situations. The first is a single Gaussian beam on zero

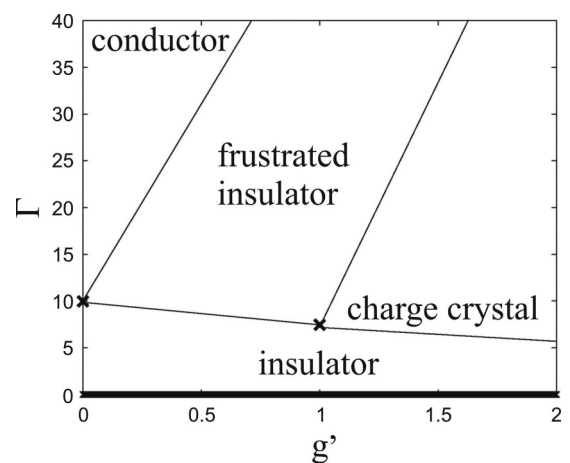


FIG. 3. Typical phase diagram for the system without disorder, in the Γ - g' plane. There are two discrete fixed points and the critical line at $\Gamma = 0$, which corresponds to the critical line $g + g' = 1$ in the previous figure. We also see two discrete fixed points, corresponding to $g_{*,**}, g'_{*,**}$. The advantage of physical parameters is that the location of these fixed points in the Γ - I plane can be calculated directly from the numerics (or measured from the experiment).

background ($I_x = 0$), with Gaussian initial intensity profile $|F(z = 0, \mathbf{r})|^2 = |B(z = L, \mathbf{r})|^2 = \mathcal{N} \exp(-r^2/2s^2)$ and possibly nonzero vortex charges: $\arg \Psi_{\pm}(\mathbf{r}) \sim \exp(Q_{F,B}\phi)$, with $\mathbf{r} = (r \cos \phi, r \sin \phi)$. The second case is a quadratic vortex lattice of F and B beams, so the initial beam intensity is $I_0 = \sum_{i,j} \exp[-(x - x_i)^2/2s_0^2 - (y - y_j)^2/2s_0^2]$, with $x_{i+1} - x_i = y_{i+1} - y_i \equiv b = \text{const.}$, the situation particularly relevant for analogies with condensed matter systems. In the third case, we have again a quadratic vortex lattice but now on top of the background photonic square lattice, which is either coincident or off phase (shifted for half a lattice spacing) with the beam lattice. The background intensity is thus of the form $I_x = \sum_{i,j} \exp[-(x - x_i)^2/2s^2 - (y - y_j)^2/2s^2]$.

First of all, it is important to notice that there are two kinds of instabilities that can arise in a vortex beam:⁴

(1) There is an instability which originates in the imbalance between the diffusion and self-focusing (crystal response) in favor of diffusion in *high-gradient regions*: If a pattern $I(x, y)$ has a large gradient ∇I , the kinetic term in the Lagrangian (4), i.e., the diffusion term in (1) is large and the crystal charge response is not fast enough to balance it as we travel along the z axis, so the intensity rapidly dissipates and the pattern changes. Obviously, the vortex core is a high-gradient region so we expect it to be vulnerable to this kind of instability. This is indeed the case: In the center of the vortex the intensity diminishes, a dark region forms, and the intensity moves toward the edges. We dub this the core or central instability (CI), and in the effective theory it can be understood as the decay of states with low fugacity y , i.e., high self-interaction constants g_0, g_1 . This instability prevents the formation of vortices in the insulator phase, or limits it in the frustrated insulator and conductor phases.

(2) There is an instability stemming from the dominance of diffusion over self-focusing in *low-intensity regions of sufficient size and/or convenient geometry*. At low intensity, the charge response is nearly proportional to I [from Eq. (2)], so if I is small diffusion wins and the intensity dissipates. If there is sufficient inflow of intensity from more strongly illuminated regions, it may eventually balance the diffusion, but if the pattern has a long “boundary”, i.e., outer region of low intensity, it will not happen and the pattern will dissipate out or reshape itself to reduce the low-intensity region. We call this case the edge instability (EI). For a vortex, it happens when the positive and negative vortex charges tend to redistribute due to Coulomb attraction and repulsion. In our field theory Hamiltonian (7), this instability dominates in the conductor and perfect conductor phases.

Let us first show how the CI and EI work for a single beam with nonzero vortex charge. In Fig. 4, we show the intensity patterns for a single vortex with charges (1,0) and (3,0) as the x - y cross sections (transverse profiles) in the middle of the

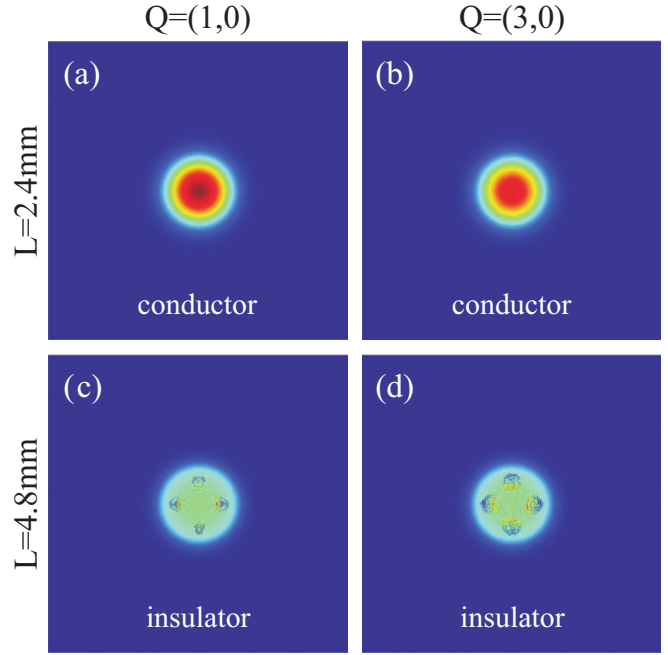


FIG. 4. Transverse profiles for a single Gaussian beam for two different propagation distances, $L = 2.4$ mm (top) and $L = 4.8$ mm (bottom), with vortex charges (1,0) [(a), (c)] and (3,0) [(b), (d)], at the back face of the crystal ($z = L$). The regime on top [(a), (b)] corresponds to the conductor phase, which has a single conserved vortex charge Q_F . This vortex charge conservation prevents significant instabilities; nevertheless, the multi-quantum vortex (3,0) shows the onset of CI; notice the reduced intensity and incoherent distribution of the beam in the central region in the top right panel (the CI is expected to grow roughly as $Q_+^2 + Q_-^2$). The insulator phase only preserves the $F - B$ invariance but not the vortex charge, and in the absence of topological protection the vortices can annihilate into the vacuum. Here we see the EI taking over for both charges; four unstable regions appear near the boundary, violating the circular symmetry and dissipating away the intensity of the vortex. Parameter values: FWHM $40 \mu\text{m}$, $\Gamma I_0 = 41$, $t = 10\tau$.

crystal, i.e., for $z = L/2$. The parameters chosen (Γ, I_0, R, L) correspond to the conductor phase (top) and the insulator phase (bottom). In top panels, for $Q_+^2 + Q_-^2 = 1$, the core energy is not so large and CI is almost invisible. For $Q_+^2 + Q_-^2 = 9$, we see the incoherence and the dissipation in the core region, signifying the CI. The conductor phase allows the proliferation of vortices but only those with $|Q_{\pm}| \leq 1$ are stable. In the bottom panels, both vortices have almost dissipated away due to EI, which starts from discrete poles near the boundary.⁵ Indeed, the insulator phase has no free vortices, no matter what the charge. In Fig. 5, we see no instability even for a high-charge vortex in the perfect conductor phase (top), whereas the frustrated insulator phase (bottom) shows strong EI for the like-charged vortex (3,3) since this fixed point has $g'_* > 0$, but the (3, -3) vortex is stable. Notice that we could not expect

⁴They are distinct from the bifurcations which happen also in topologically trivial beam patterns and lead to the instability which eventually destroys optical (nontopological) solitons. These instabilities have been analyzed in Appendix C and in more detail in Ref. [32], where the authors have found them to start from the edge of the beam and result in the classical “walk through the dictionary of patterns.”

⁵As a rule, it follows the sequence (C9) found in Appendix C from the pole structure of the propagator, though some of the steps can be absent, e.g., for a single Gaussian vortex there is no C_2 stage.

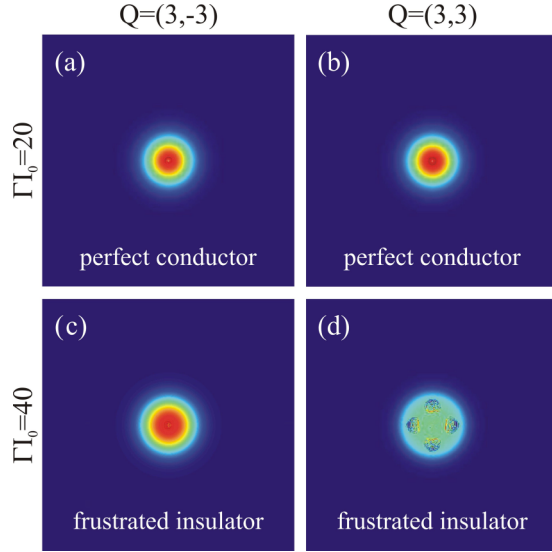


FIG. 5. Transverse profiles for a single Gaussian beam for two different coupling strengths, $\Gamma I_0 = 20$ (top) and $\Gamma I_0 = 40$ (bottom), with vortex charges $(3, -3)$ [(a), (c)] and $(3, 3)$ [(b), (d)] at the back face of the crystal ($z = L$). The regime on top corresponds to the perfect conductor phase, where the vortices of all charges freely proliferate—both vortices are reasonably stable. The bottom case is in the frustrated insulator phase—the forward-backward coupling makes the $(3, 3)$ vortex unstable from EI while the $(3, -3)$ vortex survives. Parameter values: FWHM $40 \mu\text{m}$, $L = 2 \text{ mm}$, $t = 10\tau$.

CI for this case since the sum $Q_+^2 + Q_-^2 = 9$ is the same in both cases—if for $Q_- = -Q_+$ the vortex has no CI, then for $Q_- = Q_+$ it cannot have it either (since the value $Q_+^2 + Q_-^2$ is the same).

We have thus seen what patterns to expect from CI and EI and also what kind of stable vortices to expect in different phases: The perfect conductor phase allows free proliferation of vortices of any charge, the conductor phase allows only single-quantum vortices (or vortices with sufficiently low $Q_+^2 + Q_-^2$) while others dissipate from CI, the frustrated insulator supports the vortices with favorable charges (or favorable charge distribution in multiple-vortex systems) while others disintegrate from EI, and the insulator phase supports no vortices—they all dissipate from CI or EI, whichever settles first (depending on the vortex charges).

The case rich with analogies with condensed matter systems is the square vortex lattice on the background photonic square lattice, Fig. 6. Here we can also appreciate the transport processes. The photonic lattice is coincident with the beam lattice and equal in intensity, so $\Gamma(I_0 + I_x) = 2\Gamma I_0$. In the perfect conductor phase [Fig. 6(a)], the vortices are stable and coherent and keep the uniform lattice structure. In the conductor phase [Fig. 6(b)], the CI is visible but the lattice structure survives. The bottom panels show the nonconducting phases: frustrated insulator [Fig. 6(c)] and insulator [Fig. 6(d)]. The insulator loses both lattice periodicity and the Gaussian profile of the vortices but the frustrated insulator keeps the regular structure: From EI the intensity is *inverted* and the resulting lattice is *dual* to the original one [compare Fig. 6(c) to Fig. 6(a)]. The phase patterns $\theta_F(x, y; z = L/2)$

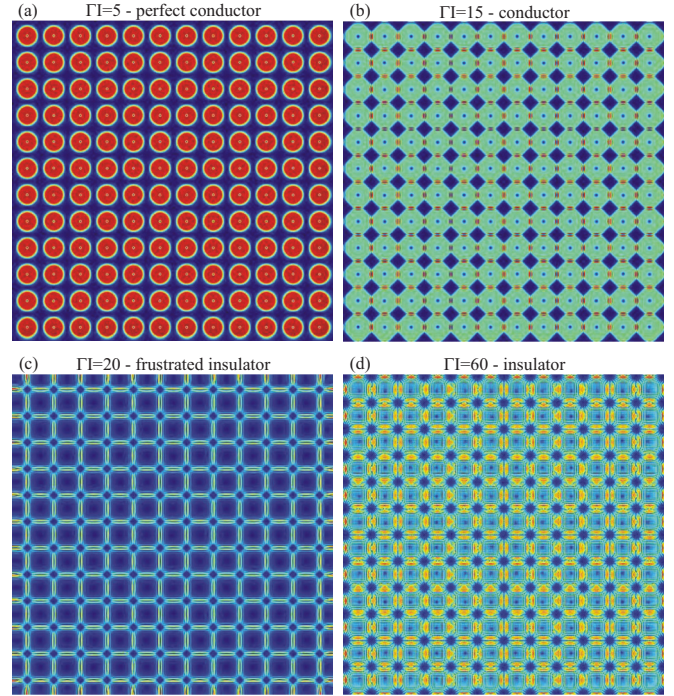


FIG. 6. Vortex lattice with Gaussian profile for $\Gamma I = 5$ [PC, panel (a)], $\Gamma I = 15$ [conductor, panel (b)], $\Gamma I = 20$ [FI, panel (c)], and $\Gamma I = 60$ [insulator, panel (d)]. The perfect conductor phase has a coherent vortex lattice and no instabilities. Conductor exhibits a deformation of the vortex lattice and the reduction of the full $O(2)$ symmetry, starting from the *center*, whereas the FI exhibits the reduction of symmetry and the inversion of the lattice due to *edge* effects. Notice how both phases have reduced symmetry compared to PC but retain coherence. Only the insulator phase loses not only symmetry but also coherence; i.e., the intensity diffuses and the pattern is smeared out. Transverse size of the lattice is 512×512 in computational space; same lattice size, FWHM, and lattice spacing are used for all subsequent figures unless specified otherwise. Parameter values: $L = 4.8 \text{ mm}$, $t = 10\tau$, FWHM $10 \mu\text{m}$, and lattice spacing equal to FWHM.

and $\theta_F(x, z; y = 320 \mu\text{m})$ for the perfect conductor (top) and the frustrated insulator phase (bottom) are shown in Fig. 7. Here we see the vortex charge transport mechanism in a PC: The vortices are connected in the sense that the phase θ_F is coherently traveling from one vortex to the next. In the FI phase, the phase is initially frozen along the z axis, until the transport starts at some $z \approx L/2$.

It may be instructive to take a closer look at the lattice dynamics of the most interesting phase: the frustrated insulator. In Fig. 8, we inspect square lattices on the photonic lattice background for several charges of the form $(Q_+ = 3, Q_-)$. The first row shows how the vortices lose stability and develop CI as the total square of the charge grows [from Fig. 8(a) to Fig. 8(c)]. Figures 8(d)–8(i) show how the g' coupling favors the opposite sign of Q_+ and Q_- and how the optimal configuration is found for $Q_- = -3$. This is easily seen by minimizing the free energy over Q_- : It leads to the conclusion that the forward-backward coupling favors the “antiferromagnetic” ordering in the sense that $Q_+ + Q_- = 0$.

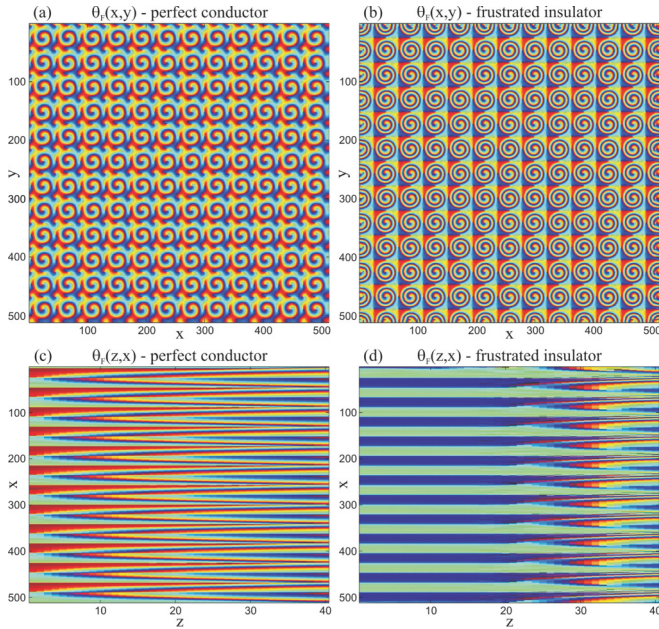


FIG. 7. Same system as in panels (a) and (c) from the previous figure (PC and FI phases) but now we plot the phase θ_F , as the transverse cross section $\theta_F(x,y; z = L/2)$ [(a), (b)] and as the longitudinal section along the PR crystal $\theta_F(x,z; y = 320 \mu\text{m})$ [(c), (d)]. The perfect conductor phase has well-defined vortices in contact which allows the transport of the vortex charge through the lattice and shows as the periodical modulation of the phase along the z axis (vortex lines). The frustrated insulator keeps well-defined vorticity even though the intensity map undergoes inversion [Fig. 6(c)] with frozen phase along the z axis, so there is no vorticity transport until some $z \approx L/2 = 2.4 \text{ mm}$, when the phase stripes develop into vortex lines. The unit on the x and y axis is $1 \mu\text{m}$ (1 in computational space) and on the z axis 0.12 mm (120 in computational space).

Finally, it is interesting to see how the FI phase at high intensities and coupling strengths contains a seed of translation symmetry breaking which will become important in the presence of disorder. In Figs. 9 and 10, we give intensity and phase transverse profiles across the PC-FI transition and deep into the FI phase at large couplings. The intensity maps show the familiar inverse square lattice but the phase maps show stripelike ordering, i.e., translation symmetry breaking along one direction in Figs. 10(c) and 10(d)—horizontal and vertical lines with a repeating constant value of the phase θ_F on all lattice cells along the line. This is a new instability, distinct from CI and EI. We cannot easily derive this instability from the perturbation theory in Appendix C as it is a collective phenomenon and cannot be understood from a single beam.

IV. THE SYSTEM WITH DISORDER

Consider now the same system in the presence of quenched disorder. This is a physically realistic situation: The disorder corresponds to the holes in the photonic lattice which are caused by the defects in the material. The defects are in fixed positions, i.e., they are quenched, whereas the beam is dynamical and can fluctuate. Now $I_x(\mathbf{r}) \rightarrow I_x(\mathbf{r}) + I_h(\mathbf{r})$; i.e.,

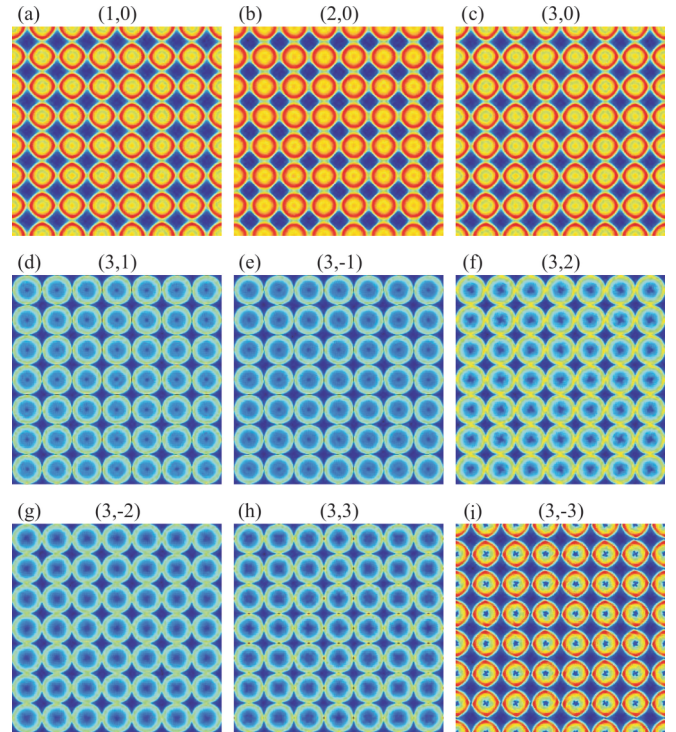


FIG. 8. Transverse profiles for vortex lattices with different charges in the FI phase. In the first row [(a)–(c)], we see how the CI gets stronger and stronger as the total vortex core energy grow (with the square of the total charge). The second and third rows show the growth of CI from $(3,0)$ to $(3, \pm 3)$ (notice the increasingly reduced intensity in the center and the strong ringlike structure of the beams) but also the forward-backward interaction which favors the configurations $(3, -3), (3, -2), (3, -1)$ over $(3, 3), (3, 2), (3, 1)$. In particular, the $(3, -3)$ lattice is the optimal configuration of all $(3, Q_-)$ configurations even though it has greater CI than say $(3, 0)$ (notice the small dark regions in the center), because the $\sum_{ij} gg' Q_{i+} Q_{i-} \ln r_{ij}$ term minimizes the EI—notice there is no “spilling” of intensity from one vortex to the next. The parameters are $\Gamma I = 20, L = 2.5 \text{ mm}$.

the quenched random part $I_h(\mathbf{r})$ is superimposed to the regular background (whose intensity is I_x). The disorder is given by some probability distribution, assuming no correlations between defects at different places. As in the disorder-free case, the lattice is static and “hard”, i.e., does not backreact due to the presence of the beams. One should, however, bear in mind that the backreaction on the background lattice can sometimes be important as disregarding it violates the conservation of the angular momentum [37]. Disregarding the backreaction becomes exact when $I_x + I_h \gg |\Psi|^2$, i.e., when the background irradiation is much stronger than the propagating beams.

To treat the disorder, we use the well-known replica formalism [54]. For vortex-free configurations, typical experimental values of the parameters suggest that the influence of disorder is small [31, 33, 35]. However, the influence of disorder becomes dramatic when vortices are present. This is expected, since holes in the lattice can change the topology of the phase field θ_{\pm} (the phase now must wind around the holes). Our equations of motion are still given by the Lagrangian (4), but with $I_x \mapsto I_x + I_h$. In our analytical calculations, we assume

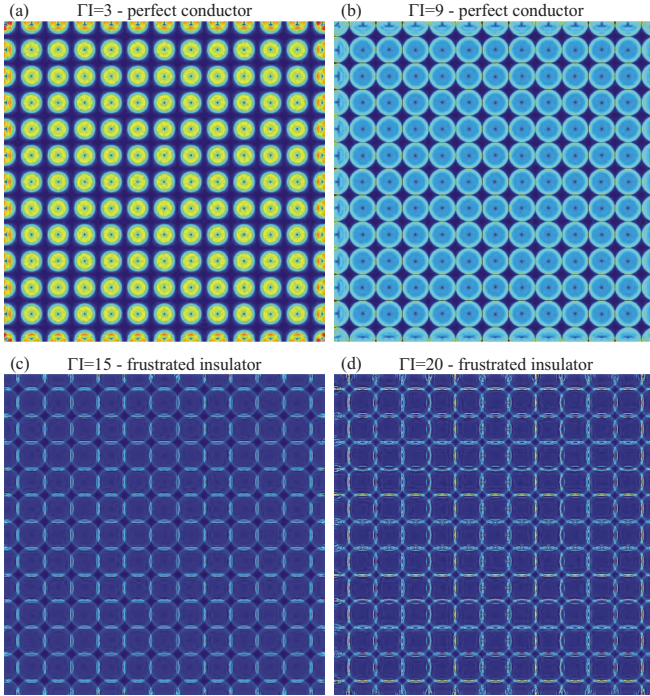


FIG. 9. Intensity maps for the quadratic vortex lattice with charges (1,1), for increasing values of $\Gamma I = \Gamma(I_0 + I_x)$. The transition from the PC phase [(a), (b)] into the FI phase [(c), (d)] happens at about $\Gamma I \approx 12$. The edge instability sets in progressively, in accordance with what we saw in the previous figure, leading eventually to an inverse square lattice. Propagation length $L = 5$ mm.

that a defect in the photonic lattice changes the lattice intensity from I_x to $I_x + I_h$, with Gaussian distribution of “holes” in I_h , which translates to the approximately Gaussian distribution of the couplings g, g', g_0, g_1 . In the numerics, however, we do a further simplification and model the defects in a discrete way; i.e., at a given spot either there is a lattice cell of intensity I_1 (with probability h), or there is not (the intensity is zero, with probability $1 - h$). This corresponds to $I_x = I_1/2, I_h = \pm I_1/2$ so the disorder is discrete. Due to the central limit theorem, we expect that the Gaussian analytics should be applicable to our numerics.

A. The replica formalism at the mean-field level

To study the system with quenched disorder in the photonic lattice, we need to perform the replica calculation of the free energy of the vortex Hamiltonian (7). We refer the reader to the literature [41,42] for an in-depth explanation of the replica trick. In short, one needs to average over the various realizations of the disorder *prior* to calculating the partition function, i.e., prior to averaging over the dynamical degrees of freedom (vortices in our case). This means that we need to perform the disorder average of the free energy, i.e., the logarithm of the original partition function $-\ln \mathcal{Z}$, and not the partition function \mathcal{Z} itself. The final twist is the identity $\ln \mathcal{Z} = \lim_{n \rightarrow 0} (\mathcal{Z}^n - 1)/n$: We study the Hamiltonian consisting of n copies (replicas) of the original system and then *carefully*

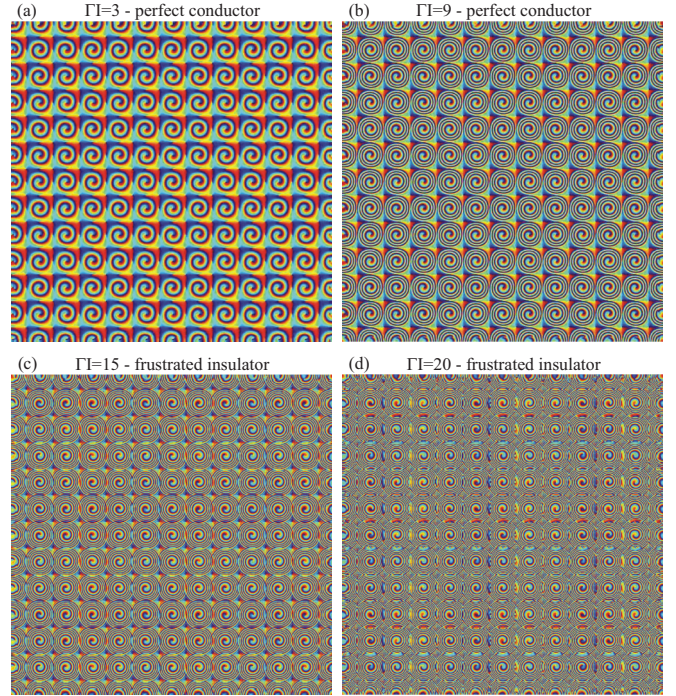


FIG. 10. Transverse phase maps for the F beam for the same cases as in Fig. 9. As the coupling strength ΓI grows toward very large values (d), the violation of translation symmetry becomes obvious: Notice the vertical and horizontal phase stripes. This instability gives rise to the charge density wave ordering in the presence of disorder.

take the $n \rightarrow 0$ limit.⁶ The partition function of the replicated Hamiltonian reads

$$\mathcal{Z} = \lim_{n \rightarrow 0} \text{Tr} \exp \left[- \sum_{\mu=1}^n \mathcal{H}_{\text{vort}}(Q^{(\mu)}) \right], \quad (15)$$

where $Q^{(\mu)}$ are the vortex charges in the μ th replica of the system. In the original Hamiltonian (7), the disorder turns the interaction constants into quenched random quantities $g_{ij}, g'_{ij}, g_{0;ij}, g_{1;ij}$, so we can compactly write our interaction term as

$$\mathcal{H}_{\text{vort}} = \sum_{ij} \sum_{\alpha\beta} Q_{i\alpha} J_{ij}^{\alpha\beta} Q_{j\beta} \quad (16)$$

with $J_{ij}^{++} = J_{ij}^{--} = g_{ij}(1 - \delta_{ij}) \ln r_{ij} + g_0 \delta_{ij}$, $J_{ij}^{+-} = J_{ij}^{-+} = g'_{ij}(1 - \delta_{ij}) \ln r_{ij} + g_1 \delta_{ij}$. Now we again make the mean-field approximation for the long-ranged logarithmic interaction. Similar to the clean case, for $i \neq j$ we approximate $g \ln r_{ij} \sim g' \ln r_{ij} \sim \ln \Lambda$, knowing that $g, g' \sim 1$ and assuming that average intervortex distance is of the same order of magnitude as the system size Λ , and for the core energy we likewise get $g_0, g_1 \sim \ln a/\epsilon \sim -\ln \epsilon \sim \ln \Lambda$. The result is that all terms in $J_{ij}^{\alpha\beta}$, both for $i \neq j$ and $i = j$, are on average of the order $\ln \Lambda \gg 1$, and the mean-field approach is justified. We will sometimes denote the 2×2 matrices in the flavor space by hats (e.g., $\hat{J} = J^{\alpha\beta}$).

⁶Care is needed as the $n \rightarrow 0$ limit does not in general commute with the thermodynamic limit.

The final Hamiltonian (16) has the form of the random-coupling and random-field Ising-like model: Random couplings stem from the stochasticity of J_{ij} values and random field from the fact that $\langle J_{ij} \rangle \neq 0$ introduces terms linear in $Q_{i\alpha}$, i.e., an effective external field coupling to the “spins.” We have arrived at this model through three steps of simplification: our microscopic model is a type of the XY -glass model (Cardy-Ostlund model [55]), a well-known toy model for disorder. At this stage, our model is similar to the work of Refs. [9,10], only with two components instead of one. Then we have written the effective vortex Hamiltonian with Coulomb-like interaction, disregarding the topologically trivial configurations. This is a rather extreme approximation but a necessary one as it is very complicated to consider the full model with vortices. Finally, we have approximated the logarithmic potential with a constant all-to-all vortex coupling. Such an approximation

(essentially the infinite dimension limit) is frequently taken and lies at the heart of the solvable Sherington-Kirkpatrick Ising random coupling model [41]. Our case differs from the Sherington-Kirkpatrick model as it (i) has also a random field, (ii) has two flavors, and (iii) has the Ising spins taking arbitrary integer values. From the random XY model it differs by (i) and (ii) above, and also by considering only vortices and no nontopological spin configurations. The additional phases we get in comparison to Refs. [9,10] and its generalization in Refs. [11,12,46] come from the interactions between the forward and backward flavors. But bearing in mind the drastic approximations we take, we stress that we cannot aspire to solve either the XY model or the resulting Ising-like model in any rigorous way (certainly not at the level of rigor of mathematical physics). We merely try to obtain a crude understanding.

The Gaussian distribution of defects reads $p(J_{ij}^{\alpha\beta}) = \exp[-(J_{ij}^{\alpha\beta} - J_0^{\alpha\beta})(\hat{\sigma}^{-2})_{\alpha\beta}(J_{ij}^{\alpha\beta} - J_0^{\alpha\beta})]$, where the second moments are contained in the matrix $\sigma_{\alpha\beta}$, with $\sigma_{+-} = \sigma_{-+}$. In this case, we get the replicated partition function

$$\bar{Z}^n = \int \mathcal{D}[Q_{i\alpha}^{(\mu)}] \int \mathcal{D}[J_{ij}^{\alpha\beta}] \exp \left[-\frac{1}{2} \sum_{i,j=1}^N \sum_{\alpha,\beta} (J_{ij}^{\alpha\beta} - J_0^{\alpha\beta}) \sigma_{\alpha\beta}^{-2} (J_{ij}^{\alpha\beta} - J_0^{\alpha\beta}) - \sum_{\mu=1}^n \sum_{i,j=1}^N \sum_{\alpha,\beta} \beta J_{ij}^{\alpha\beta} Q_{i\alpha}^{(\mu)} Q_{j\beta}^{(\mu)} \right]. \quad (17)$$

We can now integrate out the couplings $J_{ij}^{\alpha\beta}$ in (17) and get

$$\bar{Z}^n = \text{const.} \int \mathcal{D}[Q_{i\alpha}^{(\mu)}] \exp \left[\frac{1}{2} \beta^2 \sum_{\mu,\nu=1}^n \sum_{i,j=1}^N \sum_{\alpha,\beta} Q_{i\alpha}^{(\mu)} Q_{j\alpha}^{(\nu)} (\hat{\sigma}^2)_{\alpha\beta} Q_{j\alpha}^{(\mu)} Q_{j\beta}^{(\nu)} - \beta \sum_{\mu=1}^n \sum_{i,j=1}^N \sum_{\alpha,\beta} J_0^{\alpha\beta} Q_{i\alpha}^{(\mu)} Q_{j\beta}^{(\mu)} \right]. \quad (18)$$

Integrating out the disorder has generated the nonlocal quartic term proportional to the elements of $\sigma_{\alpha\beta}^2$. The additional scale given by the average disorder concentration means we cannot scale out $\beta = L$ anymore, and it becomes an additional independent parameter. The partition function can be rewritten in the following way, usual in the spin-glass literature [42,54]. We can introduce the nonlocal order parameter fields

$$p_\alpha^{(\mu)} = \frac{1}{N} \sum_{i=1}^N Q_{i\alpha}^{(\mu)}, \quad q_{\alpha\beta}^{(\mu\nu)} = \frac{1}{N} \sum_{i,j=1}^N Q_{i\alpha}^{(\mu)} Q_{j\beta}^{(\nu)}, \quad (19)$$

which have the meaning of overlap between different metastable states. The rest is just algebra, although rather tedious: One rewrites the Hamiltonian in terms of new order parameters, and then one can solve the saddle-point equations for p_α and $q_{\alpha\beta}$, or do an RG analysis. The calculation is found in Appendix G.

The mean-field analysis yields six phases:

(1) One phase violates both the replica symmetry and the flavor symmetry, breaking it down to identity. We dub this phase *vortex charge density wave* (CDW), as it implies spatial modulation of the vortex charge, leading to nonzero net charge density $\sum_i Q_{i\alpha}^{(\mu)}$ in some parts of the system even if the boundary conditions are electrically neutral (the *total* net charge density must still be zero due to charge conservation). Vortices take their charges from $\mathbb{Z} \otimes \mathbb{Z}$.

(2) The second phase violates the replica symmetry in both flavors and reduces the flavor symmetry but does not break it down to identity. Instead, it reduces it to the diagonal subgroup

$U(1)_F \otimes U(1)_B \rightarrow U(1)_d$, so it has nonzero density of the vortex charge in a given replica $\sum_i Q_{i+}^{(\mu)} = -\sum_i Q_{i-}^{(\mu)}$. Again, the charge density is locally nonzero but now with an additional constraint resulting in frustration (multiple equivalent free energy minima). This is thus the dirty equivalent of the frustrated insulator phase and we dub it *vortex glass*, as it has long-range correlations (because of the logarithmic interactions between charged areas), does not break spatial symmetry, and exhibits frustration; its charges are from $\pi_1[U(1)_d] = \mathbb{Z}$.

(3) The remaining phases have no nonzero vortex charge density fluctuation and are similar to the phases in the clean system. Vortex perfect conductor violates the replica symmetry of all three fields q^{++}, q^{--}, q^{+-} and allows free proliferation of vortices with charges $(Q_+, Q_-) \in \mathbb{Z} \otimes \mathbb{Z}$.

(4) Frustrated vortex insulator preserves the replica symmetry of $q^{\pm\pm}$ but has nonzero value, with broken replica symmetry, of the mixed q^{+-} field, which gives $U(1)_d$ vortices, with charges $Q_+ = -Q_- \in \mathbb{Z}$.

(5) Vortex conductor preserves the replica symmetry of the mixed q^{+-} order parameter but violates it in $q^{\pm\pm}$, resulting in the proliferation of single-flavor vortices with \mathbb{Z} charge.

(6) Vortex insulator fully preserves the replica symmetry, all order parameters are zero, and vortices cannot proliferate. RG analysis will show that insulator survives only at zero disorder; otherwise it generically becomes CDW.

The phase diagram (given in Fig. 11 in the next subsection) now contains six phases (only five are visible for the parameters chosen in the figure): CDW, insulator, FI,

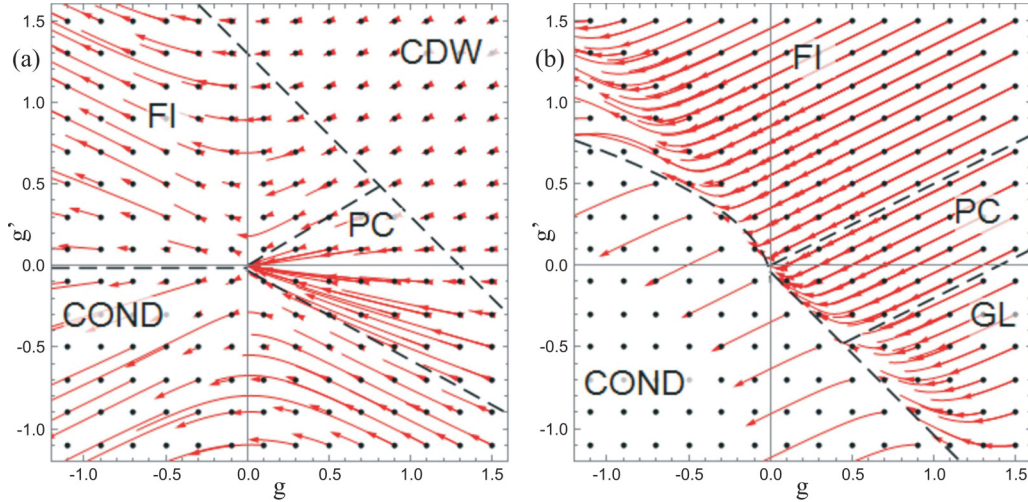


FIG. 11. Phase diagram for the system with lattice disorder in the g - g' plane together with RG flows, with red lines denoting the flows starting at the initial conditions denoted by black points. The dashed black lines are approximate phase boundaries from mean-field theory, for $\sigma^2 = 0.4$ (a) and $\sigma^2 = 1.2$ (b). In panel (a), the area where $g + g' + \beta^2\sigma^2 > 1$ is inhabited by the flows toward nonuniversal values of (g, g') which belong to the CDW phase and the opposite region is divided between the attraction regions of $(0, 0)$, $(g_* \rightarrow \infty, g'_* \rightarrow \infty)$, and $(g_{**} \rightarrow \infty, g'_{**} \rightarrow -\infty)$ —the familiar PC, FI, and conductor phases. In panel (b), for $\sigma^2 = 1.2$, the disorder becomes relevant in the glass phase (denoted by “GL”), whose RG flows end on the half-line of fixed points $g + g' + \beta^2\sigma^2 = 1, g' < 0$. For our parameter values, this line happens to pass almost through the origin; in general, this is not necessarily the case. The nondisordered phases (flowing to $\sigma^2 = 0$) FI, conductor, and PC have survived. Propagating length is $L = 3.0$ nm.

conductor, PC, and the glassy phase. The insulator phase is now of measure zero in the (g, g', σ^2) plane, existing only for the points at $\sigma^2 = 0$; for generic nonzero values we have a CDW. For simplicity, we have plotted the phase diagram for $\sigma_{++}^2 = \sigma_{--}^2 = \sigma_{+-}^2 \equiv \sigma^2$.

B. RG analysis and the phase diagram

To study the RG flow, we can start from the replicated partition function (18), inserting the definition of the couplings $J_{ij}^{\alpha\beta}$ and keeping the vortex charges $Q_{i\alpha}^{(\mu)}$ as the degrees of freedom (without introducing the quantities $p_\alpha^\mu, q_{\alpha\beta}^{(\mu\nu)}$). The basic idea is the same: We consider the fluctuation $\delta(\bar{\mathcal{Z}}^n)$ upon the creation of a vortex pair at $\mathbf{r}_{1,2}$ with charges $\vec{q}_1^{(\mu)}, -\vec{q}_2^{(\mu)}$, in the background of the vortices $\vec{Q}_{1,2}^{(\nu)}$ at positions $\mathbf{R}_{1,2}$. Likewise, we introduce the fugacity parameter $y^{(\mu)}$ to account for the vortex core energy. However, this problem is much harder than the clean problem and one has to resort to many approximations to perform the calculation. In its most general form, the problem is still open, in the sense that all known solutions suppose a certain form of replica symmetry breaking or truncate the RG equations [42]. The RG analysis is thus less useful in the disordered case but at least the numerical integration of the flow equations is supposed to give a more precise rendering of the phase diagram compared to the mean field theory. We again describe the calculation in Appendix G and jump to the results.

The fixed point of the flow equations lies either at infinite y or at $y = 0$ like in the clean case. This is again controlled by the equation for $\partial y / \partial \ell$ but now depending on the combination $g + g' + \beta^2\sigma^2$ instead of $g + g'$ in the clean case (for simplicity, we consider the case where $\sigma_{\alpha\beta}^2$ are all equal). The following cases appear:

(1) When the fugacity flows toward infinity, we reproduce the phases and the fixed point values (g, g', σ^2) from the clean case: The PC flows toward $(0, 0, 0)$, the FI toward $(g_*, g'_*, 0)$, and the conductor toward $(g_*, g'_{**}, 0)$ with $g_* \rightarrow -\infty, g'_* \rightarrow -\infty, g'_{**} \rightarrow \infty$. Notice that all these phases flow to $\sigma^2 = 0$; i.e., disorder is irrelevant.

(2) When the fixed point lies at $y = 0$, one possibility is that all parameters (g, g', σ^2) flow toward some nonuniversal nonzero values. The attraction region of this point is the CDW phase: The disorder term stays finite as well as the couplings. In particular, the points on the half-plane $g + g' > 0, \sigma^2 = 0$ stay at $\sigma^2 = 0$ (with constant coupling values) and this is the insulator phase from the clean case. Notice that $\sigma^2 > 0$ now; i.e., disorder is relevant. For $\sigma^2 < 1$, this are the only fixed points when $y = 0$.

(3) However, for sufficiently strong disorder ($\sigma^2 > 1$), there is a new line of fixed points at $y = 0$ with a finite attraction region, corresponding to a new phase. For $\beta > 1$, the right-hand side of the second RG equation in (G19) has a zero at nonzero g' and there are trajectories flowing toward $(y, g, g', \sigma^2) = [0, g, g'(g), \sigma^2(g)]$ and not toward an arbitrary nonuniversal value of σ^2 . This is precisely the glass phase, where disorder is again relevant. At the lowest order, the relation between g, g', σ^2 at the fixed point line is given by the relation $g + g' + \beta^2\sigma^2 = 1$.

Now we have made contact between the mean-field classification of phases and the fixed points and regions of the RG flow. The flows in the (g, g') plane are given in Fig. 11. The parameter space is four-dimensional so the phase structure is different at different disorder concentrations σ^2 . In Fig. 11(a) for $\sigma^2 = 0.4$, the phase structure is similar to the clean case; we see the same four phases except that insulator (no stable vortices) is replaced by the CDW phase with localized vortices. In Fig. 11(b) for $\sigma^2 = 1.2$, the CDW phase is replaced by

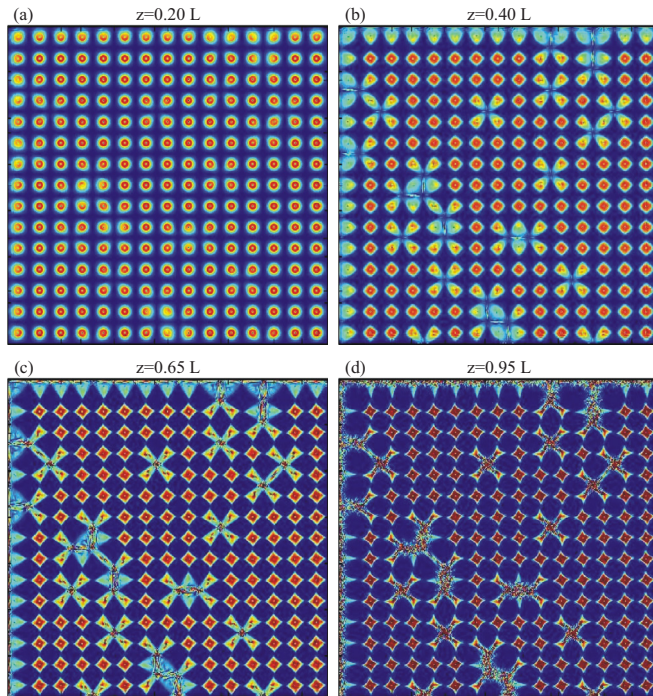


FIG. 12. Transverse profile for the PC phase in a Gaussian beam lattice on a background lattice, for four different propagation distances. The vortex charge is $(1, 1)$, which is sufficiently low that the CI does not destroy the vortices. We see some CI-induced symmetry reduction from $O(2)$ to C_4 but the overall lattice structure is preserved. Parameter values are $\sigma^2 = 0.1, \Gamma I = 20, L = 2$ mm, FWHM for the CP beams is $9 \mu\text{m}$ and for the photonic lattice $6 \mu\text{m}$.

another disordered phase, the glasslike regime. Importantly, the glass phase does not cross the $g' = 0$ axis, meaning that a single-flavor system even with disorder could not support a glass. We thus conjecture that the transition at $\sigma^2 = 1$ is of first order, as the change is the structure of the (g, g') phase diagram is discontinuous, and we do not see how this could happen if the first derivative $\partial\mathcal{F}/\partial\rho_{\pm}$ (the derivative of the free energy with respect to vortex charge density) is continuous. However, we have not checked the order of this transition by explicit calculation. The phase structure is further seen in the $\sigma^2 - g'$ diagram, where we see the glass phase emerge at some value of the disorder. This is discussed further in the next section, where we study the equivalent antiferromagnetic system (with the same structure of the phase diagram, Fig. 16).

C. Geometry of patterns

The two previously considered mechanisms of instability—central instability and edge instability—remain active also in the presence of disorder. However, in the presence of disorder there is a third, inherently collective effect that we dub *domain instability* (DI). It follows from the fact that the self-focusing term ΓE grows with intensity I : More illuminated regions react faster [Eqs. (1) and (2)]. In the presence of background lattice, there will be regions of initially zero beam intensity I_0 where the regular lattice cells have some nonzero intensity I_x . Approximating $I = I_0 + I_x \approx I_x = \text{const.}$, our equations in the vicinity of the defect (hole) in the background lattice

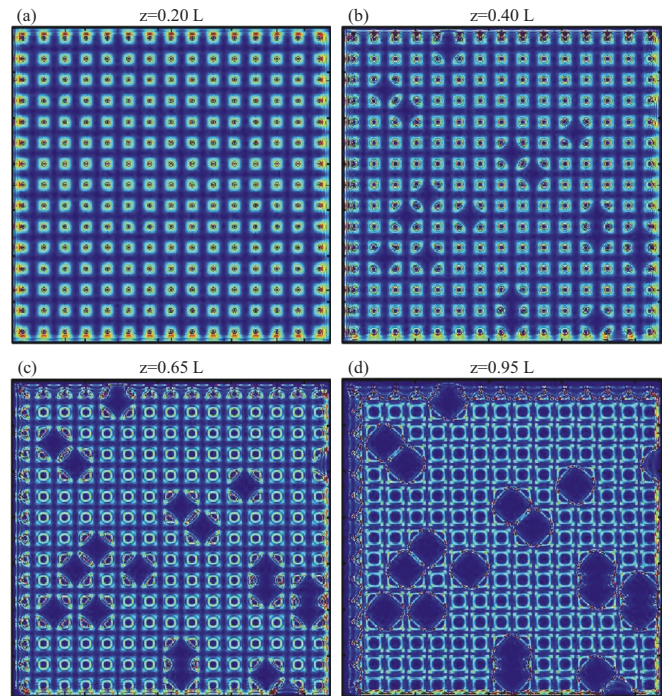


FIG. 13. Transverse profile for the FI phase, present in the same system as in Fig. 12 but for $\Gamma I = 40$. Now both the CI [low-intensity regions in the beam center in panels (a) and (b)] and the EI [lattice inversion in panels (c) and (d)] are present. The net result is the lattice inversion, and the vortex charge dissipates along the inverse lattice.

becomes the Schrödinger equation in a step potential (equal to I_x in the regular parts of the photonic lattice, and equal to zero where a hole is found), so the z -dependent part of the solution is of the form $\sum_k e^{i\lambda_k z}$ and the eigenenergies along z are gapped by the inverse length: $\lambda_k > 1/L$. For small eigenenergies, the transmission coefficient is very low, whereas for large energies it approaches unity. Thus for $1/L$ large (i.e., there are few λ_k 's which are larger than $1/L$), most of the intensity remains confined by the borders of the defect and the intensity does not spill but for small $1/L$ the beam profile is deformed by the “spilling” into the hole regions. For vortices, there is an additional Coulomb interaction in the x - y plane, meaning the effective potential is not piecewise constant anymore (even in the simplest approximation) but the qualitative conclusion remains: Large L brings global reshaping of the intensity profile.

The other phases are analogous to the ones in the clean case, though with a general trend that the presence of disorder decreases the stability of vortex patterns. The PC and FI phases are shown in Figs. 12 and 13. In this section, we only look at the lattices, as the notion of disorder is inapplicable for a single beam. Consider first the patterns in the PC phase (Fig. 12). Compared to the clean case [Fig. 6(a)], the symmetry is much reduced, from $O(2)$ to C_4 , but the vortices are conserved and the original lattice structure (outside the holes) is clearly visible. The FI (Fig. 13) shows mainly EI (and to a smaller extent CI), which together lead to the lattice inversion. The rule of thumb for differentiating the conductor and PC on one side from the CDW and FI on the other side is precisely the presence of the lattice inversion. The absence of the charge transport is

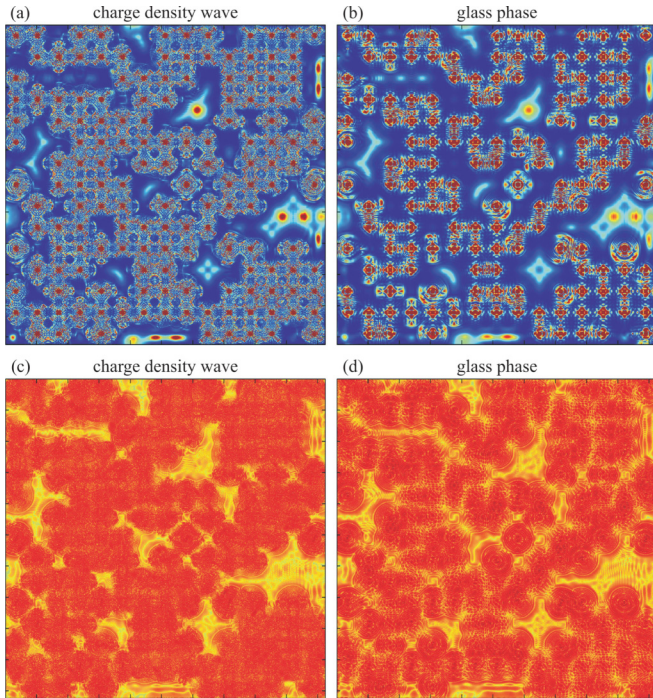


FIG. 14. Transverse profiles for the charge density wave [panels (a) and (c)] and the glass phase [panels (b) and (d)]: intensity maps (top) and vortex charge density maps (down). The telltale difference is that the CDW loses the regular lattice as the intensity “flows” between the regular and the defect regions and we see the DI at work. Glass, on the other hand, consists of domains with coherent (well-defined) vortices though with reduced symmetry (C_4) mostly due to EI. The charge density forms a connected network in the glass phase and transport is possible, whereas in a frustrated insulator the charge is stuck in isolated points.

best appreciated in the phase images: The charge pins to the defects and localizes toward the end of the crystal (i.e., for z near L). Only near the edges we see high vorticity, somewhat analogous to topological insulators, which only have nonzero conductivity along the edges of the system.

The CDW versus the glass phase is given in Fig. 14. The charge density wave [Figs. 14(a) and 14(c), $L = 240 \mu\text{m}$] exhibits the diffusion of intensity due to DI, and the vortex beams are in general asymmetric and not clearly delineated. In Figs. 14(b) and 14(d), where $L = 120 \mu\text{m}$ with all other parameters the same, there is a clear border between defects and the regular parts of the lattice and the intensity is concentrated in the vortex cores. We give also the vortex charge density map in Figs. 14(c) and 14(d) in addition to the intensity maps in Figs. 14(a) and 14(b) as the charge density shows why the CDW is insulating: Even though individual beams diffuse and smear out in *intensity*, the regions of nonzero vortex charge are disjoint and no global conduction can occur. Glass is divided into ordered domains in intensity but the vortex charges form a connected network which supports transport. This is analogous to the percolation transition in a disordered Ising model [56,57] and we may expect that the CDW-glass transition follows the same scaling laws near the critical point. However, we have not checked this explicitly and we leave it for further work.

V. THE CONDENSED MATTER ANALOGY: COLLINEAR DOPED HEISENBERG ANTIFERROMAGNET

The two-beam photorefractive system can serve as a good model for quantum magnetic systems. The most obvious connection is to multicomponent XY antiferromagnets (i.e., two-dimensional Heisenberg model): Planar spins are nothing but complex scalars, and the vortex Hamiltonian remains identical ($\pi_1[\text{SO}(2)] = \pi_1[\text{U}(1)] = \mathbb{Z}$). The nonlinearity in the spin system is different and usually much simpler, but that typically does not influence the phase diagram (the symmetry structure remains the same). Such connection is so obvious it does not require further explanations. Our point is that the CP beams in a PR crystal can also describe more general magnetic systems in the presence of topological solutions described by homotopy groups different from \mathbb{Z} . In particular, we want to point out a connection with a two-sublattice antiferromagnetic system which has some time ago enjoyed considerable popularity as a possible description of magnetic ordering in numerous planar strongly coupled electron systems, including cuprate high- T_c superconductors [5,38,58]. This is the collinear doped antiferromagnet defined on two sublattices. When coupled to a charge density wave (speaking about the usual $\text{U}(1)$ electromagnetic charge) and a superconducting order parameter, it becomes a toy model of cuprate materials (one variant is given in Ref. [58]). In the light of what we know today, the ability of this model to realistically describe the cuprate physics is quite questionable; but even so it is an interesting magnetic system on its own, and it was already found in Refs. [39,47] to exhibit a spin-glass phase, though in a slightly different variant (in particular, with spiral instead of collinear ordering).

Let us formulate the model. While the material is a lattice on the microscopic level, here we are talking about an effective field theory model. The order parameter is the staggered magnetization

$$M(\mathbf{r}) = \sum_{\alpha=1,2} \mathbf{M}_\alpha(\mathbf{r}) \cos(\mathbf{n} \cdot \mathbf{r}), \quad (20)$$

where $\alpha \in \{1,2\}$ is the sublattice “flavor” index (analogous to the α index for the F and B beam in the previous sections)⁷ and each component \mathbf{M}_α is a three-component spin, describing the internal, i.e., spin degree of freedom (we label the spin axes as X, Y, Z). The total spin is thus the sum of the spins of the two components, and \mathbf{n} is the modulation vector. The modulation gives rows of alternating staggered magnetization in opposite directions as in Fig. 15(a). This stands in contrast with the spiral order, where the modulation vectors become \mathbf{n}_α , i.e., differ for the two sublattices, and are themselves space dependent [39]. The ordered phase of the collinear system has the nonzero expectation value of the staggered magnetization along one direction, which can be chosen as the Z axis (“easy axis”), where the spin fluctuations about the easy axis remain massless, and the symmetry is broken from $\text{O}(3)$ to $\text{O}(3)/\text{O}(2)$. The spiral order, on the other hand, breaks the symmetry down to identity, as the order parameter is a dreibein [39].

⁷Sometimes we will denote the sublattices by \pm instead of 1,2 for compactness of notation.

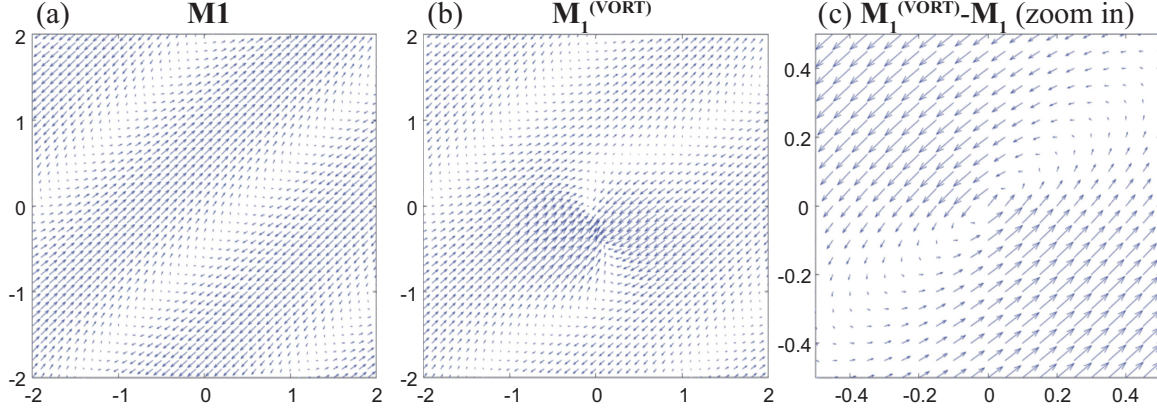


FIG. 15. Numerical realization of the spin pattern (staggered magnetization \mathbf{M}_1) in the collinear $O(3)$ antiferromagnet. Magnetization is three dimensional and we give the projection in the XY plane, $\mathbf{M}_1 \cdot \mathbf{n}_{XY} \equiv M_\perp$. In panel (a), we show the characteristic collinear spin pattern in absence of vortices. In panel (b), we plot $\mathbf{M}_1^{(\text{VORT})}$, a \mathbb{Z}_2 -charged point vortex defect with $Q = 1$. In panel (c), we give an enlargement of the vortex from panel (b) shown as the difference $\mathbf{M}_1^{(\text{VORT})} - \mathbf{M}_1$ to show more clearly the structure of the vortex—now the regular periodic pattern is absent and we appreciate the pointlike structure of the vortex. The parameters are $u = r = 1$ and $v = 0.5$.

The symmetry conditions (isotropy in absence of external magnetic field) determine the Hamiltonian up to fourth order, as discussed in Ref. [58]:

$$\mathcal{H}_{\text{af}} = \frac{1}{2g_M} \left[\left(\frac{1}{c_M} \partial_\tau \mathbf{M}_\alpha \right)^2 + |\nabla \mathbf{M}_\alpha|^2 + \frac{r}{2} |\mathbf{M}_\alpha|^2 \right] + \frac{u_0}{2} |\mathbf{M}_\alpha|^4 - v_0 (|\mathbf{M}_1|^2 + |\mathbf{M}_2|^2)^2. \quad (21)$$

The antiferromagnetic coupling is g_M , the spin stiffness is c_M , and the effective mass of spin wave excitations is r . The fourth-order coupling u_0 comes from the “soft” implementation of the constraint $|\mathbf{M}_\alpha| = 1$ ⁸ and v_0 is the anisotropy between the two sublattices, justified by the microscopic physics [5,58]. The Hamiltonian can be transformed by rescaling τ and x, y , together with the couplings $u_0 \mapsto u$ and $v \mapsto v_0$ to set $g_M = c_M = 1$ so that the kinetic term becomes isotropic, giving

$$\mathcal{H}_{\text{af}} = \frac{1}{2} (\partial_\tau M)^2 + \frac{1}{2} |\nabla M|^2 + \frac{r}{2} |M|^2 + \frac{u}{2} (|M|^2)^2 - v |\mathbf{M}_1|^2 |\mathbf{M}_2|^2, \quad (22)$$

where we have also rewritten the quartic terms for convenience. Without anisotropy, the energy of the system is a function of $|\mathbf{M}_1|^2 + |\mathbf{M}_2|^2$ only and the symmetry group is the full $O(6)$. With $v \neq 0$, the symmetry is reduced to $O(3)_1 \otimes O(3)_2$: The internal spin symmetry in each sublattice remains unbroken but the spatial rotation symmetry between the layers is broken down to just the discrete flip. Compare this to the $U(1) \otimes U(1)$ symmetry in the PR system: There, it is the internal phase symmetry that remains unbroken.

⁸One could also enforce the constraint exactly, through the nonlinear σ model, as was done in Ref. [39]. While the leading term of the “vortex” Hamiltonian would remain the same in that case, the amplitude fluctuations have different dynamics which influences some terms of the Hamiltonian and thus its RG flow (though probably not the very existence of the glass phase).

A. \mathbb{Z}_2 vortices

Remember that topological solitons are classified by homotopy groups and that we work in a two-dimensional plane. The relevant group is again the first homotopy group, $\pi_1[O(3)] = \mathbb{Z}_2$. For simplicity, we will call these excitations “vortices,” bearing in mind that the only possible charges are $Q_\alpha = \pm 1$ and not all integers. A realization of the vortex with $Q = 1$ is shown in Fig. 15(b). Since the spins are three-dimensional (the figure shows the projection in the XY plane), it becomes clear that vortex charge is only defined modulo 2; i.e., it makes no sense to talk about charges $|Q| > 1$. For example, winding around twice in the XY plane can be done along a closed line in the XYZ space which can be contracted to a point. That could not happen for the two-dimensional phase $U(1)$ precisely because there is no extra dimension. In Fig. 15(b), the vortex is superimposed onto the regular configuration: It is recognizable as a contact point between two lines of alternating staggered magnetization. In Fig. 15(c) we have subtracted the regular part and only the vortexing spin pattern is shown: Here we see the vortex interpolates between two opposite spin orientations in two opposite directions in the plane.

Now let us derive the effective Hamiltonian of the vortices. For the \mathbb{Z}_2 vortex, a loop in real space is mapped onto a π arc in the internal space, so the vortex can be represented as

$$\mathbf{M}_\alpha(r, \phi) = \int d\phi' e^{\frac{i}{2}(\phi' - \phi)\hat{\ell}_3} \mathbf{m}_\alpha, \quad (23)$$

giving (the matrices $\ell_{1,2,3}$ represent the $so(3)$ algebra)

$$\mathbf{M}_\alpha = \begin{pmatrix} \cos \phi & \mp \sin \phi & 0 \\ \pm \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} m_{1\alpha} \\ m_{2\alpha} \\ m_{3\alpha} \end{pmatrix}, \quad (24)$$

where \mathbf{m}_α is the magnetization amplitude, analogous to the beam amplitude ψ_α in the optical system. The leading-order, noninteracting term in (22) gives the following for the energy of a single vortex of charge \tilde{Q} :

$$E_1 = 2\pi (|\mathbf{m}_X \times \mathbf{e}_Z|^2 + |\mathbf{m}_Y \times \mathbf{e}_Z|^2) \ln \Lambda = 2\pi |\mathbf{m}_{\perp\alpha}|^2 \ln \Lambda, \quad (25)$$

which is in fact independent of the sign of \vec{Q} (as could be expected, as it is in general proportional to $\vec{Q} \cdot \vec{Q}$ which is a constant for parity vortices). The vortex singles out an easy axis (Z axis) around which the staggered magnetization winds (ϕ being the winding angle). This allows one to introduce $\mathbf{m}_{\alpha\perp} \equiv (m_{X\alpha}, m_{Y\alpha}, 0)$. A vortex pair with charges \vec{Q}_i and \vec{Q}_j has the binding energy

$$\begin{aligned} E_2 &= 2\pi \vec{Q}_i \cdot \vec{Q}_j (|\mathbf{m}_1 \times \mathbf{e}_Z|^2 + |\mathbf{m}_2 \times \mathbf{e}_Z|^2) \ln r_{ij} \\ &= 2\pi |\mathbf{m}_{\perp\alpha}|^2 \vec{Q}_i \cdot \vec{Q}_j \ln r_{ij}. \end{aligned} \quad (26)$$

Now we should integrate out the amplitude fluctuations as we did in Appendix D for the CP beams. This again leads to the coupling between different flavors, giving a vortex Hamiltonian analogous to (7):

$$\mathcal{H}_{\text{vort}} = \sum_{i<j} (g \vec{Q}_i \cdot \vec{Q}_j + g' \vec{Q}_i \times \vec{Q}_j) \ln r_{ij} + \sum_i \vec{\mu} \cdot \vec{Q}_i. \quad (27)$$

Two obvious differences with respect to the optical system are (i) the charges are now limited to the values ± 1 , and (ii) there is a term linear in charge density, which acts as a chemical potential. The latter arises from the coupling of the *three-dimensional* spin waves (i.e., the topologically trivial excitations of the amplitude \mathbf{m}_α) to the vortices. Remember that in the CP system, the amplitude fluctuations also couple to the vortices, but there is no third, Z axis of the order parameter so no linear term appears. The microscopic expressions for the effective parameters g, g', μ_α read

$$g = m_\perp^2 + \frac{4r + 6um_\perp^2}{(2v + \frac{3}{2}um_\perp^2 + \frac{v}{2}m_\perp^2)(2r + \frac{3}{2}um_\perp^2 - \frac{v}{2}m_\perp^2)}, \quad (28)$$

$$g' = -\frac{4vm_\perp^2}{(2v + \frac{3}{2}um_\perp^2 + \frac{v}{2}m_\perp^2)(2r + \frac{3}{2}um_\perp^2 - \frac{v}{2}m_\perp^2)}, \quad (29)$$

$$\mu_\alpha = \frac{1}{2} m_\perp m_z, \quad (30)$$

assuming $m_{1\perp} = m_{2\perp} \equiv m_\perp$. Now the RG calculation is similar to the optical case but the nonzero chemical potential introduces two differences. First, there is obviously the additional term proportional to the total charge of the virtual pair of vortices, $\mu_\alpha(q_{1\alpha} + q_{2\alpha})$. Second, there is no charge conservation as the expectation value of the total vortex charge is now $\langle \vec{Q} \rangle = \partial \mathcal{F} / \partial \vec{\mu} \neq 0$. Thus we need to take into account not only the fluctuations with zero net charge (virtual vortex pairs with charges $\vec{q}_1 \equiv \vec{q}$ and $\vec{q}_2 \equiv -\vec{q}$) but also the situations with arbitrary pairs \vec{q}_1, \vec{q}_2 .⁹ This modifies the variation of the

partition function from (12) and (13) to

$$\begin{aligned} \frac{\delta \mathcal{Z}}{\mathcal{Z}} &= 1 + \frac{y^4}{4} \sum_{\vec{q}_{1,2}} \int dr r^2 r_{12}^3 e^{-g\vec{q}_1 \cdot \vec{q}_2 - g'\vec{q}_1 \times \vec{q}_2 - \vec{\mu} \cdot \vec{q}'} \\ &\times \left[\int dr r^2 (g \vec{Q}_1 \cdot \vec{q} + g' \vec{Q}_1 \times \vec{q}) \nabla \ln |\delta \mathbf{R}_1| \right. \\ &\left. + (g \vec{Q}_2 \cdot \vec{q} + g' \vec{Q}_2 \times \vec{q}) \nabla \ln |\delta \mathbf{R}_2| \right]^2 \\ &+ \frac{y^4}{4} \sum_{\vec{q}_{1,2}} \int dr r^2 r_{12}^3 e^{-g\vec{q}_1 \cdot \vec{q}_2 - g'\vec{q}_1 \times \vec{q}_2 - \vec{\mu} \cdot \vec{q}_1} \\ &\times \left[\int dr r^2 (g \vec{Q}_1 \cdot \vec{q}_0 + g' \vec{Q}_1 \times \vec{q}_0) \nabla \ln |\delta \mathbf{R}_1| \right. \\ &\left. + (g \vec{Q}_2 \cdot \vec{q}_0 + g' \vec{Q}_2 \times \vec{q}_0) \nabla \ln |\delta \mathbf{R}_2| \right]^2, \end{aligned}$$

where we have introduced $2\vec{q} \equiv \vec{q}_1 - \vec{q}_2, \vec{q}_0 \equiv \vec{q}_1 + \vec{q}_2$ and $\delta \mathbf{R}_{1,2} \equiv \mathbf{R}_{1,2} - \mathbf{r}$. The mixed term which includes both \vec{q} and \vec{q}_0 vanishes due to isotropy. By matching the terms in the resulting expression with the original Hamiltonian, we find the recursion relations:

$$\begin{aligned} \frac{\partial g}{\partial \ell} &= -16\pi y^4 (g^2 + g'^2), & \frac{\partial g'}{\partial \ell} &= -16\pi y^4 g g', \\ \frac{\partial \vec{\mu}}{\partial \ell} &= 0, & \frac{\partial y}{\partial \ell} &= (1 - g - g' - \mu_+ - \mu_-) y. \end{aligned} \quad (31)$$

Crucially, the chemical potential does not run which could be guessed from dimensional analysis (it couples to dimensionless charge). This is the same system as (14) up to the trivial rescaling of the coupling constants and the shift of the critical line $g + g' = 1$ in the PR system to the line $g + g' + \mu_+ + \mu_- = 1$. It becomes obvious that the phase diagrams are equivalent and can be mapped onto each other.

B. Influence of disorder

The disorder in a doped antiferromagnet comes from electrically neutral metallic grains quenched in the bipartite lattice. Being metallic and neutral, they are naturally modeled as magnetic dipoles \mathbf{X} quenched in the bipartite lattice. This picture stems from the microscopic considerations in Ref. [48]. We again assume the Gaussian distribution of the disorder as $p(X) \propto \exp(-|\mathbf{X}|^2/2\sigma_X^2)$. The disorder dipoles are one and the same for both sublattices, so \mathbf{X} has no flavor (sublattice) index. The minimal coupling of the dipoles to the lattice spins $\partial_i \mapsto \partial_i - i \hat{\ell}_i X_i$ gives

$$\mathcal{H}_{\text{af}} \mapsto \mathcal{H}_{\text{dis}} = \mathcal{H}_{\text{af}} + \nabla \mathbf{M}_\alpha \cdot (\mathbf{X} \times \mathbf{M}_\alpha) + M^2 X^2. \quad (32)$$

Now the replica calculation requires the multiplication of the M field into n copies and performing the Gaussian integral over the disorder. The initial distribution of the disorder $p(X)$ gives rise to two independent Gaussian distributions: for the couplings $J_{ij}^{\alpha\beta}$ with dispersion matrix $\sigma_{\alpha\beta}^2$ and for the chemical potential μ_i^α with the dispersion vector ξ_α^2 . The resulting

⁹In the CP beam system, the total vortex charge can be nonzero if the boundary conditions at $z = 0, L$ have nonzero total vorticity. But there we had no *bulk* chemical potential so the total vorticity in the crystal could not change during the propagation along z . Here, we have a *bulk* term in the Hamiltonian which violates charge conservation.

Hamiltonian is

$$\begin{aligned} \mathcal{H}_{\text{dis}} = & \sum_{\mu=0}^n \left(\frac{1}{2} |\partial_{\tau} \mathbf{M}_{\alpha}^{(\mu)}|^2 + \frac{1}{2} |\nabla \mathbf{M}_{\alpha}^{(\mu)}|^2 \right. \\ & \left. + \frac{u}{2} |\mathbf{M}_{\alpha}^{(\mu)}|^2 - v |\mathbf{M}_1^{(\mu)}|^2 |\mathbf{M}_2^{(\mu)}|^2 \right) \\ & + \frac{\sigma^2}{4} \sum_{\mu, \nu=0}^n (\nabla \mathbf{M}_{\alpha}^{(\mu)} \times \mathbf{M}_{\alpha}^{(\mu)}) \cdot (\nabla \mathbf{M}_{\alpha}^{(\nu)} \times \mathbf{M}_{\alpha}^{(\nu)}), \quad (33) \end{aligned}$$

where we have disregarded the subleading logarithmic term ($\sim \ln |\mathbf{M}_{\alpha}^{(\mu)}|$). Now making use of the representation (23) and plugging it into (33) gives the disordered vortex Hamiltonian

$$\begin{aligned} \beta \mathcal{H}_{\text{vort}} = & \sum_{\mu, \nu=1}^n \sum_{i, j=1}^N \left[\frac{\beta^2}{2} \mathcal{Q}_{i\alpha}^{(\mu)} \mathcal{Q}_{i\beta}^{(\nu)} \mathcal{Q}_{j\alpha}^{(\mu)} \mathcal{Q}_{j\beta}^{(\nu)} \right. \\ & \left. - \beta \mathcal{Q}_{i\alpha}^{(\mu)} J_0^{\alpha\beta} \mathcal{Q}_{j\beta}^{(\mu)} + \beta^2 \mathcal{Q}_{i\alpha}^{(\mu)} \xi^2 \mathcal{Q}_{j\alpha}^{(\nu)} \right] \\ & - \sum_{\mu=1}^n \sum_{i=1}^N \beta \xi^2 \mu_0^{\alpha} \mathcal{Q}_{i\alpha}^{(\mu)}. \quad (34) \end{aligned}$$

Of course, we could have arrived at the same effective action starting from the vortex Hamiltonian (27), taking the infinite-range approximation and identifying $J_{ij}^{\alpha\beta} = g_{ij} \ln r_{ij}$ and similarly for other components of $J_{ij}^{\alpha\beta}$ as we demonstrated for the PR system. The final result has to be same at leading order.

The next step is to rewrite the Hamiltonian in terms of the order parameters $p_{\alpha}^{(\mu)}, q_{\alpha\beta}^{(\mu)}$ defined in (19). Compared to the effective action for the photonic lattice with disorder in Eq. (G4), there are two extra terms in the resulting action S_{eff} : One is proportional to the dispersion ξ^2 and the other to the mean chemical potential $\bar{\mu}_0$. The former term just introduces the shift $J_0^{\alpha\beta} \mapsto J_0^{\alpha\beta} - \sigma^2/2\beta$ and the latter term, linear in the vortex charges and proportional to the chemical potential, introduces solutions with nonzero net vortex charge density. Looking back at the results of the saddle-point calculation in Eqs. (19) and (G14), this tells us that the relation between the phase diagrams is the following. The phases with no net vortex charge density—insulator, conductor, frustrated insulator, and perfect conductor—remain the same as in the PR system, since both the average coupling value $J_0^{\alpha\beta}$ (which gets shifted) and the term proportional to the chemical potential μ_{α} couple only to $\vec{p}^{(\mu)}$. For brevity, denote $J_0^{\pm\pm} \equiv J_0^{\pm}$ and notice that $J_0^{-+} = J_0^{+-}$. The structure of phases with nonzero $\vec{p}^{(\mu)}$ depends on the zeros of the saddle-point equation

$$\begin{aligned} J_0^{\pm} p^{\pm} + \left(\frac{J_0^{+-}}{\beta} - \frac{\beta}{2} \xi^{\pm} \right) p^{\mp} + (p^{\pm})^{-1} \\ - \frac{\mu_0^{\pm} (\sigma^{\pm\pm})^2 + \mu_0^{\mp} \sigma_{\pm}^2}{\beta} = 0, \quad (35) \end{aligned}$$

analogous to (G13), where the one-step replica symmetry breaking implies $p_{(\mu)}^{\pm} = (p^{\pm}, \dots, p^{\pm})$. Now the equation is cubic and the structure of solutions is different from (G14). We could not find the solution in the closed form but it is clear that a pair of cubic equations will have either a

single solution (p^+, p^-) or nine combinations (p^+, p^-), not necessarily all different. Numerical analysis of (35) reveals only two inequivalent solutions, analogous to (G14), i.e., one of them has a single free energy minimum and the other one a pair of degenerate minima. Therefore, we again have two disordered solutions, one of which is glassy (frustrated).

Now we can write down also the RG equations for the effective action (34). In this calculation, we put $\xi_{\alpha}^2 = \sigma_{\alpha\beta}^2 \equiv \sigma^2$ for simplicity. Following the same logic as earlier, the equations are found to be¹⁰

$$\begin{aligned} \frac{\partial g}{\partial \ell} &= -8\pi(g + g')^2 y^4 \cosh(2\beta^2 \sigma^2) \\ &\quad \times \cosh(2\beta^2 \sigma^2) - 8\pi(g - g')^2 y^4, \\ \frac{\partial g'}{\partial \ell} &= -\pi(g + g')^2 y^4 \cosh(2\beta^2 \sigma^2) \\ &\quad \times \cosh(2\beta^2 \sigma^2) - \pi(g - g')^2 y^4, \\ \frac{\partial y}{\partial \ell} &= 2\pi(1 - g - g' - \mu_+ - \mu_- - \beta^2 \sigma^2) y, \\ \frac{\partial \mu}{\partial \ell} &= -8\pi \mu, \\ \frac{\partial \sigma^2}{\partial \ell} &= -2\pi \beta^4 \sigma^4 y^4. \quad (36) \end{aligned}$$

Like in the clean case, the chemical potential is irrelevant and the solutions for fixed point are the same as for the PR beams, including the spin-glass fixed point. We conclude that the phase structure of the optical system is repeated in strongly correlated doped antiferromagnets, which also exhibit the spin-glass phase and have the phase diagram sketched in Fig. 16. In this context, it is more interesting to plot the phase diagram in the $\sigma^2 - 1/g'$ plane, mimicking the $x - T$ phase diagram of quantum critical systems [38] (remember that the coupling constants g, g' behave roughly as inverse temperature in XY -like models). Bear in mind that all phases shown are about vortex dynamics; i.e., one should not compare Fig. 16 to the textbook phase diagram of high-temperature superconductors, which accounts also for the charge or stripe order and the superconducting order. All vortex phases would be located inside the pseudogap regime of the superconductor, where various exotic orders can coexist (assuming, of course, that our model is an adequate approximation of the magnetic order in a cuprate or similar material, which is a complex question). Crucially, the spin-glass phase (blue curves) flows toward *finite disorder* σ^2 , whereas the remaining two phases end up at zero disorder, either at infinite $1/g'$ (PC, red flows) or at zero $1/g'$ (conductor, green flows). The RG flows in the conductor phase are almost invisible in the figure, as the flows are much slower than in the remaining two phases.

Discussion

Early papers which found and explored the spin-glass phase in a very similar model are Refs. [39,40,47,49]. The

¹⁰For the most general case of different and nonscalar $\sigma_{\alpha\beta}^2$ and ξ_{α}^2 , the flow equations for them complicate significantly and we will not consider them.

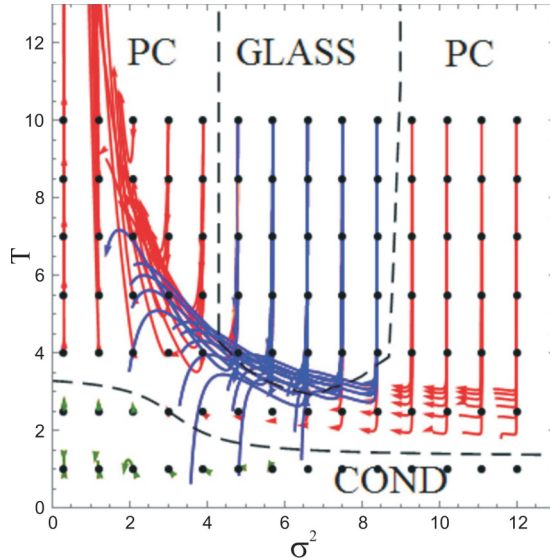


FIG. 16. The phase diagram of the two-sublattice-doped Heisenberg antiferromagnet model in the σ^2 - T plane (we have rescaled $\sigma^2 \mapsto 12\sigma^2$). Since $T \sim 1/g'$, we can alternatively understand the vertical axis as $1/g'$. Black dashed curves are approximate phase boundaries. RG flows (starting from black dots) are colored differently according to the phase they belong to: spin glass (blue), PC (red), and conductor (green). At high temperatures, the vortex conductor becomes either a perfect vortex conductor or a spin glass. Spin glass (blue) is recognized by the fact that the RG equations flow to nonzero disorder at finite and large g' (low temperatures). The PC phase (red) flows toward zero disorder and zero coupling (infinite T), collapsing practically to a single trajectory. The flows for the conductor (green) end up at $T = \sigma^2 = 0$ but are not shown to scale in the figure. Parameter values are $u = r = 1$ with varying v so as to have $g = -0.5$ for all trajectories.

main difference is that the papers cited consider the spiral (noncollinear) spin order. These works are all inspired by the cuprate materials, the most celebrated brand of high-temperature superconductors. While Refs. [40,47] explore in detail the transport properties, we have no pretension either to provide a realistic model of cuprates or to explore in detail all the properties of the spin-glass phase. We are content to see that the PR system of \mathbb{Z} vortices reproduces the phase structure of a certain kind of dirty Heisenberg antiferromagnets (with $O(3)$ spins and \mathbb{Z}_2 vortices), besides the more obvious connection to systems which directly reproduce the \mathbb{Z} vortices in multicomponent $U(1)$ systems like multicomponent Bose-Einstein condensates and type-1.5 superconductors.

VI. CONCLUSIONS

We have investigated the light intensity patterns in a nonlinear optical system consisting of a pair of counterpropagating laser beams in a photorefractive crystal. We have studied this system as a strongly interacting field theory and have focused mostly on the formation and dynamics of vortices. The vortices show a remarkable collective behavior and their patterns are naturally classified in the framework of statistical field theory: The effective action shows several different

phases with appropriate order parameters, and the system is essentially an XY model with two flavors, i.e., two kinds of vortex charge, for the two beams. The interaction between the flavors is the central reason that the total energy of the Coulombic interactions between the vortices in general cannot be locally minimized at every point. In the presence of disorder, a phase with multiple free energy minima arises, where the absence of long-range order is complemented by the local islands of ordered vortex structure, and which resembles spin glasses.

The phase diagram is simple in terms of the effective parameters—vortex coupling constants—and quite complex when expressed in terms of the experimentally controllable quantities—the intensity of the laser beams, the intensity of the background photonic lattice, and the properties of the photorefractive crystal (the last is not controllable but can be estimated reasonably well [3]). The lesson is that the approach we adopt can save us from demanding numerical work if the space of original parameters is blindly explored. Our phase diagrams can serve as a starting point for guided numerical simulations, suggesting what phenomena one should specifically look for. So far the field-theoretical and statistical approach was not much used in nonlinear optics (important exceptions are Refs. [9–12, 14, 15, 50, 51, 59, 60]). We hope to stimulate work in this direction, which is promising also because of the potential of the photorefractive systems to serve as models of strongly correlated condensed matter systems. They make an excellent testing ground for various models because of the availability and relatively low cost of experiments.

In this work, we have focused on the relation of the photorefractive counterpropagating system to the model of an $O(3)$ doped antiferromagnet with two sublattices. The authors of previous works on this model [40,47,48,58] were motivated mainly by the ubiquitous problem of understanding the pseudogap phase in cuprate superconductors. The applicability of the model to this particular problem is still an open question; it may well be that cuprate physics goes far beyond. Nevertheless, it is an important quantum magnetic system in its own right and serves as an illustration of how one can simulate condensed matter systems in photorefractive optics.

Another field where vortices are found as solutions of a nonlinear Schrödinger equation are cold atom systems and Bose-Einstein condensates [26]. Notice, however, that Bose-Einstein condensates in optical traps are usually (but not always; see Ref. [30]) three-dimensional systems with vortex lines (rather than XY -type systems with point vortices) and our formalism would be more complicated there: In three spatial dimensions, vortices give rise to emergent gauge fields. The multicomponent systems of this kind give rise to so-called type-1.5 superconductors [53], which are a natural goal of further study.

A more complete characterization of the glasslike phase is also left for further work. The reader will notice we have devoted very little attention to the correlation functions in various parameter regimes or the scaling properties of susceptibility, which should further corroborate the glassy character of the system. This is quite difficult in general but very exciting as it offers an opportunity to tune the parameters (e.g., disorder strength) freely in the optical system and study the glasslike phase and its dynamics.

ACKNOWLEDGMENTS

We are grateful to Mariya Medvedyeva and Vladimir Juričić for careful reading of the manuscript. Work at the Institute of Physics is funded by Ministry of Education, Science and Technological Development, under Grants No. OI171033 and No. OI171017. M.P. is also supported by the NPRP 7-665-1-125 project of the Qatar National Research Fund (a member of the Qatar Foundation).

APPENDIX A: NUMERICAL ALGORITHM

In order to solve numerically the system [(1) and (2)], we employ a variation on the method of Refs. [61,62]. The method does not make use of any analytical ansatz: It is an *ab initio* numerical procedure which integrates the equations of motion. The system has four independent variables: the transverse coordinates (x, y) , the longitudinal coordinate (formal time) z , and the (physical) time t . That means we have essentially three nested loops: (i) At every z slice we integrate the transverse Laplacian and the interaction terms for the whole z axis, (ii) we advance the time t , and (iii) we repeat the whole procedure until reaching some time t_f , which certainly should be much longer than the relaxation time τ .

The important point is the very different natures of the initial and boundary conditions for various coordinates. The boundary conditions in the (x, y) plane, i.e., at the crystal edge are not crucial: We have either just one or a few Gaussian beams whose intensity drops exponentially away from the center and is practically zero at the crystal edge, or we have a large lattice consisting of many (of the order of 50–100) Gaussian beams so the edge effects only affect a small portion of the whole lattice. Therefore, imposing periodic boundary conditions (stemming naturally from the integration in Fourier space, see the next paragraph) are perfectly satisfying. Crucially, however, the CP geometry means that $F(t; z = 0; x, y) = F_0(x, y)$ and $B(t; z = L; x, y) = B_0(x, y)$ are given functions, fixed for all times. We thus have a two-point boundary value problem along z and have to iterate the z integration several times until we reach the right solution. Finally, the initial condition for the relaxation equation (2) is that the crystal is initially at equilibrium, meaning that $E(t = 0) = -I_x/(1 + I_x)$; specifically, for zero background lattice, $E(t = 0) = 0$.

The algorithm now has the following structure:

(1) The innermost loop integrates in the x - y plane. This is a Poisson-type (elliptic) equation, thus we employ the operator-split method, integrating the Laplacian operator in the Fourier space and the interaction term (the EF and EB terms) in real space, in the second-order leapfrog scheme. Thus, at every time instant $t_i = i \Delta t$, we start from $z = 0$ where we set the condition $F(i \Delta t; z = 0; x, y) = F_0(x, y)$, divide the z axis into N steps of size $\Delta z = L/N$, and at every slice $z = j \Delta z$ perform the frog’s leap: We do the fast Fourier transform (FFT) to turn the (x, y) dependence into (q_x, q_y) dependence,¹¹ then we advance the Laplacian for $\Delta z/2$ as $F(i \Delta t; j \Delta z; \mathbf{q}) \equiv \tilde{F}_{i,j}^{(0)} \mapsto \tilde{F}_{i,j}^{(1)} = \exp(-iq^2 \Delta z/2) \tilde{F}_{i,j}^{(0)}$, and then we do the inverse FFT and advance the interaction in real space as

$F_{i,j}^{(2)} = \exp[i\Gamma E(i \Delta t; j \Delta z; x, y)] F_{i,j}^{(1)}$. Finally we do the FFT again and advance the Laplacian for the remaining half-step, $\tilde{F}_{i,j+1} = \exp(-iq^2 \Delta z/2) \tilde{F}_{i,j}^{(2)}$. Once we reach $j = N$, the integration goes backward, along the same lines, updating now the B field [starting from $B_0(x, y)$], where all signs in the exponents of the above formulas are to be reversed. When we reach $z = 0$ again, we are done. In this loop, we use the field $E_{1,j}$ as already known for all j .

(2) The above loop will, in general, produce results inconsistent with the charge field $E_{i,j}$ because the equation for E couples F and B and we have ignored that by integrating the two fields one after the other instead of simultaneously. This is, of course, commonplace in two-point boundary value problems: Either only one boundary condition can be imposed exactly and the other is shot for or, as in our case, both are imposed exactly but at the cost of the solution being inconsistent with the equations, so we have to iterate the system to arrive at the correct solution everywhere. The second loop thus iterates the first loop A times, at each step updating the charge field as $E_{i,j}^{(a-1)} \mapsto E_{i,j}^{(a)} = E_{i-1,j} - \tau [E_{i,j}^{(a-1)} + I_{i,j}^{(a-1)} / (1 + I_{i,j}^{(a-1)})] / (1 + I_{i,j}^{(a-1)})$. The number of iterations A is not fixed: We stop iterations when the intensity pattern stabilizes, $\sum_j \sum_{x,y} (I_{i,j}^{(a)} - I_{i,j}^{(a-1)}) < \epsilon$, for some tolerance ϵ . Here, $I_{i,j}$ refers to total intensity, i.e., $|F|^2 + |B|^2 + I_x$.

(3) Finally, the outermost loop integrates in time t , from $t = 0$, with the initial condition $E(t = 0) = -I_x / (1 + I_x)$ given above. The integration time t_f is divided into $M = t_f / \Delta t$ intervals, and at the end of each step we update $(F_{i,j}, B_{i,j}, E_{i,j}) \mapsto (F_{i+1,j}, B_{i+1,j}, E_{i+1,j})$. Only the charge field is directly integrated (as written above), in the first-order, Euler scheme. The beam envelopes depend on time only parametrically, through $E(t)$, and they evolve by using an updated $E_{i,j}$ in the first two loops at every time step.

This procedure is very close to that in Ref. [61]; the main difference is that we use a second-order (leapfrog) scheme, while on the other hand our time integration is of the lowest, linear order instead of second order as in Ref. [61].

APPENDIX B: TIME-DEPENDENT PERTURBATION THEORY AND THE EXISTENCE OF EQUILIBRIUM CONFIGURATIONS

1. Stability analysis: fixed points and limit cycles

In this appendix, we consider the time evolution of the CP beams and show the existence of a stable equilibrium point with nonzero intensity. This means that the system reaches a stationary state for long times, justifying the basic assumption of the paper that one can study the vortex configurations within equilibrium statistical mechanics. Not all patterns are stable: Depending on the boundary conditions and parameter values, the system may or may not have a stable equilibrium, and nonequilibrium solutions in photorefractive optics are well known [37,50]. For our purposes, however, it is enough to identify the region of parameter space where the equilibrium exists; other cases are not the topic of this paper.

The time evolution of the beams Ψ_α and the charge field E in (k, q) space is obtained by differentiating Eqs. (1) with respect to time and plugging in $\partial E / \partial t$ from the relaxation

¹¹We denote the fields in Fourier space with a tilde, e.g., \tilde{F} .

equation (2):

$$\begin{aligned}\frac{\partial \Psi_{\alpha}^{\pm}}{\partial t} &= -\frac{\Gamma}{\tau} \frac{[(1+I)E + I]}{\alpha k - q^2 - \Gamma E} \Psi_{\alpha}^{\pm}, \\ \frac{\partial E}{\partial t} &= -\frac{1}{\tau} [(1+I)E + I].\end{aligned}\quad (\text{B1})$$

This system has three equilibrium points. One is the 0 point,

$$(\Psi_{+}^{\pm}, \Psi_{-}^{\pm}, E) = \left(0, 0, -\frac{I_x}{1+I_x}\right),$$

and the remaining two are related by a discrete symmetry $\Psi_{\pm} \mapsto \Psi_{\mp}$, so we denote them as “ \pm ” points, with the “+” point being

$$(\Psi_{+}^{\pm}, \Psi_{-}^{\pm}, E) = \left(\sqrt{\frac{E(1+I_x) + I_x}{1+E}} e^{i\phi_{\pm}}, 0, E\right),$$

and the “-” point has instead $\Psi_{+} = 0$ and $\Psi_{-} = \sqrt{(E(1+I_x) + I_x)/(1+E)} \exp(i\phi_{-})$. Notice that the phase ϕ_{\pm} remains free to vary, so this solution supports vortices. The 0 point is the trivial vacuum, i.e., the zero-intensity configuration with only background lattice. The fluctuation equations about this point to quadratic order read

$$\begin{aligned}\partial_t X &= -\left[-f_{+} X_1 X_5, -f_{+} X_2 X_5, -f_{-} X_3 X_5, -f_{-} X_4 X_5, \right. \\ &\quad \left. -\frac{1}{1+I_x} (X_1^2 + X_2^2 + X_3^2 + X_4^2) - (1+I_x) X_5 \right],\end{aligned}\quad (\text{B2})$$

where we have introduced the real variables $X_{1,3} = \text{Re} \delta \Psi_{\pm}$, $X_{2,4} = \text{Im} \delta \Psi_{\pm}$, $X_5 = \delta E$ and

$$f_{\pm} = \frac{\Gamma(1+I_x)^2}{\Gamma I_x \mp (1+I_x)(k \pm q^2)}.\quad (\text{B3})$$

The system (B2) is degenerate at linear order; thus, we need a quadratic order expansion to analyze stability. The simplest approach is to construct a Lyapunov function for Eq. (B2). The function $V(X) = X^2$ is positive for and only for $X \neq 0$, and its derivative is

$$\begin{aligned}\frac{dV}{dt} &= -2f_{+}(X_1^2 + X_2^2)X_5 - 2f_{-}(X_3^2 + X_4^2)X_5 \\ &\quad -\frac{1}{1+I_x} (X_1^2 + X_2^2 + X_3^2 + X_4^2)X_5 - (1+I_x)X_5^2,\end{aligned}\quad (\text{B4})$$

which is strictly negative for X nonzero if $f_{\pm} > 0$ and $X_5 > 0$. However, we always have $X_5 > 0$ because dX_5/dt in the full relaxation equations (B1) has a strictly negative right-hand side and E grows monotonically from zero to $-I_x/(1+I_x)$, and at any finite t we have $E(t) - E(t \rightarrow \infty) = X_5 > 0$. Thus the trivial equilibrium point is *locally* stable for $f_{+} > 0$, $f_{-} > 0$, i.e., $k > q^2$. It is much harder to construct the Lyapunov function for the global equations (B1): In this case, there are no additional symmetries and the stability of higher dimensional systems is in general an extremely difficult topic. Thus there may well be regions far away from the 0 point which do not flow toward it.

The “ \pm ” pair is quite hard to study. All hope of expanding the system to second order and understanding the resulting complicated five-variable system is lost. This time, however, we can do a nontrivial first-order analysis as the system is nondegenerate and nicely reduces to the (X_1, X_5) subsystem. Rescaling $X_1 \mapsto (1+E_0)^{-3/4} [I_x + E_0(1+I_x)]^{1/2}$ and $t \mapsto t\{(1+E_0)/[I_x + E_0(1+I_x)]\}^{1/4}$, the equation of motion for the \pm point reads

$$\begin{aligned}\partial_t \begin{pmatrix} X_1 \\ X_5 \end{pmatrix} &= \begin{pmatrix} -\frac{a_{\pm}}{\Gamma E_0 + k + q^2} & -1 \\ 1 & -\frac{a_{\pm}}{\Gamma E_0 + k + q^2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_5 \end{pmatrix} \\ &\quad + O(X_1^2 + X_5^2; X_2, X_3, X_4),\end{aligned}\quad (\text{B5})$$

with a_{\pm} being some (known) *positive* functions of Γ, E_0, I_x (independent of k, q). This is precisely the normal form for the Andronov-Hopf bifurcation [63], and the bifurcation point lies at $k = -\Gamma E_0 - q^2$. As a reminder, the bifurcation happens when the off-diagonal element in the linear term changes sign: The fixed point is stable when $a_{\pm}/(\Gamma E_0 + k + q^2)$ is positive. The sign of the nonlinear term determines the supercritical or subcritical nature of the bifurcation. A negative sign means the fixed point is stable everywhere before the bifurcation and is replaced by a stable limit cycle after the bifurcation (supercritical). A positive sign means the fixed point coexists with the stable limit cycle before the bifurcation and the (X_1, X_5) plane is divided among their attraction regions; after the bifurcation there is no stable solution at all (subcritical).¹²

In conclusion, stable + equilibrium exists for $k > -\Gamma E_0 - q^2$ where E_0 is best found numerically. Exactly the same condition holds for the - point. For $k < -\Gamma E_0 + q^2$, dynamics depends on the sign of the nonlinear term in (B5): For the positive sign, we expect periodically changing patterns. If the term is negative and the bifurcation is subcritical, various possibilities arise: The system may wander chaotically between the + and the - point, or it may end up in the attraction region of the 0 point and fall onto the trivial solution with zero intensity. Naively, the attraction regions of the two fixed points (\pm and 0) are separated by the condition $-\Gamma E_0 - q^2 = q^2$, i.e., $q_c = \sqrt{-\Gamma E_0(\Gamma, \tau)/2}$, where we have emphasized that E_0 is in general nonuniversal. The actual boundary may be more complex, however, as our analysis is based on finite-order expansion around the fixed points, which is not valid far away from them.

The outcome is that the system generically has stable trivial and nontrivial (nonzero intensity) equilibria, in addition to time-dependent, periodic, or aperiodic solutions. Numerical integration gives a similar picture of the stability diagram in Fig. 17. Numerically we find that the stability limit is $k > \Gamma - q^2$, i.e., $E_0 \approx -1$. The region of applicability of our formalism lies in the top right corner of the diagram (nontrivial equilibrium), above $k \approx 1/L$. Formally, both k and q can be any real numbers. In practice, however, k is discrete and its minimal value is of the order $1/L$. The spatial momentum q

¹²One should not take the stability in the whole (X_1, X_5) plane in the supercritical case too seriously. We have expand the equations of motion in the vicinity of the fixed points and the expansion ceases to be valid far away from the origin.

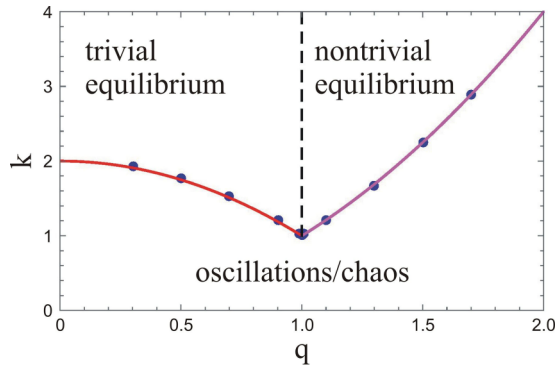


FIG. 17. Stability diagram in the q - k plane. The onset of instability for $k < k_c(q)$ is found numerically for a range of q values. The solid lines are the analytical prediction for the stability of the 0 point ($k_c = q^2$, magenta) and of the + point ($k_c = \Gamma E_0 - q^2 \approx \Gamma - q^2$, red). The black dashed line at $q = q_c \approx 1$ separates the stability regions of the two points. The domain of applicability of our main results is the top left corner (nontrivial equilibrium), above $k > k_{\min} \sim 1/L$ and for not very large q values. Parameter values: $\Gamma = 2, I_x = 0$.

lies between the inverse of the transverse length of the crystal (which is typically an order of magnitude smaller than L , i.e., minimal q can be assumed equal to zero) and some typical small-scale cutoff which in our case is the vortex core size. We made no attempt to study the nonequilibrium behavior in detail or to delineate the boundary between the oscillatory and the chaotic regime since it is irrelevant for the main story of the paper.

From a practical viewpoint, the Γ - I_x plane can be divided into two regions. One of them has a single stable “+” or “-” equilibrium or a $+ \mapsto -$ limit cycle whose amplitude vanishes in the thermodynamic limit at all scales, i.e., for all (k, \mathbf{q}) . This region can be legitimately described within the formalism of partition functions and equilibrium field theory. The second region flows toward the trivial fixed point and does not support vortices—this can also (trivially) be described by our formalism, as it always corresponds to the insulator regime, with no stable vortices. Thus the consistency check is that our method predicts no other phases in this region but insulator. In the third regime, long-term dynamics is either a limit cycle with amplitude of order unity or chaos. This regime was studied in detail in some earlier publications (e.g., Ref. [3] and references therein), and it cannot be reached within our present formalism.

2. Numerical checks

Now we complement the analytical considerations with numerical evidence that the phases described in the main text exist as long-term stable configurations. In Fig. 18, we show the time evolution of a vortex lattice in three different phases, where a visual inspection clearly suggests the system approaches equilibrium. In contrast, in Fig. 19 we see first a pattern that oscillates forever, i.e., follows a limit cycle [Fig. 19(a)], becomes incoherent [wandering chaotically over the unstable manifold, Fig. 19(b)], or dissipates away (reaching the 0-fixed point), in Fig. 19(c). The loss of stability corresponds to an Andronov-Hopf bifurcation, as found earlier for nonvortex patterns in Ref. [32].

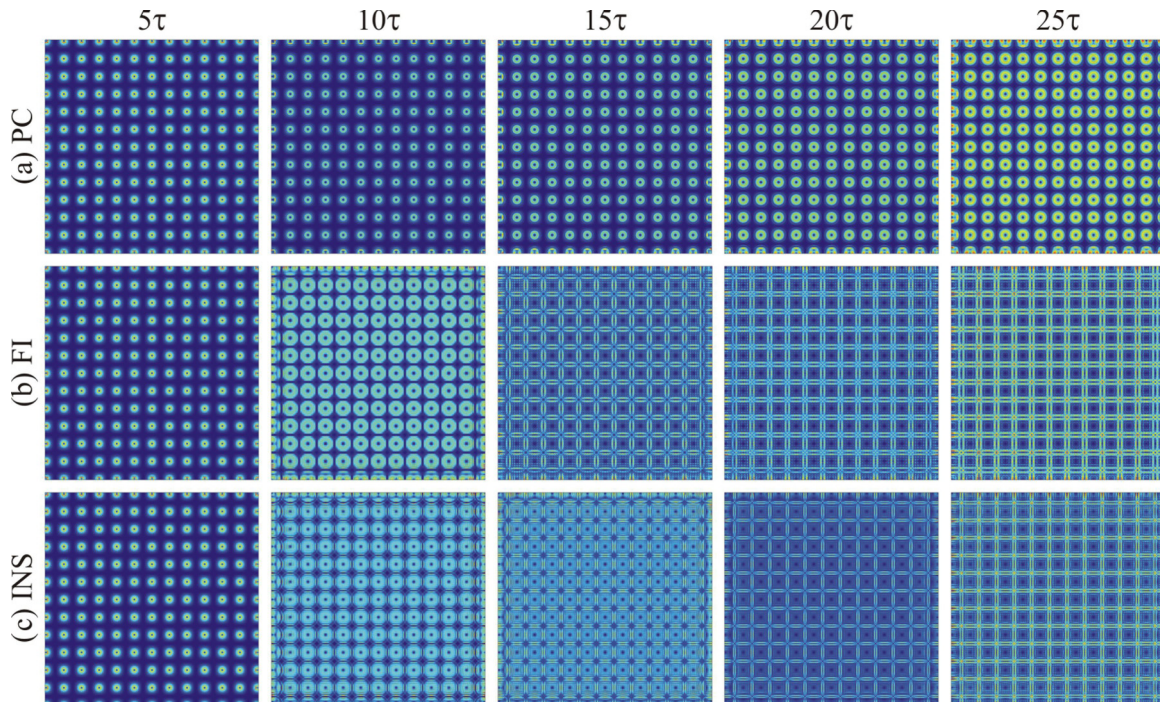


FIG. 18. Time evolution of patterns at five different times: (a) perfect conductor phase, (b) frustrated insulator phase, and (c) insulator phase. In all cases, the approach to equilibrium is obvious, and we expect that for long times a thermodynamic description is justified. The parameters are the same as in Fig. 6, for the corresponding phases.

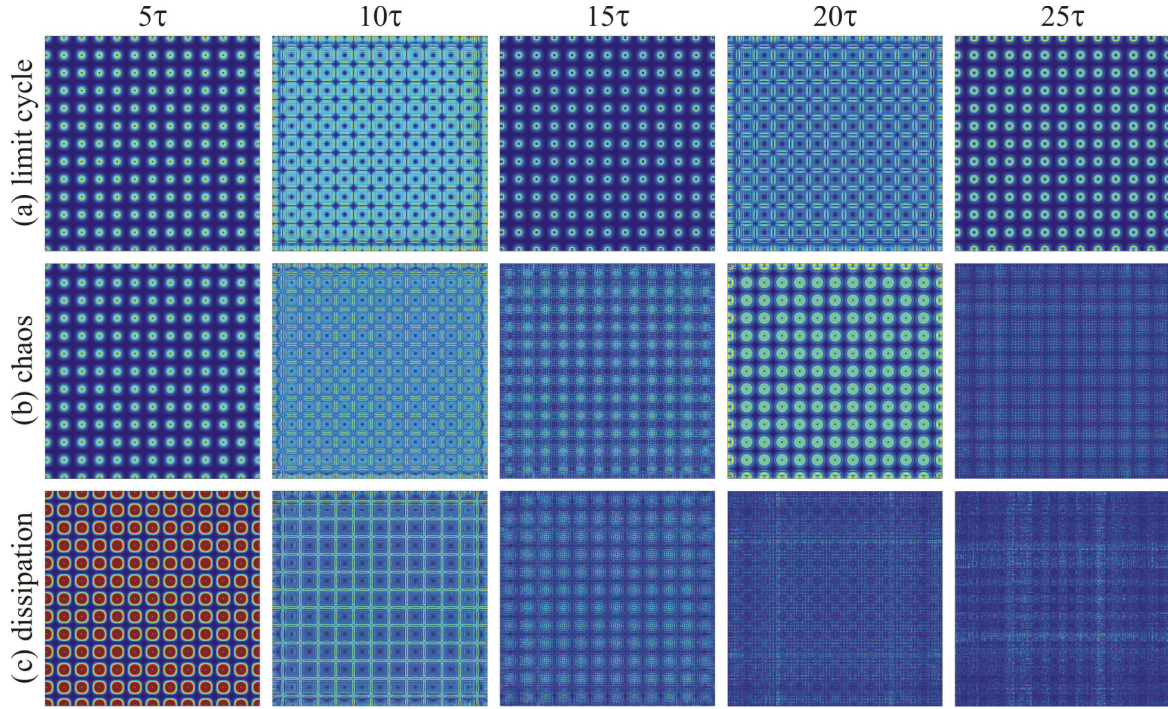


FIG. 19. Time evolution of nonequilibrium patterns. In panel (a), the limit cycle leads to permanent oscillatory behavior, in panel (b) wandering along the unstable manifold between the equilibrium points gives rise to chaos, and in panel (c) dissipation wins and dynamics dies out. The parameters are the same as in the previous figure, except that the length L is increased three times.

Dynamics can be most easily traced by looking at the numerically computed relaxation rate

$$\frac{1}{X} \frac{dX}{dt} = \frac{\sum_{x,y} |X(t_{j+1}, x, y) - X(t_j; x, y)|^2}{\sum_{x,y} |X(t_j; x, y)|^2}, \quad (\text{B6})$$

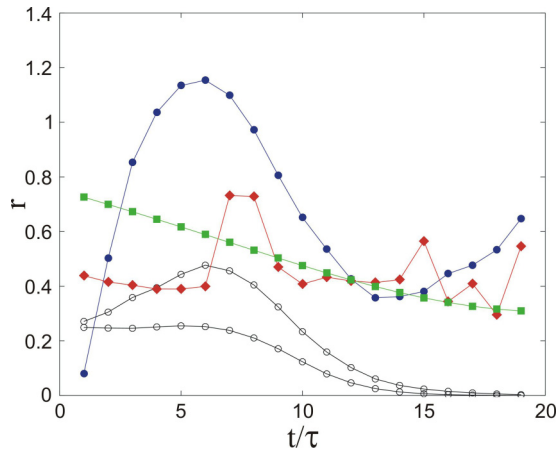


FIG. 20. Time evolution of the relaxation rate r for the various situations from Figs. 18 and 19, illustrating the relaxation to nontrivial (non-zero-intensity) equilibrium, i.e., “ \pm ”-fixed points [Figs. 18(a), 18(c), hollow black circles], limit cycle [Fig. 19(a), full blue circles], chaos [Fig. 19(b), full red romboids], and the relaxation to trivial (zero-intensity) equilibrium, i.e., 0 fixed point [Fig. 19(c), full green squares]. In the main text, we study the cases like the black curves, where time-independent stable configurations are seen. The symbols are data points from numerics and the lines are just to guide the eye.

which is expected to reach zero for a general relaxation process, where in the vicinity of an asymptotically stable fixed point $X \sim X_{\text{eq}} + xe^{-rt}$ will be generically nonzero for a limit cycle or chaos and will asymptote to a constant for the 0 point, where $X_{\text{eq}} = 0$, so we get $(1/X)dX/dt \sim r$. Figure 20 summarizes these possibilities. The black curves, corresponding to Figs. 18(a) and 18(c), show the situation which is in the focus of this work—the approach toward static equilibrium. The blue curve shows the limit cycle leading to periodic oscillations. The green curve corresponds to the chaotic regime with aperiodic dynamics and no relaxation, as in Fig. 19(b). Finally, the red curve corresponding to the pattern which radiates away in Fig. 19(b) reaches a constant value of r . In conclusion, the system shows roughly four classes of dynamics: fixed point, limit cycle, chaos, and incoherence. Our work only covers the first of the four, but the bifurcation diagrams in the previous subsection give a good hint of the part of the parameter space which contains them, facilitating experimental or numerical verification.

APPENDIX C: PERTURBATION THEORY AND STABILITY ANALYSIS

In this appendix, we develop the perturbation theory of the photorefractive beam system starting from the Lagrangian (4). The perturbation theory yields the criterion for the stability of the intensity patterns as they propagate along the z axis. Formally, it is just the perturbative diagrammatic calculation of the propagator. This calculation explicitly excludes topologically nontrivial patterns and thus is somewhat peripheral for our main goal, understanding the vortex dynamics. But the general ways by which an envelope Ψ_{\pm} can evolve along the z axis and

become unstable remain valid also for vortices. In particular, we will end up with a classification of geometrical symmetries of the intensity pattern $\Psi^\dagger\Psi$; the same symmetries are seen in vortex patterns and are an important guide for numerical and experimental work—how to recognize instabilities and also phases of the system.

Our system is strongly nonlinear, thus a naive perturbation theory about the trivial vacuum, i.e., constant beam intensity is out of question. The right way is to perturb about a nontrivial solution, which approximates a stable pattern. This means we treat the light intensity as constant in time z but nonconstant in space (x, y) . This is the hallmark of spatial dynamical solitons: They propagate with a constant profile along the z axis and to a good approximation do not interact with each other and do not radiate [3]. We thus write $\Psi = \Psi_0 + \delta\Psi$, giving $\Psi^\dagger\Psi = I_0 + F_0(\delta\Psi_+^\dagger + \delta\Psi_+) + B_0(\delta\Psi_-^\dagger + \delta\Psi_-) + \delta\Psi^\dagger\delta\Psi + O(|\delta\Psi|^2)$ with $F_0^2 + B_0^2 = I_0$. The lowest-order Lagrangian for Ψ_0 now reads

$$\begin{aligned} \mathcal{L}_0 = & \Psi_0^\dagger \Delta \Psi_0 + \Gamma I_0 - \Gamma(1 + \tau u + \tau E_0) \\ & \times \ln(1 + \tau u + I_x + I_0), \end{aligned} \quad (\text{C1})$$

which determines the shape of the solution $\Psi_0(x, y)$ in the first approximation. The dynamical term with $\partial_z\Psi$ drops out (it is proportional to the equation of motion for Ψ). Nontrivial propagation in time z is obtained from second-order expansion of the potential which is given in the next appendix in Eq. (D1) and we will not copy it here. Varying the quadratic expansion with respect to the fluctuation $\delta\Psi$ gives the linearized equation of motion for $\delta\Psi$:

$$\begin{aligned} [\pm i\sigma_3\partial_z - q^2 + \Gamma - (1 + \tau u + \tau E_0)]\delta\Psi^\mp \\ \mp \Gamma \frac{1 + \tau u + \tau E_0}{(1 + \tau u + I_x + I_0)^2} \delta\Psi^\pm = 0, \end{aligned} \quad (\text{C2})$$

where $\delta\Psi^+ \equiv \delta\Psi^\dagger, \delta\Psi^- \equiv \delta\Psi$. In homogenous spacetime (z, x, y) , we can transform to momentum space in both transverse and longitudinal directions. In the transverse plane, we get $(x, y) \mapsto (q_x, q_y)$ and $\Delta \mapsto -q^2$. The longitudinal coordinate or time z transforms as $z \mapsto k_n$ where $k_n = \pi n/L$, so the time maps to discrete frequencies. The reason is, of course, that its domain is finite, corresponding to the crystal length L .

Now we can derive the bare propagator (Green's function) of the fluctuating dynamical field $\delta\Psi$ by inserting the appropriate source $S(z)$ on the right-hand side of Eq. (C2). Normally, the source in the equation for the Green's function is just the Dirac δ function but the counterpropagating nature of our beams imposes a two-sided source:

$$S(z) = \begin{pmatrix} \delta(z) & 0 \\ 0 & \delta(z - L) \end{pmatrix}. \quad (\text{C3})$$

With this source (also Fourier-transformed in z), Eq. (C2) gives the bare propagator $G_{\alpha\beta}^{(0)}$ for the fields $\delta\Psi_{\alpha\beta}^\pm$:

$$\begin{aligned} G_{\alpha\beta}^{(0)}(k_n, q) = & [-ik_n S_{\alpha\gamma}(k_n) + A_{\alpha\delta}^* S_{\delta\gamma}(k_n) - B_{\alpha\delta} S_{\delta\gamma}(k_n)] \\ & \times [-k_n^2 + A_{\gamma\delta}^* A_{\delta\beta} - B_{\gamma\delta} B_{\delta\beta}^* + [A^*, B]_{\gamma\beta}]^{-1}. \end{aligned} \quad (\text{C4})$$

The auxiliary matrices A, B are defined as follows:

$$\begin{aligned} A_{\alpha\beta} = & i \begin{pmatrix} P_0 + P_1 - q^2 & P_0 \\ -P_0 & -P_0 - P_1 + q^2 \end{pmatrix}, \\ B_{\alpha\beta} = & i \begin{pmatrix} P_0 & P_0 \\ -P_0 & -P_0 \end{pmatrix}, \end{aligned} \quad (\text{C5})$$

where $P_1 = (1/4)I_0\Gamma(1 + \tau u + \tau E_0)/(1 + \tau u + I_x + I_0)^2$, $P_0 = \Gamma - \Gamma(1 + \tau u + \tau E_0)/(1 + \tau u + I_x + I_0)$, and $S(k_n) = \text{diag}(1, e^{ik_n L})$.

Now we have the basic ingredient of the perturbation theory: the bare propagator. The self-energy correction Σ of the propagator from the potential V_{eff} can be expanded in a power series over $\delta\Psi$, which gives an infinite tower of vertices. Simple combinatorial considerations give the expansion

$$\begin{aligned} \Sigma = & \sum_{j_1, j_2, j_3 \in \mathbb{N}} \frac{(-1)^{j_1+j_2+j_3} (j_1 + j_2 + j_3 - 1)!}{j_1! j_2! j_3!} \\ & \times \frac{\Gamma(1 + \tau u + \tau E_0)}{(1 + \tau u + I_0 + I_x)^{j_1+j_2+j_3+1}} (\Psi_0^\dagger \delta\Psi)^{j_1} \\ & \times (\Psi_0 \delta\Psi^\dagger)^{j_2} (\delta\Psi^\dagger \delta\Psi)^{j_3}, \end{aligned} \quad (\text{C6})$$

and the contraction over the internal indices of $\Psi^\pm, \delta\Psi^\pm$ is understood. Now we can formulate the diagrammatic rules. We have two kinds of propagators, $G^{(0)}$ and its Hermitian conjugate. The mean-field values Ψ_0^\pm are external sources. The term of order (j_1, j_2, j_3) contains $j_1 + j_3$ propagator lines $G^{(0)}$ (j_1 of them ending with the source Ψ_0) and $j_2 + j_3$ lines $(G^{(0)})^\dagger$ (j_2 of them ending with a source Ψ_0^\dagger); altogether, there are $j \equiv j_1 + j_2 + 2j_3$ lines. The expansion has to be truncated at some j . Since the mass dimension of Ψ is 1, the (j_1, j_2, j_3) diagram has the scaling dimension $2 - 2(j_1 + j_2 + 2j_3) < 0$, so *all* diagrams are irrelevant in the IR. This means we can make a truncation at small j .¹³ The leading terms are those where the order of the perturbation in $\delta\Psi^\pm$, which equals $j_1 + j_2 + 2j_3$, is the smallest. This gives two classes of diagrams, one with $j_1 = 1, j_2 = j_3 = 0$ and another with $j_2 = 0, j_1 = j_3 = 0$. They contain a single external source and introduce the wave-function renormalization, $G^{(0)} \mapsto ZG^{(0)}$, which does not influence the stability analysis. The four quadratic terms [with $(j_1, j_2, j_3) = (2, 0, 0), (0, 2, 0), (1, 1, 0), (0, 0, 1)$] introduce a mass operator. Only the terms $(1, 1, 0)$ and $(0, 0, 2)$ are trivial (noninteracting); the other two are interacting as they contain $(\delta\Psi^\pm)^2$ and require the calculation of an internal loop, giving the dressed propagator

$$G_{\alpha\beta}^{-1}(k_n, q) = [G^{(0)}(k_n, q)]_{\alpha\beta}^{-1} + (m^2)_{\alpha\beta}, \quad (\text{C7})$$

where the mass squared is a *positive* matrix, because the corresponding coefficients in (C6) have positive signs [from the term $(-1)^{j_1+j_2+j_3}$ with $j_1 + j_2 + j_3 = 2$] and the integral of the bare propagator is also positive. Explicitly, it reads

$$(m^2)_{\alpha\beta} = \frac{\Gamma(1 + \tau u + \tau E_0)}{(1 + \tau u + I_0 + I_x)^2} \sum_{k_n} \int_0^\infty dq q G_{\alpha\beta}^{(0)}(k_n, q), \quad (\text{C8})$$

¹³We do not worry about the UV divergences: We have an effective field theory and the UV cutoff is physical and finite.

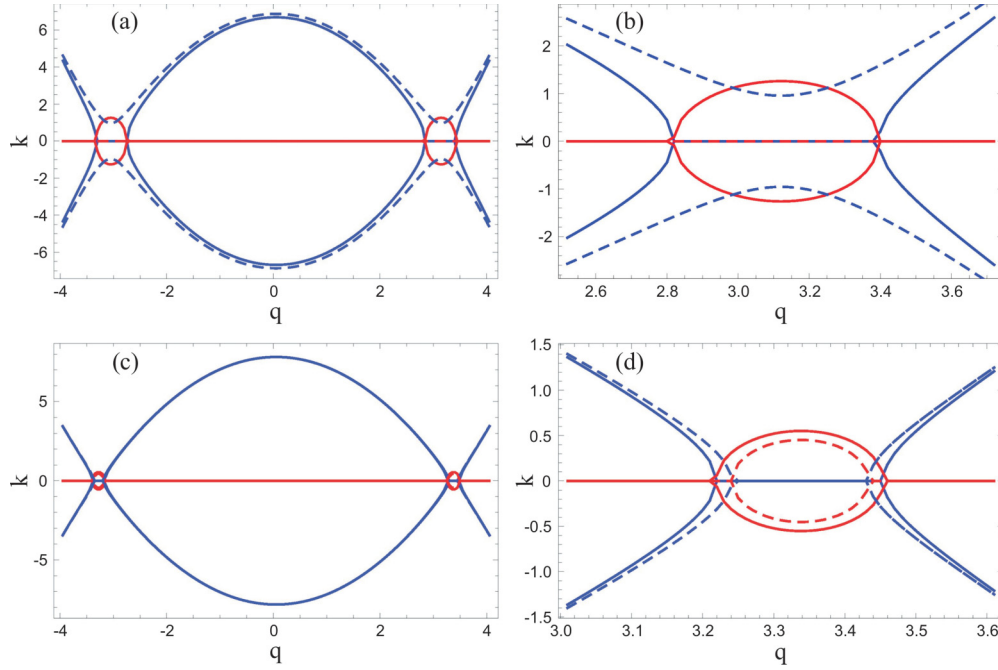


FIG. 21. Dispersion relation (position of the poles of the propagator) $k(q)$, where k is the continuous approximation of the discrete effective momentum $k_n = \pi nL$, for $I_x = 0$ [(a), (b)] and $I_x = 1$ [(c), (d)]. The plots (b) and (d) are enlargements of the plots (a) and (c). Blue lines denote $\text{Re}k$ and red lines show $\text{Im}k$. Notice that the propagator contains only k_n^2 and q^2 , so the pole has two copies with opposite signs and is either real or pure imaginary. Dashed lines are the corrected relations, with dressed propagator instead of the bare one. In the top panels, the region of instability, with $\text{Im}k_n \neq 0$, is cured by the nonlinear corrections, whereas in the bottom panels the instability remains. This generically happens at finite q and corresponds to the edge instability. Parameter values: $I_0 = 1$, $\Gamma = 15$, $L = 10$ mm.

where the discrete “frequency” k_n is summed in steps of π/L . Other than the mass renormalization, the dressed propagator has the same structure as the bare one. Now we will consider what this means for the stability of the patterns.

1. The pole structure, stability, and dispersion relations

Consider the poles of the propagator defined by the zeros of the eigenvalues of the matrix $G_{\alpha\beta}^{-1}(k_n, q)$. The stable solution corresponds to the situation where the perturbation $\delta\Psi_{\pm}$ dies out along z , so the stability of the solution is determined by the condition that the pole in q should have a nonpositive imaginary part, i.e., that a small perturbation decays. The denominator depends on k_n, q solely through k_n^2, q^2 ; it is linear in k_n^2 and quadratic in q^2 . Therefore, each of the two eigenvalues λ_{\pm} defines two pairs of opposite poles, $\pm q_{*+}, \pm q_{*-}, \pm q_{**+}, \pm q_{**-}$. Out of these, two pairs are positive for all parameter values, so no imaginary part can arise, and we have either two pairs of centrally symmetric imaginary poles, or one such pair, or none at all. We thus expect the sequence of symmetry-breaking transitions:

$$O(2) \longrightarrow \mathbb{C}_4 \longrightarrow \mathbb{C}_2. \quad (\text{C9})$$

Full circular symmetry is expected when there is no instability. With a single pair of unstable eigenvalues, we expect a square-like pattern with \mathbb{C}_4 symmetry, and with two pairs only a single reflection symmetry axis remains, yielding the group \mathbb{C}_2 . Only in the presence of disorder in the background lattice intensity pattern I_x can we expect the full breaking of the symmetry

group down to unity, but this is an *explicit* breaking and is not captured by this analysis.

The dispersion relation for a typical choice of parameter values is represented in Fig. 21, where we plot the location of the pole $k(q)$ in the continuous approximation (interpolating between the k_n values), with real parts of the pole in blue and imaginary in red. Since we have two pairs of opposite eigenvalues, the dispersion is P symmetric in x, y , and z (remember that time is really another spatial dimension), and any dispersion relation with a nonzero imaginary part will have a branch in the upper half-plane, i.e., an unstable branch. The only way out of instability is that the pole is purely real, i.e., infinitely sharp—this quasiparticle-like excitation signifies a solitonic solution. In Fig. 21, the dashed lines are drawn with the bare propagator $G^{(0)}$ and the full lines with the dressed propagator G , for the sets of parameter values. The perturbation always reduces the instability, i.e., the magnitude of the imaginary part of the poles—in Figs. 21(a) and 21(b) completely, resulting in zero imaginary part, and in Figs. 21(c) and 21(d) only partially. This reduction of instability likely explains the fact that linear stability analysis works extremely well for hyper-Gaussian beams (which have most power at small values of q), as found in Ref. [32].

The fact that the imaginary region always lies at finite q implies that the instability always starts at a finite scale, which corresponds to the behavior seen in the edge instability, which is shown, e.g., in Fig. 4. In order to understand the central instability, which starts from a single point, corresponding to $q \rightarrow \infty$, one needs to take into account also the higher order corrections from the potential (C6) which, as we discussed,

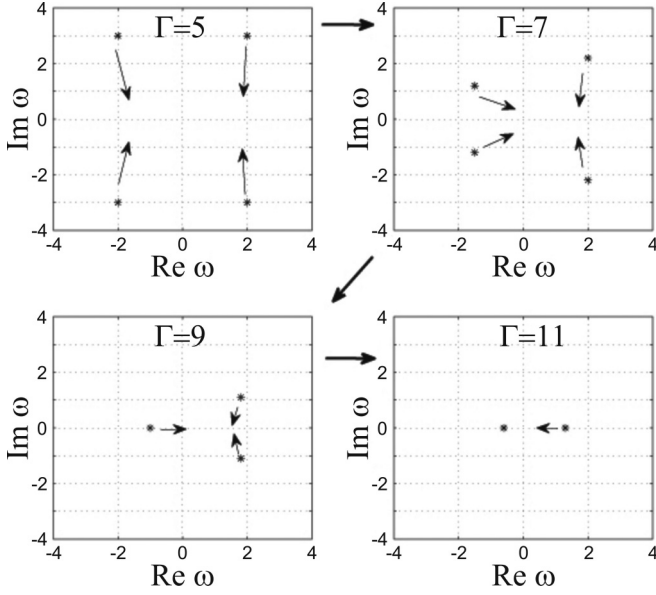


FIG. 22. The movement of the poles in the complex momentum plane in the case of central instability, for four different values of the PR coupling constant $\Gamma = 5, 7, 9, 11$. The complex momentum k is denoted by ω . Starting from the \mathbb{C}_4 -symmetric situation with two pairs of complex-conjugate poles, we first break the symmetry down to \mathbb{C}_2 and eventually lose all geometric symmetry as the two pairs merge into two real poles. Parameter values $I_0 = I_x = 1, L = 10$ mm.

diverge at $q \rightarrow \infty$. While we always have a natural UV cutoff, it may happen that the corrections become large (though finite) before that UV scale is reached. We postpone a detailed account for the subsequent publication, and content ourselves to give only the diagram of the movement of the poles in the complex plane. Higher order terms bring q -dependent corrections and break the inversion symmetry, resulting in the evolution of poles, as in Fig. 22. The instability corresponds to the situations where at least one pole has a positive imaginary part, i.e., the first three situations in the figure. The last pattern, with no symmetry at all and two real poles, is stable (but not asymptotically stable, as there is no pole with nonzero negative imaginary part).

The analysis performed here is obviously incomplete, and we have contented ourselves merely to give a sketch of how the instabilities considered in the main text arise, as well as to formulate a perturbation scheme which allows one to study such phenomena. Further work along the lines of Refs. [20,32,36] is possible by making use of our formalism, and we plan to address this topic in the future.

APPENDIX D: DERIVATION OF THE VORTEX HAMILTONIAN FROM THE MICROSCOPIC LAGRANGIAN

Starting from the vortex solution (6), we want to obtain an effective Hamiltonian for the vortex-vortex interaction. The task is to separate the kinetic term of the winding phase (with $\nabla\theta_{0\pm} = \sum_i Q_{\pm} \ln |\mathbf{r} - \mathbf{r}_i|$) from (i) the intensity fluctuations $\delta\psi_{\pm}$ about some background value $\psi_{0\pm}$ and (ii) the nonvortex phase fluctuations ($\delta\theta_{\pm}$) in (6). The first task requires us

to integrate out the amplitude fluctuations in the quadratic approximation. We first write $\Psi = \Psi_0 + \delta\Psi$ and expand the Lagrangian to quadratic order:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_2, \\ \mathcal{L}_0 &= \frac{1}{2} \partial_r \psi_0^\dagger \partial_r \psi_0 + \frac{I_0}{2r^2} |\nabla\theta_{0\alpha}|^2, \\ \mathcal{L}_2 &= \frac{1}{2} \partial_r \delta\Psi_\alpha^\dagger \partial_r \delta\Psi_\alpha \\ &\quad + \frac{1}{2r^2} \delta\Psi_\alpha^\dagger |\nabla\theta_{0\alpha}|^2 \delta\Psi_\alpha + V_{\text{eff}}(\delta\Psi_{\pm}), \\ V_{\text{eff}}(\delta\Psi^{\pm}) &= -\Gamma \delta\Psi^\dagger \delta\Psi - \Gamma \frac{1 + \tau u + \tau E_0}{2(1 + \tau u + I_x + I_0)^2} \\ &\quad \times [(\Psi_0 \delta\Psi^\dagger)^2 + (\Psi_0^\dagger \delta\Psi)^2 - 2(1 + \tau u + I_x + I_0) \\ &\quad \times \delta\Psi^\dagger \delta\Psi - (\Psi_0 \sigma_2 \delta\Psi^\dagger)(\Psi_0^\dagger \sigma_2 \delta\Psi)]. \end{aligned} \quad (\text{D1})$$

The zeroth-order (nonfluctuating) term \mathcal{L}_0 determines the intensity $I_0 = \psi_0^\dagger \psi_0$ and produces the kinetic term for the vortex phase θ_α , which gives just two decoupled copies of the conventional XY vortex gas. The quadratic part \mathcal{L}_2 becomes quite involved when we separate the amplitude $\delta\psi$ and the phase $\delta\theta$. Inserting (6) into (D1), one gets a quadratic action for $\delta\psi_\alpha$ and $\delta\theta_\alpha$. The rest gives a coupled quadratic action for the amplitude and phase fluctuations. Altogether, the Lagrangian is

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} (\delta\psi_+^2 + \delta\psi_-^2) + \frac{1}{2r^2} (\delta\psi_+^2 + \delta\psi_-^2) |\nabla\theta_\alpha|^2 \\ &\quad + \delta\psi_\alpha \hat{K}_{\alpha\beta} \delta\psi_\beta + (\delta\psi_\alpha^\dagger \psi_\alpha \nabla\theta_\alpha \nabla\delta\theta_\alpha + \text{H.c.}) + \dots, \end{aligned} \quad (\text{D2})$$

where (\dots) denote all terms of cubic or higher order in amplitude or phase fluctuations $\delta\psi_\alpha, \delta\theta_\alpha$, and we have left out the constant terms independent of all field values. Primes denote the derivatives with respect to r . The first term defines the intensity fluctuations through $\psi_\alpha(r)$, and the second term (transformed through partial integration) yields the aforementioned conventional Coulomb gas of vortices after inserting the vortex solution from (6) for θ_α . The third term has the meaning of stiffness or mass matrix for intensity fluctuations and the last term gives rise to the coupling between the flavors, upon integrating out $\delta\psi$. The matrix \hat{K} is

$$\begin{aligned} \hat{K} &= \frac{1}{(b + \frac{3}{2}I_0)(b - \frac{1}{2}I_0)} \begin{pmatrix} b + \frac{I_0}{2} & I_0 \\ I_0 & b + \frac{I_0}{2} \end{pmatrix}, \\ b &= \Gamma \frac{1 + \tau E_0}{2(1 + \tau E_0)^2} (2 + 2I_x + 3I_0). \end{aligned} \quad (\text{D3})$$

The action is quadratic in $\delta\psi$; therefore, we know how to integrate it out and obtain an effective action depending only on phase fluctuations. To do that, we need to solve the eigenvalue equation for $\delta\psi$ obtained from (D2), which reads

$$\partial_{rr} \delta\psi_\alpha - K_{\alpha\beta} \delta\psi_\beta = \left(\frac{|\nabla\theta|^2}{2r^2} + \lambda_{\pm} \right) \delta\psi_\alpha, \quad (\text{D4})$$

and is solved by diagonalizing the system $\delta\psi_\alpha \mapsto \delta\chi_\alpha = U_{\alpha\beta}\delta\psi_\beta$ and reducing it to the Bessel equation:

$$\delta\chi_\pm(r) = \sqrt{r}\{C_\pm J_\nu[\sqrt{2(K_{11} \mp K_{12} - \lambda_\pm)}r] + D_\pm Y_\nu[\sqrt{2(K_{11} \mp K_{12} - \lambda_\pm)}r]\}, \quad (\text{D5})$$

with $\nu = \sqrt{1/4 + |\nabla\theta|^2}$, where J_ν, Y_ν are Bessel functions of the first and second kinds, respectively. For well-defined behavior close to the vortex core (for $r \sim a$), we have $D_\pm = 0$, and C_\pm are arbitrary as the equation is linear. The eigenvalues $\lambda_{n\pm}, n = 0, 1, \dots$ are obtained by requiring that the fluctuation decays to zero at the crystal edge $r = \Lambda$:

$$\sqrt{2K_{11} \mp K_{12} - \lambda_{n\pm}}\Lambda = j_n(\nu), \quad (\text{D6})$$

where j_n is the n th zero of the Bessel function J_ν . The values $\lambda_{n\pm}$ are impossible to express analytically in closed form; however, it is not necessary for our purposes as $\Lambda \gg a$. The functional determinant obtained after integrating out χ_\pm is now expressed in terms of the eigenvalues:

$$\begin{aligned} \mathcal{K}_{\alpha\beta} &= \ln \left(\prod_n \lambda_{n\alpha} \lambda_{n\beta} \right)^{-1/2} \\ &= -\frac{1}{2} \sum_n \ln \left(K'_\alpha + K'_\beta + \frac{2j_n(\nu)^2}{\Lambda^2} \right) \sim \\ &\quad -\frac{\Lambda}{2} (K'_\alpha + K'_\beta) + O\left(\frac{1}{\Lambda}\right), \end{aligned} \quad (\text{D7})$$

where $K'_\pm = 2K_{11} \mp K_{12}$. Now we are left with a solely phase-dependent quadratic Lagrangian:

$$\mathcal{L} = \psi_{0\alpha} \nabla\theta_\alpha \nabla\delta\theta_\alpha \hat{\mathcal{K}}_{\alpha\beta} \psi_{0\beta} \nabla\theta_\beta \nabla\delta\theta_\beta + \frac{I_0}{2r^2} |\nabla\theta_\alpha|^2. \quad (\text{D8})$$

The final task is to integrate out the phase fluctuations, which is a trivial Gaussian integration, yielding

$$\mathcal{L} = \frac{I_0}{2r^2} |\nabla\theta_\alpha|^2 + I_0 \nabla\theta_\alpha \mathcal{K}_{\alpha\beta}^{-1} \nabla\theta_\beta. \quad (\text{D9})$$

The resulting Lagrangian now depends only on the vortexing phases θ_α . The first term is carried from the original Lagrangian, and it does not mix the flavors. But the second term, stemming from the fluctuations, has nonzero mixed \pm cross terms. The quadratic derivative terms can be transformed by partial integration to the familiar Coulomb gas form of the XY model, with the same-flavor coupling which is already present in absence of fluctuations, *and* the coupling between the vortices of different flavors. Thus the existence of two beams together with the fact that amplitude and phase fluctuations do not decouple give us a richer system, with interaction between two vortex flavors. For future use, it is more convenient to look at the vortex Hamiltonian $\mathcal{H}_{\text{vort}}$ —the difference from the Lagrangian lies just in the sign of the term V_{eff} . This finally yields the Hamiltonian [for Eq. (7), repeated here for convenience]:

$$\begin{aligned} \mathcal{H}_{\text{vort}} &= \sum_{i < j} (g \vec{Q}_i \cdot \vec{Q}_j + g' \vec{Q}_i \times \vec{Q}_j) \ln r_{ij} \\ &\quad + \sum_i (g_0 \vec{Q}_i \cdot \vec{Q}_i + g_1 \vec{Q}_i \times \vec{Q}_i), \end{aligned} \quad (\text{D10})$$

with $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$, and the indices $1 \leq i, j \leq N$ sum over all the vortices. The coupling constants g, g', g_0, g_1 are the result of integrating out the intensity fluctuations and in general are given by rather cumbersome (and not very illustrative) functions of Γ, I_0, τ . We give the expressions at leading order just for comparison with numerics:

$$\begin{aligned} g &= I_0 + \frac{4b + 2I_0}{(2b + 3I_0)(2b - I_0)}, \\ g' &= \frac{4I_0}{(2b + 3I_0)(2b - I_0)}, \\ b &= \Gamma \frac{1 + \frac{\tau}{L} - \tau \frac{I_0 + I_x}{1 + I_0 + I_x}}{2(1 + \frac{\tau}{L} - \tau \frac{I_0 + I_x}{1 + I_0 + I_x})^2} \left(2 + 2\frac{\tau}{L} + 2I_x + 3I_0 \right). \end{aligned} \quad (\text{D11})$$

These expressions are used later to redraw the phase diagram in the space of physical parameters Γ, I_0, I_x, L .

APPENDIX E: MULTIVORTEX MEAN-FIELD THEORY

For a mean-field treatment of a system with multiple vortices, we start from the Hamiltonian (7) and introduce the order parameter fields in the following way. Denote the number of vortices with charge (1, 1) by ρ_{2+} and the number of vortices (1, -1) by ρ_{2-} ; due to charge conservation, this means we also have ρ_{2+} vortices of type (-1, -1) and ρ_{2-} vortices with charge (-1, 1). The number of single-charge vortices of type (1, 0) and (0, 1) is denoted by ρ_{1+} and ρ_{1-} , respectively. Denote also $\rho_2 \equiv \rho_{2+} + \rho_{2-}$ and $\delta\rho_2 \equiv \rho_{2+} - \rho_{2-}$ (notice that $-\rho_2 \leq \delta\rho_2 \leq \rho_2$), and finally $\rho_1 \equiv \rho_{1+} + \rho_{1-}$. We insert this into the vortex Hamiltonian $\mathcal{H}_{\text{vort}}$ and assume that the long-ranged logarithmic interaction $\ln r_{ij}$ justifies the mean-field approximation: For $i \neq j$, we can approximate $\ln r_{ij} \sim \ln \Lambda$, assuming that average intervortex distance is of the same order of magnitude as the system size. For the core energy, we know that $g_0, g_1 \sim \ln(a/\epsilon) \sim -\ln\epsilon \sim \ln \Lambda$, where in the last equality we have assumed that the UV cutoff ϵ is of similar order of magnitude as the inverse of the IR cutoff $1/\Lambda$, which is natural.¹⁴ Thus all terms are proportional to $L \ln \Lambda$ and we can write

$$\begin{aligned} \mathcal{F}_{\text{mf}} &= \beta \ln \frac{\Lambda}{a} [2(g - 1)\rho_2 + 2g'\delta\rho_2 + (g - 1)\rho_1] \\ &\equiv A\rho_2 + B\delta\rho_2 + \frac{B}{2}\rho_1. \end{aligned} \quad (\text{E1})$$

We use the notation $\beta \equiv L$ to emphasize the analogy with the free energy of spin vortices, where β is the inverse temperature. The analogy is purely formal as our system is not subject to thermal noise. Now the ground state is determined by minimizing the free energy, i.e., the effective action of the system. Notice that \mathcal{F}_{mf} is linear in the fields $\rho_2, \delta\rho_2, \rho_1$ so the optimal configurations have either $\mathcal{F}_{\text{mf}} = 0$ or $\mathcal{F}_{\text{mf}} \rightarrow -\infty$, and the mean-field densities $\rho_{1,2}$ are either zero or arbitrary

¹⁴Nevertheless, this is clearly not a rigorous argument. Our mean-field calculation is somewhat sketchy and merely assumes that the long-range interactions can safely be modeled as a uniform vortex charge field.

(formally infinite). This is a well-known property of the 2D Coulomb gas and has to do with the fact that (assuming the cutoff dependence has been eliminated) this system is conformal invariant in the insulator phase, so all finite densities ρ are equivalent: There is no other scale to compare ρ to. Likewise, the prefactor β can be absorbed into the definition of the coupling constants g, g' and thus is not an independent parameter (this is well known also from the single-flavor case). Minimizing (E1) is an elementary exercise and we find again four regimes, corresponding to the four phases we guessed based on the single-vortex free energy $\mathcal{F}^{(1)}$:

(1) For $A > 0, A > |B|$, the minimum is reached for $\rho_2 = \delta\rho_2 = \rho_1 = 0$. In the ground state, there are no vortices at all—the system is a vortex insulator.

(2) For $B > 0, A + B < 0$, giving $g + g' < 1, g' > 0$, the free energy has its minimum for $\rho_2 > 0$ and $\delta\rho_2 = -\rho_2$ (notice that $-\rho_2 \leq \delta\rho_2 \leq \rho_2$). This means $\rho_{2+} = 0, \rho_{2-} > 0$, so opposite-charged vortices $(Q, -Q)$ proliferate, and the system is dominated by the interactions between the charges. This is the frustrated vortex insulator regime. Since $g' < 0$, the single-charge vortices (density ρ_1) are suppressed.

(3) For $B < 0, A + B < 0$, i.e., $g + g' < 1, g' < 0$, the minimum is reached for $\rho_2 = \delta\rho_2 > 0$, i.e., $\rho_{2-} = 0$, so the vortices (Q, Q) can proliferate. However, since $g' < 0$, there is also nonzero single-flavor density ρ_1 and the proliferation of vortices $(Q, 0)$ and $(0, Q)$ which generically dominate over two-flavor vortices. This is the conductor phase, with mostly single-flavor vortices (as in the standard XY model).

(4) The point $A = B = 0$ is special: Naively, from (E1), arbitrary nonzero $\rho_1, \rho_2, \delta\rho_2$ are allowed. Of course, higher order corrections will change, this but the energy cost of vortex formation will generically be smaller than in previous phases. This is the vortex perfect conductor phase. In the mean-field approach, it looks like a single point, but that will turn out to be an artifact of the mean-field approach: For small nonzero A, B the system still remains in this phase.

In terms of the original parameters g, g' , one sees the insulator phase is given by $g + g' > 1$, and the conductor and the FI are separated by the line $g' = 0$. We can now sketch the phase diagram, which is given in Fig. 2(a), side by side with the more rigorous diagram obtained by the RG flow, in Sec. III B 2.

APPENDIX F: COUNTERPROPAGATING BOUNDARY CONDITIONS

In the derivation of the vortex Hamiltonian and its RG analysis, we have pulled under the rug the treatment of the CP boundary conditions: The effective Hamiltonian (and consequently the partition function and the phase diagram) depends solely on the bulk configuration, and nowhere can one see the fact that $\Psi_+(z = 0; \mathbf{r}, t)$ and $\Psi_-(z = L; \mathbf{r}, t)$ are fixed. Now we will explicitly show that these boundary conditions are irrelevant in the RG sense; i.e., they contribute additional, boundary terms to the effective Hamiltonian, but these terms do not change the fixed points to which the solution flows.

The full Hamiltonian with correct CP boundary conditions is obtained by adding the F source at $z = 0$ and the B source at $z = L$ to the Lagrangian \mathcal{L} from (4) or, equivalently, to the equations of motion. The sources impose the conditions $F(z =$

$0; x, y, t) = F_0(x, y)$ and $B(z = L; x, y, t) = B_0(x, y)$ so they equal

$$J_+ = F_0(x, y)\delta(z), \quad J_- = B_0(x, y)\delta(z - L), \quad (\text{F1})$$

and the full Lagrangian is

$$\mathcal{L}_{\text{CP}} = \mathcal{L} + J_+\Psi_+ + J_-\Psi_- \mapsto \mathcal{L} + F_0(x, y)\Psi_+(z; x, y; t) + e^{ikz}B_0(x, y)\Psi_-(z; x, y; t). \quad (\text{F2})$$

Unlike the Dirac δ source (C3) for the Green's function, now the source has nontrivial dependence on transverse coordinates. Now we can insert the vortex solution (6) in both Ψ_{\pm} and F_0, B_0 and repeat the steps from the subsequent derivation. The vortex charges in F_0 can be denoted by $\vec{P}_{i'}^{(+)} \equiv (P_{i'+}, 0)$ and $\vec{P}_{i'}^{(-)} \equiv (0, P_{i'-})$; by definition, the $+$ component of B_0 as well as the $-$ component of F_0 are zero and thus carry no vorticity. The primed indices refer to the vortices in the input beams, and the nonprimed, like before, to the bulk vortices. Notice the source term changes sign upon performing the Legendre transform, appearing as $-J_+\Psi_+ - J_-\Psi_-$ in the Hamiltonian.

Now we will check if the RG flow of the Hamiltonian with boundary terms is affected by the sources. In the notation introduced above, the total vortex Hamiltonian is

$$\begin{aligned} \mathcal{H}_{\text{CP}} = & \sum_{i,j} (g\vec{Q}_i \cdot \vec{Q}_j + g'\vec{Q}_i \times \vec{Q}_j) \ln r_{ij} \\ & + \sum_{i',j} \delta(z) P_{i'+} (gQ_{j+} + g'Q_{j-}) \ln r_{i'j} \\ & + \sum_{i',j'} \delta(z - L) P_{j'-} (gQ_{i-} + g'Q_{i+}) \ln r_{i'j'}. \end{aligned} \quad (\text{F3})$$

Notice there is no source-source interaction: Same-flavor interaction cannot exist as $P_{j'-} = P_{j'+} = 0$, and cross-flavor interaction does not exist as J_+ and J_- exist at different z values, i.e., the cross term would be proportional to $\delta(z)\delta(z - L)$ and thus vanishes. The presence of sources breaks the spatial homogeneity, complicating the traces (integrals over the positions of virtual vortex-antivortex pairs), but does not change the main line of the calculation. The fluctuation of the partition function due to vortex pair creation is now

$$\begin{aligned} \frac{\delta\mathcal{Z}}{\mathcal{Z}} = & 1 + \frac{y^4}{4} \sum_{\vec{q}} \int d^2r \\ & \times \int d^2r_{12} e^{-C(\vec{q}, \mathbf{r}_1; -\vec{q}, \mathbf{r}_2) - \sum_j [D_j^+(\vec{q}, \mathbf{r}_1) - D_j^-(\vec{q}, \mathbf{r}_2)]} \\ & \times [e^{C(\vec{Q}_1, \mathbf{R}_1; \vec{q}, \mathbf{r}_1) + C(\vec{Q}_1, \mathbf{R}_1; -\vec{q}, \mathbf{r}_2) + C(\vec{Q}_2, \mathbf{R}_2; \vec{q}, \mathbf{r}_1) + C(\vec{Q}_2, \mathbf{R}_2; -\vec{q}, \mathbf{r}_2)} \\ & - 1]. \end{aligned} \quad (\text{F4})$$

We have denoted $C(\vec{Q}_1, \mathbf{R}_1; \vec{Q}_2, \mathbf{R}) \equiv (g\vec{Q}_1 \cdot \vec{Q}_2 + g'\vec{Q}_1 \times \vec{Q}_2) \ln R_{12}$, and the coupling to the sources is encapsulated in the function

$$D_j^{\pm}(\vec{q}, \mathbf{r}) \equiv \delta(z - z_{\pm})(g_{\pm}\vec{q} \cdot \vec{P}_{j'} \mp g'_{\pm}\vec{q} \times \vec{P}_{j'}) \ln |\mathbf{r} - \mathbf{r}_{j\pm}|, \quad (\text{F5})$$

with $z_+ = 0, z_- = L$. The coupling constants g_{\pm}, g'_{\pm} are obtained from g, g' in (F4) by replacing

$$I_0 \mapsto \sqrt{I_0 I_{\pm}}, \quad (\text{F6})$$

with $I_+ = |F_0|, I_- = |B_0|$. Now the exponential of the extra term with sources also needs to be expanded in r_{12} , as the combinations $2\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2, \mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$ do not decouple anymore and exact integration is impossible. After writing $D_{j'}^{\pm}(\vec{q}, \mathbf{r}_{1,2}) = D_{j'}^{\pm}(\vec{q}, \mathbf{r}) \pm \mathbf{r}_{12} \cdot \nabla D_{j'}^{\pm}(\vec{q}, \mathbf{r}) + \dots$ and similarly for C , we notice first that the zeroth-order terms from $D_{j'\pm}$ cancel out: $D_{j'}^{\pm}(\vec{q}, \mathbf{r}) + D_{j'}^{\pm}(-\vec{q}, \mathbf{r}) = 0$. Then we get (at quadratic order in r_{12} in the integrand):

$$\begin{aligned} \frac{\delta \mathcal{Z}}{\mathcal{Z}} &= 1 + \frac{y^4}{4} \sum_{j'} \int d^2 r \int d^2 r_{12} e^{-C(\vec{q}, \mathbf{r}_{12}; -\vec{q}, \mathbf{0})} \\ &\times \left\{ \mathbf{r}_{12} \cdot [\nabla C(\vec{Q}_1, \mathbf{R}; \vec{q}, \mathbf{r}) + \nabla C(\vec{Q}_2, \mathbf{R}; -\vec{q}, \mathbf{r})] \right. \\ &+ \frac{r_{12}^2}{2} \left| \nabla C(\vec{Q}_1, \mathbf{R}; \vec{q}, \mathbf{r}) + \nabla C(\vec{Q}_2, \mathbf{R}; -\vec{q}, \mathbf{r}) \right|^2 \left. \right\} \\ &\times [\mathbf{r}_{12} \cdot [1 - \nabla D_{j'}^+(\vec{q}, \mathbf{r}) + \nabla D_{j'}^-(\vec{q}, \mathbf{r})] \\ &+ \{\mathbf{r}_{12} \cdot [\nabla D_{j'}^+(\vec{q}, \mathbf{r}) + \nabla D_{j'}^-(\vec{q}, \mathbf{r})]\}^2] + \dots \\ &\equiv 1 + \frac{y^4}{4} [I_{10} + I_{20} - I_{11} + O(r_{12}^3)]. \quad (\text{F7}) \end{aligned}$$

The integral I_{mn} is the term with the contribution of order r_{12}^m from the second line in the integrand and with the contribution of order r_{12}^n from the third line. The integrands in I_{mn} are thus of order $m+n$ in r_{12} , m coming from the expansion of $D_{j'}^{\pm}$ and n from the expansion of C . By homogeneity, $I_{01} = 0$ and I_{02} is the same integral that appears in absence of sources, whose calculation was used in obtaining (13) and which gives the right-hand side of the RG flow (14). The remaining integral I_{11} is the new ingredient, and the only one which depends on the sources. Representing it as

$$\begin{aligned} I_{11} &= \frac{\pi^2}{4} \sum_{j'} \sum_{\alpha=\pm} \sum_{\sigma=1,2} \delta(z - z_{\alpha}) \\ &\times (g_{\alpha} \vec{Q}_{\sigma} \cdot \vec{P}_{j'\alpha} + g'_{\alpha} \vec{Q}_{\sigma} \times \vec{P}_{j'\alpha}) \\ &\times \nabla \frac{1}{|\mathbf{R}_{12}|} \cdot \nabla \frac{1}{|\mathbf{r}_{j'\alpha}|} \tilde{I}_{j'\alpha}, \quad (\text{F8}) \end{aligned}$$

we compute the integral $\tilde{I}_{j'\alpha}$ in polar coordinates:

$$\tilde{I}_{j'\alpha} = \frac{1}{2} \int_0^{2\pi} d\theta_j \ln(r_{12}^2 - 2r_{12}r_{j'\alpha} \cos \theta_j + r_{j'\alpha}^2) \Big|_{\Lambda_1}^{\Lambda_2}, \quad (\text{F9})$$

where $\theta_{j'\alpha}$ is the angle between $\mathbf{r}_{j'\alpha}$ and \mathbf{r}_{12} . Assuming the RG scale changes as $\Lambda_1 = \Lambda, \Lambda_2 = \Lambda(1 + \ell)$, for small ℓ we can expand the integrand, getting

$$\begin{aligned} \tilde{I}_{j'\alpha} &= \int_0^{2\pi} d\theta_j \frac{\Lambda^2 - \Lambda r_{j'\alpha} \cos \theta_j}{\Lambda^2 - 2\Lambda r_{j'\alpha} \cos \theta_j + r_{j'\alpha}^2} \ell + O(\ell^2) \\ &= 2\pi \ell + O(\ell^2). \quad (\text{F10}) \end{aligned}$$

The complicated dependence on the positions of the sources disappears completely in the first order in ℓ .¹⁵ Altogether, by comparing the outcome of (F7) to the original Hamiltonian (F3), we see that the renormalization of the bulk interaction between \vec{Q}_1 and \vec{Q}_2 is unaffected by the sources, given as before by the I_{02} term, and the source-bulk coupling renormalizes with a strictly negative shift (as $\tilde{I}_{j'\alpha} = 2\pi > 0$). The flow equations for g, g' couplings are unchanged, being the same as in (14). The bulk-to-boundary couplings g_{\pm}, g'_{\pm} have the flow equations

$$\frac{\partial g_{\pm}}{\partial \ell} = -\pi^3 N \ell, \quad \frac{\partial g'_{\pm}}{\partial \ell} = -\frac{\pi^3}{2} N \ell, \quad (\text{F11})$$

where $N = \sum_{j'} \sum_{\alpha} 1$ is the total vorticity of the sources. This obviously flows to $g_{\pm}, g'_{\pm} = 0$.

Intuitively, one may wonder how come such an important thing as the CP geometry has no bearing on the vortex dynamics; surely the behavior of a copropagating system would be expected to differ from a counterpropagating system. The answer is that the CP geometry does enter our calculations—the B beam has an extra minus sign in the equations of motion (1) (alternatively, in the Lagrangian in Eq. (4)); equivalently, the symmetry group of the effective potential in the Lagrangian is $SU(1, 1)$, not $SU(2)$ as it would be for two copropagating beams. Finally, let us emphasize again that in the numerical simulations we directly solve the propagation equations (1) together with (2); i.e., we directly take into account the CP boundary conditions—no analytical approximations whatsoever are used in the numerics, and no use is made of the effective vortex Hamiltonian.

APPENDIX G: ORDER PARAMETERS AND RG ANALYSIS OF THE CP VORTICES IN THE PRESENCE OF DISORDER

1. Saddle-point solutions

We start by rewriting the replicated partition function $\tilde{\mathcal{Z}}^n$ in terms of $p_{\alpha}, q_{\alpha\beta}$ and inserting the constraints which encapsulate their definition in Eq. (19):

$$1 \mapsto \int \mathcal{D}[\lambda_{(\mu)}^{\alpha}] \exp \left[\lambda_{(\mu)}^{\alpha} \left(p_{\alpha}^{(\mu)} - \frac{1}{N} \sum_{i=1}^N \mathcal{Q}_{i\alpha}^{(\mu)} \right) \right], \quad (\text{G1})$$

$$1 \mapsto \int \mathcal{D}[\lambda_{(\mu\nu)}^{\alpha\beta}] \exp \left[\lambda_{(\mu\nu)}^{\alpha\beta} \left(q_{\alpha\beta}^{(\mu\nu)} - \frac{1}{N} \sum_{i,j=1}^N \mathcal{Q}_{i\alpha}^{(\mu)} \mathcal{Q}_{j\beta}^{(\nu)} \right) \right]. \quad (\text{G2})$$

We have five constraints, $\lambda_{(\mu\nu)}^{++}, \lambda_{(\mu\nu)}^{--}, \lambda_{(\mu\nu)}^{+-} = \lambda_{(\mu\nu)}^{-+}, \lambda_{(\mu)}^{+}, \lambda_{(\mu)}^{-}$, for the corresponding five order parameters in (19). We can denote

$$\hat{K} \equiv \begin{pmatrix} \lambda_{(\mu\nu)}^{++} & \lambda_{(\mu\nu)}^{+-} \\ \lambda_{(\mu\nu)}^{+-} & \lambda_{(\mu\nu)}^{--} \end{pmatrix}, \quad \vec{\lambda} \equiv \begin{pmatrix} \lambda_{(\mu)}^{+} \\ \lambda_{(\mu)}^{-} \end{pmatrix}. \quad (\text{G3})$$

We will also sometimes leave out the replica indices μ, ν to avoid cramming the notation too much. Now we can first integrate out the vortex degrees of freedom $\mathcal{Q}_{i\alpha}^{(\mu)}$ from (18)

¹⁵The additional assumption is that $\Lambda > r_{j'\alpha}$ so the integrand contains no poles; this is clearly justified as Λ is the length cutoff.

to get the effective action

$$S_{\text{eff}} = -\frac{\beta^2}{4} \sum_{\mu, \nu=1}^n [\sigma_{++}^2 (q_{++}^{(\mu\nu)})^2 + 2\sigma_{+-}^2 (q_{+-}^{(\mu\nu)})^2 + \sigma_{--}^2 (q_{--}^{(\mu\nu)})^2] - \beta \sum_{\mu=1}^n [J_0^+ (p_+^{(\mu)})^2 + 2J_0^{+-} p_+^{(\mu)} p_-^{(\mu)} + J_0^- (p_-^{(\mu)})^2] + \frac{1}{2} \ln \det \hat{K} - \frac{1}{4} \vec{\lambda} \hat{K}^{-1} \vec{\lambda} - \sum_{\mu, \nu=1}^n (\lambda_{(\mu\nu)}^{++} q_{++}^{(\mu\nu)} + \lambda_{(\mu\nu)}^{+-} q_{+-}^{(\mu\nu)} + \lambda_{(\mu\nu)}^{--} q_{--}^{(\mu\nu)}) - \sum_{\mu=1}^n \vec{\lambda}_{(\mu)} \cdot \vec{p}^{(\mu)}. \quad (\text{G4})$$

The saddle-point equations for the constraints give the constraints in terms of the expectation values $q^{(\mu\nu)}, p^{(\mu)}$. Luckily, the equation for $\vec{\lambda}$ is easy:

$$\frac{\partial S_{\text{eff}}}{\partial \vec{\lambda}} = \hat{K}^{-1} \vec{\lambda} - \vec{p} = 0, \quad (\text{G5})$$

so we immediately solve $\vec{\lambda} = \hat{K} \vec{p}$. Now plugging this into the equations for the three remaining constraints yields

$$\frac{\partial S_{\text{eff}}}{\partial \lambda_{\pm\pm}} = \frac{1}{2} \frac{X \lambda_{\pm\pm}^{-1}}{X^2 - Y^2} - q_{\pm\pm} + \frac{1}{4} (p_{\pm})^2 = 0, \quad (\text{G6})$$

$$\frac{\partial S_{\text{eff}}}{\partial \lambda_{+-}} = \frac{Y \lambda_{+-}^{-1}}{X^2 - Y^2} - q_{+-} + \frac{1}{2} p_+ p_- = 0. \quad (\text{G7})$$

We have denoted $X = \det \lambda_{++} = \det \lambda_{--}, Y = \det \lambda_{+-}$ (these have a well-defined limit for $n \rightarrow 0$). It is trivial to write $\lambda_{\pm\pm}, \lambda_{+-}$ from the above expressions, and we can feed the solutions for all the constraints into the effective action and then solve the saddle-point equations for the order parameters $p_{\pm}, q_{++}, q_{--}, q_{+-}$. Full equations are too complex to be solved, even approximately. We will simplify the problem with the following reasoning. The sums over single-replica order parameters generically scale as $\sum_{\mu} p_{\pm}^{(\mu)} \sim \sum_{\mu} (p_{\pm}^{(\mu)})^2 \sim n$, whereas the double-replica parameters have $\sum_{\mu, \nu} q_{\alpha\beta}^{(\mu\nu)} \sim n^2$. This means that in the limit $n \rightarrow 0$, the p_{\pm} terms dominate over $q_{\alpha\beta}$ terms. Therefore, if $p_{\pm} \neq 0$ we can disregard the quantities $q_{\alpha\beta}$ or expand in a series over them, simplifying the equations significantly. Only if the replica symmetry breaking imposes $p_{\pm} = 0$ (not every replica-symmetry-breaking configuration does so) are the $q_{\alpha\beta}$ order parameters significant, and the saddle-point equations with $p_{\pm} = 0$ are again approachable.

Consider first the case $p_{\pm} = 0$. After some algebra, the effective action is now

$$S_{\text{eff}} = -\frac{\beta^2}{4} \sum_{\mu, \nu=1}^n [\sigma_{++}^2 (q_{++}^{(\mu\nu)})^2 + 2\sigma_{+-}^2 (q_{+-}^{(\mu\nu)})^2 + \sigma_{--}^2 (q_{--}^{(\mu\nu)})^2] + \frac{1}{2} \ln(X^2 |q^{++}|^{-1} \cdot |q^{--}|^{-1} - 4Y^2 |q^{+-}|^{-2}). \quad (\text{G8})$$

Consider first the ansatz when the $q^{\pm\pm}$ fields are nonzero, whereas the mixed-flavor field q^{+-} is zero. In this case, the second term in (G8), coming from the determinant \hat{K} , simplifies further and we get the saddle-point equation

$$-\frac{\beta^2}{2} \sigma_{\pm\pm}^2 q^{\pm\pm} - \frac{1}{2} (q^{\pm\pm})^{-1} = 0, \quad (\text{G9})$$

which is the same as for the infinite-range spin-glass Ising model [42,54]. One obvious solution is $q^{\pm\pm} = q^{+-} = 0$, the completely disordered system with no vortex proliferation—the familiar insulator phase. It is easy to check that this is indeed a minimum of the effective action S_{eff} . There is also a replica-symmetric but nontrivial solution

$$q_{(\mu\nu)}^{\pm\pm} = Q_0^{\pm\pm} + (1 - Q_0^{\pm\pm}) \delta_{\mu\nu}, \quad (\text{G10})$$

which yields the solution $Q_0^{\pm\pm} = 1 - 1/(\beta\sigma^{\pm\pm})$. However, this solution is unstable and is not observable. A stable nontrivial solution is obtained if the replica symmetry is broken. The ansatz is well known from the spin-glass literature (e.g., Ref. [42]) and has a $\rho \times \rho$ matrix $Q^{\pm\pm}$ on the block-diagonal and the constant zero elsewhere, with

$$\hat{Q}^{\pm\pm} = Q_1^{\pm\pm} + (1 - Q_1^{\pm\pm}) \delta_{\mu\nu}, \quad \mu, \nu = 1, \dots, \rho. \quad (\text{G11})$$

Equation (G9) suggests that $Q_1^{\pm\pm} > 0$ for sufficiently large β , i.e., small L . However, no analytical solution for the elements $Q_1^{\pm\pm}$ exists and they have to be solved for numerically, by plugging in the solution into the effective action and minimizing it. This is an easy task (for chosen values of the parameters and disorder statistics) but we will not do it here as we do not aim at quantitative accuracy anyway; we merely want to sketch the phase diagram. Now if the third field q^{+-} is nonzero, it satisfies the same equation as (G9) just with $\sigma_{++}^2 \mapsto 2\sigma_{+-}^2$. The three combinations of nonzero order parameters correspond to the three familiar phases: $q^{\pm\pm} \neq 0$ is the conductor, $q^{+-} \neq 0$ is the frustrated insulator, and $q^{\pm\pm}, q^{+-} \neq 0$ is the perfect conductor.

The solutions with $\vec{p} \neq 0$ yield new physics. In this case, we have at leading order $\lambda_{\pm\pm} = -2X/(X^2 - Y^2)(p_{\pm})^{-2}, \lambda_{+-} = -2Y/(X^2 - Y^2)(p_+ p_-)^{-1}$, so the effective action is

$$S_{\text{eff}} = -\beta \sum_{\mu=1}^n [J_0^+ (p_{(\mu)}^+)^2 + 2J_0^{+-} p_{(\mu)}^+ p_{(\mu)}^- + J_0^- (p_{(\mu)}^-)^2] - \ln p^+ p^- + O(|q^{\alpha\beta}|^2) + O\left(\frac{|q^{\alpha\beta}|}{|\vec{p}|^2}\right), \quad (\text{G12})$$

giving the saddle-point equation

$$J_0^{\pm} p_{(\mu)}^{\pm} + \frac{J_0^{+-}}{\beta} p_{(\mu)}^{\mp} + (p_{(\mu)}^{\pm})^{-1} = 0, \quad (\text{G13})$$

which easily gives

$$p^{\pm} = s_1 \sqrt{\frac{1}{J_0^{\pm} + s_2 \frac{J_0^{+-}}{\beta} \sqrt{\frac{J_0^{\pm}}{J_0^{\mp}}}}}, \quad (\text{G14})$$

with $s_{1,2} \in \{\pm 1\}$. The solution is the same for every μ and $p_{(\mu)}^{\pm} = (p^{\pm}, p^{\pm}, \dots, p^{\pm})$. Now, depending on the sign of the

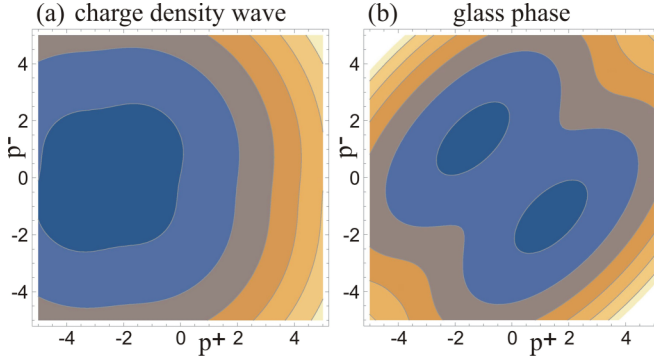


FIG. 23. Free energy (effective action $S_{\text{eff}}^{(\mu)}$) in a given replica subsystem in a photonic lattice with quenched disorder, for the case when the order parameter $p^\pm = \sum_i Q_{i\pm}$ has a nonzero saddle-point solution for the action in a given subsystem (replica). Darker (blue) tones are lower values. The ground states of the system are the local minima. In panel (a) for $J^+ = -J^- = 1$, there is a single local minimum. In case (b), for $J^+ = -J^- = 1$, we see two distinct minima of equal height, for two different nonzero values of p^\pm . Such potential energy landscape fits the description of glassy systems.

determinant $J_0^+ J_0^- - (J_0^{+-})^2$, the solutions for different $s_{1,2}$

may be minima or saddle points. In either case, we have a phase with nonzero local charge density, which is the meaning of \vec{p} . If there are multiple minima, we call this phase vortex glass. The reader may argue that true glass should satisfy more stringent conditions and that our phase is not a true glass. Depending on the viewpoint this may well be accepted, and we use the term “glass phase” merely as shorter and more convenient than “phase with power-law correlation decay, no long-range order, and frustrated free energy landscape.” The phase with a single minimum will be called charge density wave, as it has a unique ground-state configuration yielding macroscopically nonzero charge density; i.e., it has a true long-range order. On the other hand, with multiple minima the replica-averaged charge density sums to zero. The landscape, i.e., the effective action of the system for given replica (μ) as a function of p^\pm , is given in Fig. 23 as the density map of the function $S_{\text{eff}}(p^+, p^-)$ dependence for $J_0^+ = -J_0^- = 1$ (glass phase, A) and $J_0^+ = J_0^- = 1$ (charge density wave, B). We see that the glassy phase shows two inequivalent minima in each replica, with $s_1 = -s_2 = \pm 1$ in Eq. (G14), so the total action, the sum of actions of all replica subsystems, can have one and the same value for many configurations, the definition of a highly frustrated system, one of the reasons we dub this phase glass. The charge density wave only has a single minimum for $s_1 = s_2 = 1$.

2. RG flow equations

The starting Hamiltonian is the same as in (18). Now we will write it out more explicitly, keeping the distance-dependent parts:

$$\beta \mathcal{H}_{\text{eff}} = \beta \sum_{\mu=1}^n \sum_{i,j} (\bar{g}_c \vec{Q}_i^{(\mu)} \cdot \vec{Q}_j^{(\mu)} + \bar{g}'_c \vec{Q}_i^{(\mu)} \times \vec{Q}_j^{(\mu)}) \ln r_{ij} - \frac{\beta^2}{2} \sum_{\mu,v=1}^n \sum_{i,j} Q_{i\alpha}^{(\mu)} Q_{i\beta}^{(v)} \sigma_{\alpha\beta}^2 Q_{j\alpha}^{(\mu)} Q_{j\beta}^{(v)}. \quad (\text{G15})$$

We have denoted the elements of J_0 by $J_0^{++} = J_0^{--} = \bar{g}_c, J_0^{+-} = J_0^{-+} = \bar{g}'_c$ (the bars over the letter remind us that these are disorder-averaged values). The fluctuation of the partition function is completely analogous to the clean case, only it has the additional nonlocal quartic term. It can again be expanded over r_{12} as in (12) but the quartic term contains no small parameter for the power series expansion and has to be kept in the exponential form. Starting from the expression for the fluctuation analogous to the clean case (12), we get

$$\frac{\delta \mathcal{Z}}{\mathcal{Z}} = 1 + \frac{y^4}{4} \sum_{\vec{q}^{(\rho)}, \vec{q}^{(\sigma)}} e^{-\frac{\beta^2}{2} (\vec{q}^{(\rho)}, -\vec{q}^{(\sigma)}, \vec{q}^{(\rho)}, -\vec{q}^{(\sigma)}) + \frac{\beta^2}{2} (\vec{Q}^{(\mu)}, \vec{q}^{(\rho)}, \vec{Q}^{(v)}, \vec{q}^{(\sigma)})} \int dr_{12} r_{12}^3 e^{g \vec{q}^{(\rho)} \cdot \vec{q}^{(\sigma)} + g' \vec{q}^{(\rho)} \times \vec{q}^{(\sigma)}} \times \left[\int dr r^2 (g \vec{Q}_1^{(\mu)} \cdot \vec{q}^{(\rho)} + g' \vec{Q}_1^{(\mu)} \times \vec{q}^{(\rho)}) \nabla \ln |\mathbf{R}_1 - \mathbf{r}| + (g \vec{Q}_2^{(\mu)} \cdot \vec{q}^{(\rho)} + g' \vec{Q}_2^{(\mu)} \times \vec{q}^{(\rho)}) \nabla \ln |\mathbf{R}_2 - \mathbf{r}| \right]^2. \quad (\text{G16})$$

We have used the notation

$$(\vec{q}_1, \vec{q}_2, \vec{q}_3, \vec{q}_4) \equiv \sigma_{++}^2 (q_1 + q_3 + q_2 + q_4) + \sigma_{+-}^2 (q_1 + q_3 - q_2 + q_4) + \sigma_{-+}^2 (q_1 - q_3 + q_2 - q_4) + \sigma_{--}^2 (q_1 - q_3 - q_2 - q_4). \quad (\text{G17})$$

Now we trace out the fluctuations first by integrating over r and doing some simple algebra:

$$\frac{\delta \mathcal{Z}}{\mathcal{Z}} = \left[1 + 16y^4 (g + g')^2 \cosh(\beta^2 \sigma_{++}^2 + \beta^2 \sigma_{+-}^2) \cosh(\beta^2 \sigma_{-+}^2 + \beta^2 \sigma_{--}^2) (\vec{Q}_1^{(\mu)} \cdot \vec{Q}_2^{(v)} + \vec{Q}_1^{(\mu)} \times \vec{Q}_2^{(v)}) \ln R_{12} \right] \times \left[1 + 16y^4 (g - g')^2 \cosh(\beta^2 \sigma_{++}^2 - \beta^2 \sigma_{+-}^2) \cosh(\beta^2 \sigma_{-+}^2 - \beta^2 \sigma_{--}^2) (\vec{Q}_1^{(\mu)} \cdot \vec{Q}_2^{(v)} - \vec{Q}_1^{(\mu)} \times \vec{Q}_2^{(v)}) \ln R_{12} \right] \times \left[1 - 2\pi y^4 e^{-\frac{\beta^2}{2} (\sigma_{++}^2 (q_+^{(\mu)} q_+^{(v)})^2 + \sigma_{+-}^2 (q_+^{(\mu)} q_-^{(v)})^2 + \sigma_{-+}^2 (q_-^{(\mu)} q_+^{(v)})^2 + \sigma_{--}^2 (q_-^{(\mu)} q_-^{(v)})^2)} \int dr r^{1-\beta} e^{g \vec{q}^{(\mu)} \cdot \vec{q}^{(v)} + g' \vec{q}^{(\mu)} \times \vec{q}^{(v)}} \right]. \quad (\text{G18})$$

The next step is the summation over all possible ± 1 charges of virtual vortices $\vec{q}^{(\mu)}, \vec{q}^{(v)}$ (the two replica indices mean two summations from 1 to n), which requires quite some algebra. The renormalized partition function $\tilde{\mathcal{Z}}^n$ finally gives the RG flow

equations:

$$\begin{aligned} \frac{\partial g}{\partial \ell} &= -8\pi(g + g')^2 y^4 \cosh(\beta^2 \sigma_{++}^2 + \beta^2 \sigma_{+-}^2) \cosh(\beta^2 \sigma_{--}^2 + \beta^2 \sigma_{+-}^2) \\ &\quad - 8\pi(g - g')^2 y^4 \cosh(\beta^2 \sigma_{++}^2 - \beta^2 \sigma_{+-}^2) \cosh(\beta^2 \sigma_{--}^2 - \beta^2 \sigma_{+-}^2), \\ \frac{\partial g'}{\partial \ell} &= -\pi(g + g')^2 y^4 \cosh(\beta^2 \sigma_{++}^2 + \beta^2 \sigma_{+-}^2) \cosh(\beta^2 \sigma_{--}^2 + \beta^2 \sigma_{+-}^2) \\ &\quad - \pi(g - g')^2 y^4 \cosh(\beta^2 \sigma_{++}^2 - \beta^2 \sigma_{+-}^2) \cosh(\beta^2 \sigma_{--}^2 - \beta^2 \sigma_{+-}^2), \\ \frac{\partial y}{\partial \ell} &= 2\pi \left[1 - g - g' - \frac{\beta^2}{4} (\sigma_{++}^2 + 2\sigma_{+-}^2 + \sigma_{--}^2) \right] y, \quad \frac{\partial \sigma_{\alpha\beta}^2}{\partial \ell} = -2\pi \beta^4 \sigma_{\alpha\beta}^4 y^4. \end{aligned} \quad (\text{G19})$$

As discussed in the main text, the fixed point must lie either at $y = 0$ or $y \rightarrow \infty$, depending on the magnitude of $g + g' + \beta^2 \sigma^2$. For $y \rightarrow 0$, three clean fixed points remain, which flow to zero disorder: These correspond to PC, FI, and conductor. The disordered fixed point also has $y \rightarrow 0$ but the disorder is nonzero: This is the CDW phase from the mean-field analysis, the dirty analog of the insulator. Finally, when $y \rightarrow \infty$ and nonzero σ^2 at the fixed point, we expect glassy behavior.

-
- [1] M. Cross and P. Hohenberg, Pattern formation outside of equilibrium, *Rev. Mod. Phys.* **65**, 851 (1993).
 - [2] M. I. Rabinovich, A. B. Ezersky, and P. D. Weidman, *The Dynamics of Patterns* (World Scientific, Singapore, 2000).
 - [3] C. Denz, M. Schwab, and C. Weillnau, *Transverse Pattern Formation in Photorefractive Optics* (Springer, Berlin, 2003).
 - [4] L. I. Pismen, *Vortices in Nonlinear Fields* (Oxford University Press, London, 1999).
 - [5] A. Auerbach, *Interacting Electrons and Quantum Magnetism* (Springer, Berlin, 1994).
 - [6] E. Fradkin, *Field Theories of Condensed Matter Systems* (Addison-Wesley, Redwood City, CA, 1991).
 - [7] A. M. Tsvelik, *Quantum Field Theory in Condensed Matter Physics* (Cambridge University Press, Cambridge, UK, 2003).
 - [8] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay, ChaosBook.org version 15 [chaosbook.org].
 - [9] L. Leuzzi, C. Conti, V. Folli, L. Angelani, and G. Ruocco, Phase Diagram and Complexity of Mode-Locked Laser: From Order to Disorder, *Phys. Rev. Lett.* **102**, 083901 (2009).
 - [10] C. Conti and L. Leuzzi, Complexity of waves in nonlinear disordered media, *Phys. Rev. B* **83**, 134204 (2011).
 - [11] F. Antenucci, C. Conti, A. Crisanti, and L. Leuzzi, General Phase Diagram of Aultimodal Ordered and Disordered Lasers in Closed and Open Cavities, *Phys. Rev. Lett.* **114**, 043901 (2015).
 - [12] F. Antenucci, A. Crisanti, and L. Leuzzi, Complex spherical 2 + 4 spin glass: A model for nonlinear optics in random media, *Phys. Rev. A* **91**, 053816 (2015).
 - [13] N. Ghofraniha, I. Viola, F. Di Maria, G. Barbarella, G. Gigli, L. Leuzzi, and C. Conti, Experimental evidence of replica symmetry breaking in random lasers, *Nat. Commun.* **6**, 6058 (2015).
 - [14] M. C. Rechtsman, J. M. Zeuner, Y. Plotnik, Y. Lumer, D. Podolsky, F. Dreisow, S. Nolte, M. Segev, and A. Szameit, Photonic Floquet topological insulators, *Nature (London)* **496**, 196 (2013).
 - [15] Y. Plotnik, M. Rechtsman, D. Song, M. Heinrich, J. Zeuner, S. Nolte, Y. Lumer, N. Malkova, J. Xu, A. Szameit, Z. Chen, and M. Segev, Observation of unconventional edge states in “photonic graphene,” *Nat. Mater.* **13**, 57 (2014).
 - [16] R. Rajaraman, *Solitons and Instantons* (North Holland, Amsterdam, 1989).
 - [17] M. S. Petrović, M. R. Belić, C. Denz, and Yu. S. Kivshar, Counterpropagating optical beams and solitons, *Lasers Photon. Rev.* **5**, 214 (2011).
 - [18] M. I. Rodas-Verde, H. Michinel, and Yu. S. Kivshar, Dynamics of vector solitons and vortices in two-dimensional photonic lattices, *Opt. Lett.* **31**, 607 (2006).
 - [19] T. J. Alexander, A. S. Desyatnikov, and Yu. S. Kivshar, Multivortex solitons in triangular photonic lattices, *Opt. Lett.* **32**, 1293 (2007).
 - [20] B. Terhalle, T. Richter, A. S. Desyatnikov, D. N. Neshev, W. Krolikowski, F. Kaiser, C. Denz, and Yu. S. Kivshar, Observation of Multivortex Solitons in Photonic Lattices, *Phys. Rev. Lett.* **101**, 013903 (2008).
 - [21] D. Y. Tang, H. Zhang, L. M. Zhao, and X. Wu, Observation of High-Order Polarization-Locked Vector Solitons in a Fiber Laser, *Phys. Rev. Lett.* **101**, 153904 (2008).
 - [22] H. Zhang, D. Y. Tang, L. M. Zhao, and X. Wu, Dark pulse emission of a fiber laser, *Phys. Rev. A* **80**, 045803 (2009).
 - [23] V. L. Berezinsky, Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group. I Classical systems, *Sov. Phys. JETP* **32**, 493 (1971); J. Kosterlitz and D. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, *J. Phys. C* **6**, 1181 (1973).
 - [24] H. Kleinert, *Superflow and Vortex Lines*, Vol. 1 of Gauge Fields in Condensed Matter (World Scientific, Singapore, 1989).
 - [25] P. W. Anderson, Two new vortex liquids, *Nat. Phys.* **3**, 160 (2007).
 - [26] A. L. Fetter, Rotating trapped Bose-Einstein condensates, *Rev. Mod. Phys.* **81**, 647 (2009).
 - [27] J. D. Sau and S. Sachdev, Mean-field theory of competing orders in metals with antiferromagnetic exchange interactions, *Phys. Rev. B* **89**, 075129 (2014).
 - [28] E. Babaev, J. Carlström, M. Silaev, and J. M. Speight, Type 1.5-superconductivity in multicomponent systems, *Phys. C* **533**, 20 (2017).
 - [29] M. Silaev and E. Babaev, Microscopic derivation of two-component Ginzburg-Landau model and conditions of its applicability in two-band systems, *Phys. Rev. B* **85**, 134514 (2012).
 - [30] X. Ma, R. Driben, B. Malomed, T. Meoer, and S. Schumacher, Two-dimensional symbiotic solitons and vortices in binary

- condensates with attractive cross-species interaction, *Nat. Sci. Rep.* **6**, 34847 (2016).
- [31] M. Petrović, D. Jović, M. Belić, J. Schröder, P. Jander and C. Denz, Two Dimensional Counterpropagating Spatial Solitons in Photorefractive Crystals, *Phys. Rev. Lett* **95**, 053901 (2005).
- [32] D. M. Jović, M. S. Petrović, and M. R. Belić, Counterpropagating pattern dynamics: From narrow to broad beams, *Opt. Commun.* **281**, 2291 (2008).
- [33] M. S. Petrović, D. M. Jović, M. S. Belić, and S. Prvanović, Angular momentum transfer in optically induced photonic lattices, *Phys. Rev. A* **76**, 023820 (2007).
- [34] M. Petrović, D. Träger, A. Strinić, M. Belić, J. Schröder, and C. Denz, Solitonic lattices in photorefractive crystals, *Phys. Rev. E* **68**, 055601(R) (2003).
- [35] M. S. Petrović, Counterpropagating mutually incoherent vortex-induced rotating structures in optical photonic lattices, *Opt. Exp.* **14**, 9415 (2006).
- [36] B. Terhalle, T. Richter, K. Law, D. Görries, P. Rose, T. Alexander, P. Kevrekidis, A. Desyatnikov, W. Krolikowski, F. Kaiser, C. Denz, and Yu. S. Kivshar, Observation of double-charge discrete vortex solitons in hexagonal photonic lattices, *Phys. Rev. A* **79**, 043821 (2009).
- [37] D. M. Jović, S. Prvanović, R. D. Jovanović, and M. S. Petrović, Gaussian induced rotation in periodic photonic lattices, *Opt. Lett.* **32**, 1857 (2007).
- [38] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, UK, 1999).
- [39] N. Hasselmann, A. H. C. Neto, and C. M. Smith, Spin-glass phase of cuprates, *Phys. Rev. B* **69**, 014424 (2003).
- [40] V. Juričić, L. Benfatto, A. O. Caldeira, and C. Morais Smith, Dynamics of Topological Defects in a Spiral: A Scenario for the Spin-Glass Phase of Cuprates, *Phys. Rev. Lett* **92**, 137202 (2004).
- [41] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
- [42] T. Castellani and A. Cavagna, Spin-glass theory for pedestrians, *J. Stat. Mech.* (2005) P05012.
- [43] A. Perret, Z. Ristivojevic, P. Le Doussal, G. Schehr, and K. Wiese, Super-Rough Glassy Phase of the Random Field XY Model in Two Dimensions, *Phys. Rev. Lett.* **109**, 157205 (2012).
- [44] C. A. Bolle, V. Aksyuk, F. Pardo, P. L. Gammel, E. Zeldov, E. Bucher, R. Boie, D. J. Bishop, and D. R. Nelson, Observation of mesoscopic vortex physics using micromechanical oscillators, *Nature (London)* **399**, 43 (1999).
- [45] S. Bogner, T. Emig, A. Taha, and C. Zeng, Test of replica theory: Thermodynamics of two-dimensional model systems with quenched disorder, *Phys. Rev. B* **69**, 104420 (2004).
- [46] A. Crisanti and L. Leuzzi, Spherical $2 + p$ glass model: An analytically solvable model with a glass-to-glass transition, *Phys. Rev. B* **73**, 014412 (2006).
- [47] V. Juričić, L. Benfatto, A. O. Caldeira, and C. Morais Smith, Dissipative dynamics of topological defects in frustrated Heisenberg spin systems, *Phys. Rev. B* **71**, 064421 (2005).
- [48] L. Benfatto, M. Silva-Neto, V. Juričić, and C. Morais Smith, Derivation of the generalized nonlinear sigma model in the presence of the Dzyaloshinskii Moriya interaction, *Phys. B (Amsterdam, Neth.)* **378**, 449 (2006).
- [49] V. Juričić, M. B. Silva-Neto, and C. Morais Smith, Lightly Doped as a Lifshitz Helimagnet, *Phys. Rev. Lett.* **96**, 077004 (2006).
- [50] A. Gordon and B. Fisher, Phase transition theory of pulse formation in passively mode-locked lasers with dissipation and Kerr nonlinearity, *Opt. Commun.* **223**, 151 (2003).
- [51] A. Gordon and B. Fischer, Phase Transition Theory of Many-Mode Ordering and Pulse Formation in Lasers, *Phys. Rev. Lett* **89**, 103901 (2002).
- [52] N. D. Mermin, The topological theory of defects in ordered media, *Rev. Mod. Phys.* **51**, 591 (1979).
- [53] M. Silaev and E. G. Babaev, Microscopic theory of type-1.5 superconductivity in multiband systems, *Phys. Rev. B* **84**, 094515 (2011).
- [54] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, UK, 1996).
- [55] J. L. Cardy and S. Ostlund, Random symmetry-breaking fields and the XY model, *Phys. Rev. B* **25**, 6899 (1982).
- [56] M. Castellana and G. Parisi, Non-perturbative effects in spin glasses, *Nat. Sci. Rep.* **5**, 8697 (2015).
- [57] D. S. Fisher and D. A. Huse, Equilibrium behavior of the spin-glass ordered phase, *Phys. Rev. B* **38**, 386 (1988).
- [58] J. Zaanen, Quantum phase transitions in cuprates: Stripes and antiferromagnetic supersolids, *Phys. C (Amsterdam, Neth.)* **317**, 217 (1999).
- [59] O. Gat, A. Gordon, and B. Fischer, Solution of a statistical mechanics model for pulse formation in lasers, *Phys. Rev. E* **70**, 046108 (2004).
- [60] R. Weill, A. Rosen, A. Gordon, O. Gat, and B. Fischer, Critical Behavior of Light in Mode-Locked Lasers, *Phys. Rev. Lett* **95**, 013903 (2005).
- [61] O. Sandfuchs, F. Kaiser, and M. R. Belić, Dynamics of transverse waves and zigzag instabilities in photorefractive two-wave mixing with a feedback mirror, *J. Opt. Soc. Am. B* **18**, 505 (2001).
- [62] O. Sandfuchs, F. Kaiser, and M. R. Belić, Self-organization and Fourier selection of optical patterns in a nonlinear photorefractive feedback system, *Phys. Rev. A* **64**, 063809 (2001).
- [63] V. I. Arnol'd, V. S. Afraimovich, Yu. S. Il'yashenko, and L. P. Shil'nikov, *Bifurcation Theory and Catastrophe Theory* (Springer-Verlag, Berlin, 1994).



Vortex dynamics of counterpropagating laser beams in photorefractive materials

Mihailo Čubrović¹ · Milan Petrović^{2,3}

Received: 13 October 2017 / Accepted: 12 October 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

We study vortex patterns of counterpropagating laser beams in a photorefractive crystal, with or without the background photonic lattice. The vortices are effectively planar and have two “flavors” because there are two opposite directions of beam propagation. In a certain parameter range, the vortices form stable equilibrium configurations which we study using the methods of statistical field theory and generalize the Berezinsky–Kosterlitz–Thouless transition of the XY model to the “two-flavor” case. In the nonequilibrium regime, the patterns exhibit an Andronov–Hopf bifurcation which may lead to oscillations (limit cycle), chaos or decay to zero intensity due to radiation losses. We show how to identify various pathways toward instability from intensity patterns, i.e. from experiment.

Keywords Vortex · BKT transition · Photorefractive optics · Statistical field theory

1 Introduction

Nonlinear optical systems are a rich arena for studies of various fundamental physical phenomena. The strong response of the nonlinear optical medium to the propagation of light makes it a typical strongly correlated system, with many phenomena similar to those in other strongly interacting systems in areas such as condensed matter. Their complex dynamics offers an opportunity to study spatiotemporal chaos and optical turbulence

This article is part of the Topical Collection on Focus on Optics and Bio-photonics, Photonica 2017.

Guest Edited by Jelena Radovanovic, Aleksandar Krmpot, Marina Lekic, Trevor Benson, Mauro Pereira, Marian Marciniak.

✉ Mihailo Čubrović
mcubrovic@gmail.com

Milan Petrović
petrovic@ipb.ac.rs

¹ Scientific Computing Laboratory, Institute of Physics, University of Belgrade, Pregrevica 118, Belgrade 11080, Serbia

² Institute of Physics, P. O. B. 57, Belgrade 11001, Serbia

³ Texas A&M University at Qatar, Doha P.O.Box 23874, Qatar

(Cross and Hohenberg 1993; Rabinovich et al. 2000). On the other hand, they often also exhibit stable, equilibrium configurations in a certain parameter range, which are naturally studied by statistical physics methods. Vortices and other topological configurations (Alexander et al. 2007; Anderson 2007; Fetter 2009), long-range order (Anderson 2007), quenched disorder and glassy behavior (Antenucci et al. 2015a, b; Ghofraniha 2015; Perret et al. 2012) are universal in a broad range of systems such as cold atoms (Bagnato et al. 2015; Malomed et al. 2016) and magnetic systems, and the relative simplicity of experiments in optics makes it an excellent testing ground for strongly coupled models.

In this paper we study a specific and experimentally realizable nonlinear optical system: laser beams counterpropagating (CP) through a photorefractive (PR) crystal. This means we have an elongated PR crystal (with one longitudinal and two transverse dimensions) and two laser beams shone onto each end. We thus effectively have two fields, one forward-propagating and one backward-propagating. The optical response of the crystal depends nonlinearly on the *total* intensity of both beams, which means the beams effectively interact with each other. This system has been thoroughly investigated for phenomena such as dynamical solitons (Denz et al. 2003; Petrović et al. 2011, 2005; Jović et al. 2008), vortex stability on the photonic lattice (Alexander et al. 2007; Terhalle et al. 2008; Čubrović and Petrović 2017) and topological invariants (Rechtsman et al. 2013).

We first recast the system in Lagrangian and then in Hamiltonian form so it can be studied as a field theory, which depends parametrically on the time t . Then we consider the time dynamics of the system and show that in a broad parameter range the patterns relax to a static configuration which can be studied within *equilibrium* field theory. By renormalization group (RG) analysis, we obtain the phase diagram of static vortex configurations. The phase diagram is obviously closely related to the famous Berezinsky–Kosterlitz–Thouless (BKT) vortex unbinding transition in the XY model (Berezinsky 1971; Kosterlitz and Thouless 1973) except that having two components of the field produces additional phases and phase transitions, due to forward–backward beam interaction. The analytical insight we obtain also allows us to avoid overextensive numerics – analytical construction of the phase diagram tells us which patterns can in principle be expected in different corners of the parameter space.

Next we focus on the nonequilibrium regime, classify the fixed points and study possible routes of instability. We emphasize the pictorial and “rule-of-thumb” criteria to recognize various instabilities, in order to facilitate experimental checks. At the end we will discuss the perspective of studying dynamical criticality, i.e. instabilities which consist in moving from one vortex phase to another in real time, a phenomenon which is intimately connected to the difficult questions of quench dynamics and thermalization in many-body systems.

2 Counterpropagating beams in photorefractive medium: equations of motion

Consider a photorefractive crystal of length L irradiated by two paraxial head-on laser beams which propagate from the opposite faces of the crystal in the z -direction. Photorefractive crystals induce self-focusing of the beams—the vacuum (linear) wave equation is modified by the addition of a friction-like term, so the diffusion of the light intensity (the broadening of the beam) is balanced out by the self-focusing of the beam. The physical ground for this is the redistribution of the charges in the crystal due to the Kerr effect.

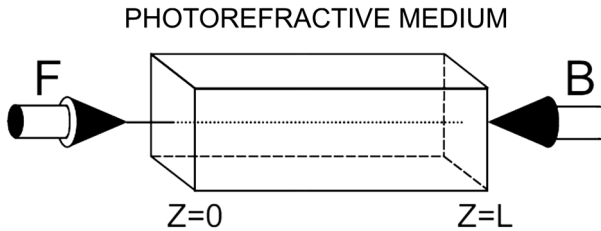


Fig. 1 Experimental setup for the study of the CP beams in the PR crystal. The crystal has the shape of a parallelepiped, and the beams propagate along the longitudinal, z -axis: the forward (F)-beam from $z = 0$ to $z = L$, and the backward (B)-beam the other way round. The intensity patterns can be observed at the transverse faces of the crystal, at $z = 0$ and $z = L$

The nonlinearity is contained in the change of the refraction index which is determined by the induced charge density. A sketch of the system is given in Fig. 1. Before entering the crystal, the laser beams can be given any desirable pattern of both intensity and phase. In particular, one can create vortices (winding of the phase) making use of the phase masks (Denz et al. 2003).

Assuming the electromagnetic field of the form $\mathbf{E} = e^{i\omega t + i\mathbf{q}\mathbf{r}} (F e^{ikz} + B e^{-ikz})$, we can write equations for the so-called envelopes F and B of the forward- and backward-propagating beams along the z -axis (the frequency, transverse and longitudinal momentum are denoted respectively by ω, \mathbf{q}, k). The wave equations for F and B are now:

$$\pm i \partial_z \Psi_{\pm}(z;x,y;t) + \Delta \Psi_{\pm}(z;x,y;t) = \Gamma E(z;x,y;t) \Psi_{\pm}(z;x,y;t), \tag{1}$$

where the plus and minus signs on the left-hand side stand for the forward- and backward-propagating component of the beam amplitude doublet $\Psi \equiv (\Psi_+, \Psi_-) \equiv (F, B)$, and Γ is the dimensionless PR coupling constant. From now on we will use $\alpha \in \{+, -\}$ to denote the two beams (F and B) and call it a flavor index, in analogy with field theory. The vorticity (winding number of the phase) will be called vortex charge as usual. The charge field E on the right-hand side of the equation is the electric field sourced by the charges in the crystal (i.e., it does not include the external electric field of the beams). Its evolution is well represented by a relaxation-type equation (notice that the derivative $\partial_t E$ is strictly negative) (Petrović et al. 2011):

$$\frac{\tau}{1 + I(z;x,y;t)} \partial_t E(z;x,y;t) + E(z;x,y;t) = - \frac{I(z;x,y;t)}{1 + I(z;x,y;t)}. \tag{2}$$

Here, $I \equiv I_p + I_x$ is the total light intensity at a given point, $I_p \equiv |F|^2 + |B|^2$ is the beam intensity and I_x the intensity of the fixed background. The meaning of I_x is that the crystal is all the time irradiated by some constant light source, independent of the counter-propagating beams with envelopes F, B . The relaxation time is τ . The form of the non-linearity accounts for the saturation of the crystal; notice that a simple quartic non-linear Schrödinger equation would not account for the saturation.¹ In the numerical calculations,

¹ One might also worry that a realistic crystal is anisotropic, while our equation is isotropic. Nevertheless, comparison to experiment (Neshev et al. 2004; Fleischer et al. 2004; Dreischuh et al. 2002) shows that this model is able to describe actual measurements rather well. Also, the effects of anisotropy can be suppressed in experiment by illuminating the crystal by *uniform* light for very long times before starting the experiment (Cohen et al. 2002).

we solve Eqs. (1), (2) with no further assumptions, using a slightly modified version of the beam propagation method (Sandfuchs et al. 2001). For analytical results we will need to transform them further assuming a vortex pattern. The Eq. (2) is completely phenomenological, but it excellently represents the experimental results (Denz et al. 2003). We will first consider the equilibrium regime, and then the nonequilibrium dynamics.

For slow time evolution (in absence of pulses), we can Laplace-transform the Eq. (2) in time ($E(t) \mapsto E(u) = \int_0^\infty dt e^{-ut} E(t)$) to get the algebraic relation

$$E(z; x, y; u) = -\frac{\Psi^\dagger \Psi + I_x - \tau E_0}{1 + \tau u + I_x + \Psi^\dagger \Psi} = -1 + \frac{1 + \tau u + \tau E_0}{1 + \tau u + I_x + \Psi^\dagger \Psi}. \quad (3)$$

The original system (1) can now be described by the Lagrangian:

$$\mathcal{L} = i\Psi^\dagger \sigma_3 \partial_z \Psi - |\nabla \Psi|^2 + \Gamma \Psi^\dagger \Psi - \Gamma(1 + \tau E_0 + \tau u) \log(1 + \tau u + I_x + \Psi^\dagger \Psi), \quad (4)$$

where σ_3 is the Pauli matrix $\sigma_3 = \text{diag}(1, -1)$. This has the form $\mathcal{L} = i\Psi^\dagger \sigma_3 \partial_z \Psi - |\nabla \Psi|^2 - V_{\text{eff}}(\Psi^\dagger, \Psi)$, i.e. the Lagrangian of a non-relativistic field theory (a two-component nonlinear Schrödinger field equation) in $2 + 1$ dimensions (x, y, z), where the role of time is played by the longitudinal distance z , and the physical time t (or u upon the Laplace transform) is a parameter.

3 Stable vortex configurations and the phase diagram

Following the same steps as for the textbook XY model we can arrive at an effective Hamiltonian for stable vortex configurations. For details we refer the reader to Čubrović and Petrović (2017). Assuming the vortex solution of the form

$$\Psi_{0\pm}(\mathbf{r}) = \psi_{0\pm}(r) e^{i\delta\theta_\pm(\phi) + i\theta_{0\pm}(\phi)}, \quad (5)$$

where $\theta_\pm(\phi)$ is the singular part of the phase and $\theta_{0\pm}(\phi)$ the regular part, we want to integrate out both the amplitude fluctuations and the regular part of the phase and arrive at a description of the systems solely in terms of vortex charges. This is done by expanding the Lagrangian (4) to quadratic order in both amplitude and phase fluctuations and integrating them out. Then the usual Legendre transform yields the vortex Hamiltonian:

$$\mathcal{H}_{\text{vort}} = \sum_{i < j} (g Q_{i\alpha} Q_{j\alpha} + g' Q_{i\alpha} \times Q_{j\beta}) \log r_{ij} + \sum_i (g_0 Q_{i\alpha} Q_{i\alpha} + g_1 Q_{i\alpha} \times Q_{i\beta}). \quad (6)$$

We denote the flavor \pm by Greek indices, and the summation convention is understood.² Furthermore, we denote $Q_\alpha \times Q_\beta \equiv Q_{i+} Q_{j-} + Q_{i-} Q_{j+}$. The first term is the expected Coulomb interaction of vortices from the XY model (Berezinsky 1971; Kosterlitz and Thouless 1973); notice that only like-flavored charges interact through this term (because the kinetic term $|\nabla \Psi|^2$ is homogenous quadratic). The second term is the forward–backward interaction, also with Coulomb-like (logarithmic) radial dependence. This interaction is generated by the coupling of amplitude fluctuations $\delta\psi_\alpha(r)$ to the phase fluctuations. In a system without amplitude fluctuations, i.e. classical spin system, this term would not be generated. The third and fourth term constitute the energy of the vortex core. The self-interaction

² There is no difference between upper and lower indices as both flavors always enter the sum with positive sign.

constants g_0, g_1 are of course dependent on the vortex core size and behave roughly as $g \log a/\epsilon, g' \log a/\epsilon$, where ϵ is the UV cutoff. The final results will not depend on ϵ , as expected, since g_0, g_1 can be absorbed in the fugacity y (see the next subsection). Expressions for the coupling constants in terms of original parameters are given in Čubrović and Petrović (2017); they can be used to relate the theoretical phase diagram to experiment.

To describe the phase diagram, we will perform the renormalization group (RG) analysis. Here we follow closely the calculation for conventional vortex systems. We consider the fluctuation of the partition function δZ upon the formation of a virtual vortex pair at positions $\mathbf{r}_1, \mathbf{r}_2$ ³ with charges $q_\alpha, -q_\alpha$, (with $\mathbf{r}_1 + \mathbf{r}_2 = 2\mathbf{r}$ and $\mathbf{r}_1 - \mathbf{r}_2 = \mathbf{r}_{12}$), in the background of a vortex pair at positions $\mathbf{R}_1, \mathbf{R}_2$ (with $\mathbf{R}_1 + \mathbf{R}_2 = 2\mathbf{R}$ and $\mathbf{R}_1 - \mathbf{R}_2 = \mathbf{R}_{12}$) with charges $Q_{1\alpha}, Q_{2\alpha}$. It is also convenient to replace the core self-interaction constants $g_{0,1}$ with the fugacity parameter defined as $y \equiv \exp[-\beta(g_0 + g_1) \log \epsilon]$. We also introduce the notation $\beta \equiv L$ in analogy with the inverse temperature β in standard statistical mechanics but of course the physical meaning of β in our system is very different: we have no thermodynamic temperature or thermal noise, and the third law of thermodynamics is not satisfied for the “temperature” $1/\beta$. We merely use the β -notation for reasons of formal similarity, not as a complete physical analogy.

This is a straightforward but lengthy calculation and we state just the resulting flow equations:

$$\frac{\partial g}{\partial \ell} = -16\pi(g^2 + g'^2)y^4, \quad \frac{\partial g'}{\partial \ell} = -2\pi g g' y^4, \quad \frac{\partial y}{\partial \ell} = 2\pi(1 - g - g')y. \tag{7}$$

Notice that if one puts $g' = 0$, they look very much like the textbook XY model RG flow, except that the fugacity enters as y^4 instead of y^2 (simply because every vortex contributes two charges). We can find fixed points analytically and then numerically integrate the flow equations to find exact phase borders. The fugacity y can flow to zero (meaning that the vortex creation is suppressed and the vortices tend to bind) or to infinity, meaning that vortices can exist at finite density. At $y = 0$ there is a fixed line $g + g' = 1$. This line is attracting for the half-plane $g + g' > 1$; otherwise, it is repelling. There are three more attraction regions when $g + g' < 1$. First, there is the point $y \rightarrow \infty, g = g' = 0$ which has no analogue in single-component vortex systems. Then, there are two regions when $g \rightarrow \infty$ and $g' \rightarrow \pm\infty$ (and again $y \rightarrow \infty$). Of course, the large g, g' regime is strongly interacting and the perturbation theory eventually breaks down. What happens when g, g' flow toward very large values is that the intensity at the vortex core becomes very large, so the lowest-order, quadratic Hamiltonian needs to be supplemented by higher-order terms in intensity fluctuations. To integrate them out, one needs to perform a diagrammatic expansion which leads to quartic- and higher-order terms in vortex charges Q_α in the effective vortex Hamiltonian [Eq. (6)], ultimately correcting the flow at large g, g' to flow toward finite values g_*, g'_* and g_{**}, g'_{**} .

The RG flows in the $g - g'$ plane are given in Fig. 2. The situation is now the following:

1. The attraction region of the fixed line is the vortex insulator phase (INS): the creation rate of the vortices is suppressed to zero. There is no vortex charge conservation.

³ The boldface vectors are the coordinate vectors in the plane.

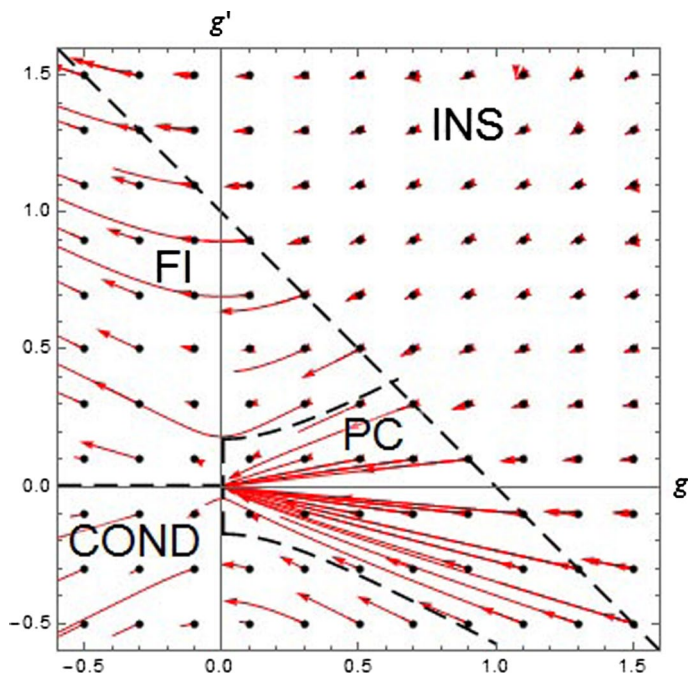


Fig. 2 Phase diagram for the clean system in the g – g' plane, at the mean-field level with RG flows. We show the flows for a grid of initial points, denoted by black dots; red lines are the flows. Four phases exist, whose boundaries are delineated by black dashed lines: conductor (COND), insulator (INS), frustrated insulator (FI) and perfect conductor (PC). The straight line $g + g' = 1$ is obtained analytically whereas the other phase boundaries can only be found by numerical integration of the flow Eq. (7). The flows going to infinity are the artifacts of the perturbative RG; they correspond to finite values which are beyond the scope of our analytical approach. Notice how the flows in the $g + g' > 1$ phase all terminate at different values

2. The zero-coupling fixed point attracts the trajectories in the vortex perfect conductor phase (PC): only the fugacity controls the vortices and arbitrary charge configurations can form. Each vortex charge, Q_+ and Q_- , is separately conserved.
3. In the attraction region of the fixed point with $g_* < 0$ and $g'_* > 0$ (formally they flow to $-\infty$ and $+\infty$, respectively), same-sign F - and B -charges attract each other and those with the opposite sign which repel each other. This is the frustrated insulator (FI): it conserves only the combination $Q_+ + Q_-$, and only vortices with charge $(Q_+, -Q_+)$ are stable.
4. The fixed point with $g_{**}, g'_{**} < 0$ (formally both flow to $-\infty$) corresponds to the conductor phase (COND). This phase preserves one of the charges, Q_+ or Q_- , i.e. either $(Q_+, 0)$ - or $(0, Q_-)$ -vortices proliferate.

In the half-plane $g + g' > 1$ every point evolves toward a different, finite point (g, g') in the same half-plane. In the other half-plane we see the regions of points moving toward the origin or toward one of the two directions at infinity. In the future we plan to apply this formalism also to multi-component vortices in Bose–Einstein condensates (Ma et al. 2016) and in particular in type-1.5 superconductors (Silaev and Babev 2012), where even more complex phenomena, including frustration, are observed as a consequence of multi-component interaction.

An interesting line of research consists in adding disorder to the above system. We consider this problem in Čubrović and Petrović (2017) and find that the system can be approximated by a random-coupling *and* random-field two-component XY-like model, related to the Cardy–Ostlund model (1982). The replica formalism (Castellana and Parisi 2015) then predicts a glassy phase with slow dynamics, strong correlations and no long-range order. This is however a separate story and we will leave it out here. Interested readers can consult (Čubrović and Petrović 2017) and look at related work in Antenucci et al. (2015a, b).

4 Time-dependent regime

Here our goal is twofolds. First, we have to show that at least for some boundary conditions and parameter values there is a stable fixed point of the time evolution, so that the system reaches a time-independent, equilibrium pattern. The reason is that the whole formalism of the previous chapter is only valid for such configurations, as it departs from equilibrium statistical mechanics. Second, we want to check other, non-static behaviors as they are interesting in their own right and experimentally relevant (but one should not expect them to be described by an equilibrium phase diagram like Fig. 2).

Time dynamics can be studied in a straightforward way, making use of the relaxation Eq. (2) to write down the first-order evolution equations for Ψ_{\pm} :

$$\frac{\partial \Psi_{\alpha}^{\pm}}{\partial t} = -\frac{\Gamma}{\tau} \frac{((1+I)E+I)}{\alpha k - q^2 - \Gamma E} \Psi_{\alpha}^{\pm}, \quad \frac{\partial E}{\partial t} = -\frac{1}{\tau} ((1+I)E+I). \tag{8}$$

This system has three equilibrium points. One is the trivial equilibrium with zero intensity (“0” point):

$$(\Psi_{+}^{\pm}, \Psi_{-}^{\pm}, E) = \left(0, 0, -\frac{I_x}{1+I_x}\right),$$

and the remaining two are related by a discrete symmetry $\Psi_{\pm} \mapsto \Psi_{\mp}$ (“±” points). The “+” point is

$$(\Psi_{+}^{\pm}, \Psi_{-}^{\pm}, E) = \left(\sqrt{\frac{E(1+I_x)+I_x}{1+E}} e^{i\phi_{+}}, 0, E\right),$$

and the “-” point has instead $\Psi_{+} = 0$ and $\Psi_{-} = \sqrt{(E(1+I_x)+I_x)/(1+E)} \exp(i\phi_{-})$. Notice that the phase ϕ_{\pm} remains free to vary so the “±” solutions support vortices. We first ask what is the stability criterion for a nontrivial solution, i.e. one of the “±” points, as this is the main criterion for the applicability of the equilibrium statistical mechanics methods in the previous section. Introducing the amplitudes of the fluctuations from equilibrium as $X_{1,3} = \Re \delta \Psi_{\pm}, X_{2,4} = \Im \delta \Psi_{\pm}, X_5 = \delta E$, we can do a first-order stability analysis as the system is non-degenerate. Rescaling $X_1 \mapsto (1+E_0)^{-3/4} (I_x + E_0(1+I_x))^{1/2}$ and $t \mapsto t((1+E_0)/(I_x + E_0(1+I_x)))^{1/4}$, the equation of motion for the “±” point reads

$$\partial_t \begin{pmatrix} X_1 \\ X_5 \end{pmatrix} = \begin{pmatrix} -\frac{a_{\pm}}{\Gamma E_0 + k + q^2} & -1 \\ 1 & -\frac{a_{\pm}}{\Gamma E_0 + k + q^2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_5 \end{pmatrix} + O(X_1^2 + X_5^2; X_2, X_3, X_4), \tag{9}$$

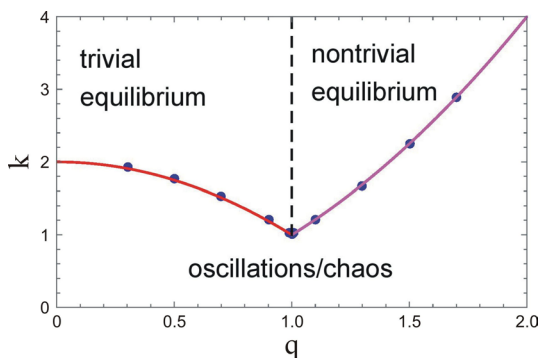


Fig. 3 Stability diagram in the q - k plane. The onset of instability for $k < k_c(q)$ is found numerically for a range of q values. The solid lines are the analytical prediction for the stability of the “0” point ($k_c = q^2$, magenta) and of the “+” point ($k_c = \Gamma E_0 - q^2 \approx \Gamma - q^2$, red). The black dashed line at $q = q_c \approx 1$ separates the stability regions of the two points. The domain of applicability of our main results is the top right corner (nontrivial equilibrium), above $k > k_{\min} \sim 1/L$ and for not too large q values. Parameter values: $\Gamma = 2, I_x = 0$

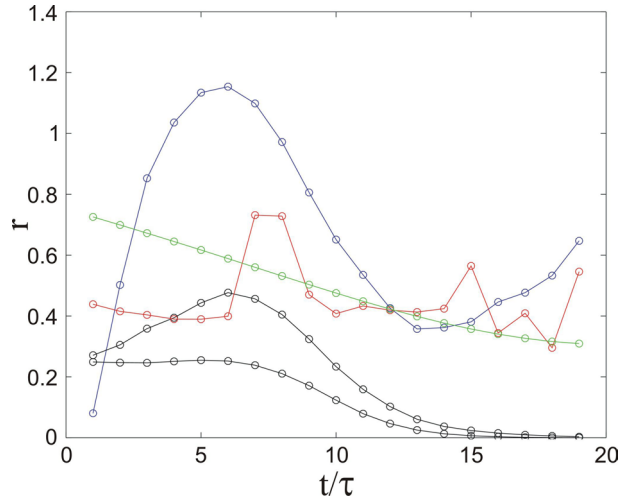
with a_{\pm} being some (known) *positive* functions of Γ, E_0, I_x (independent of k, q). This is precisely the normal form for the Andronov–Hopf bifurcation (Arnol’d et al. 1994), and the bifurcation point lies at $k = -\Gamma E_0 - q^2$. To remind, the bifurcation happens when the off-diagonal element in the linear term changes sign: the fixed point is stable when $a_{\pm}/(\Gamma E_0 + k + q^2)$ is positive. The sign of the nonlinear term determines the supercritical/subcritical nature of the bifurcation.⁴

Now the textbook analysis of the Andronov–Hopf bifurcation tells us that stable “+” equilibrium exists for $k > -\Gamma E_0 - q^2$ where E_0 is best found numerically. Exactly the same condition holds for the “-” point. For $k < -\Gamma E_0 + q^2$, dynamics depends on the sign of the nonlinear term in Eq. (9). For the positive sign we expect periodically changing patterns and for the negative sign (subcritical bifurcation), various possibilities arise: the system may wander chaotically between the “+” and the “-” points, or it may end up in the attraction region of the “0” point and fall onto the trivial solution with zero intensity. Naively, the attraction regions of the two fixed points (“±” and “0”) are separated by the condition $-\Gamma E_0 - q^2 = q^2$, i.e. $q_c = \sqrt{-\Gamma E_0(\Gamma, \tau)/2}$, where we have emphasized that E_0 is in general non-universal. The actual boundary may be more complex however, as our analysis is based on finite-order expansion around the fixed points, which is not valid far away from them.

The numerical stability diagram is given in Fig. 3. The stability limit turns out to be $k > \Gamma - q^2$, i.e. $E_0 \approx -1$. The curves separating the attraction regions of the three equilibrium points follow exactly the quadratic scaling in q as predicted by the analytical stability analysis. The equilibrium region lies in the top right corner of the diagram (nontrivial equilibrium), above $k \approx 1/L$. This is where the patterns evolve towards static

⁴ Negative sign means the fixed point is stable everywhere before the bifurcation and is replaced by a stable limit cycle after the bifurcation (supercritical). Positive sign means the fixed point coexists with the stable limit cycle before the bifurcation and the (X_1, X_5) plane is divided among their attraction regions; after the bifurcation there is no stable solution at all (subcritical). However, one should not take the stability in the whole (X_1, X_5) plane in the supercritical case too seriously. We have expand the equations of motion in the vicinity of the fixed points and the expansion ceases to be valid far away from the origin.

Fig. 4 Time evolution of the relaxation rate r for the various situations from Figs. 5 and 6, illustrating the relaxation to non-trivial (non-zero intensity) equilibrium, i.e. “ \pm ” fixed points (Fig. 5a, c, black), limit cycle (6a, blue), chaos (6b, red) and the relaxation to trivial (zero intensity) equilibrium, i.e. “0” fixed point (6c, green). In the main text we mainly study the cases like the black curves, where time-independent stable configurations are seen. The circles are data points from numerics and the lines are just to guide the eye



long-time configurations. The top left corner describes “boring” situations, when all light ultimately radiates away from the crystal and intensity drops to zero. The bottom region contains nontrivial dynamics: depending on parameters, it may contain a limit cycle (corresponding to oscillating patterns) or aperiodic wandering among an alphabet of unstable patterns (chaos).

Formally, both k and q can be any real numbers. In practice, however, k is discrete and its minimal value is of the order $1 / L$. The spatial momentum q lies between the inverse of the transverse length of the crystal (which is typically an order of magnitude smaller than L , i.e. minimal q can be assumed equal to zero) and some typical small-scale cutoff which in our case is the vortex core size.

Now we test our conclusions numerically. A convenient quantity to differentiate between different stability regimes is the relaxation rate

$$r \equiv \frac{1}{X} \frac{dX}{dt} = \frac{\sum_{x,y} |X(t_{j+1};x,y) - X(t_j;x,y)|^2}{\sum_{x,y} |X(t_j;x,y)|^2}, \tag{10}$$

which is expected to reach zero for a generic relaxation process, where in the vicinity of an asymptotically stable fixed point $X \sim X_{\text{eq}} + xe^{-rt}$ will be generically nonzero for a limit cycle or chaos, and will asymptote to a constant for the “0” point, where $X_{\text{eq}} = 0$ so we get $(1/X)dX/dt \sim r$.

Figure 4 summarizes these possibilities in terms of the relaxation rate r , whereas Figs. 5 and 6 show how the patterns evolve in some representative cases. The black curves in Fig. 4 show the situation which is in the focus of this work – the approach toward static equilibrium. This corresponds to the phases from Fig. 2. In Fig. 5 we see how the equilibrium configurations are reached (for three phases). In each case we start with a regular lattice of circular vortices. In the PC phase (Fig. 5a) the vortices expand somewhat but in principle retain the original configuration (and charges). The other two phases (Fig. 5b, c) have nontrivial transient dynamics and undergo the lattice inversion, but eventually (for times about $t \approx 20 - 25\tau$) they stabilize and form a static inverse lattice (with charges $(3, -3)$ in the FI case and with zero charge in the INS case).

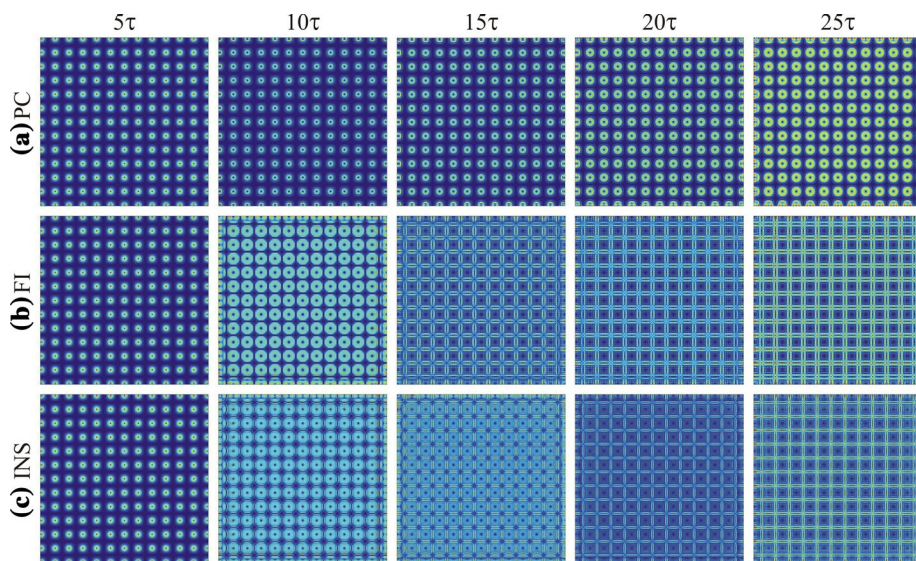


Fig. 5 Time evolution of patterns at five different times: **a** perfect conductor phase, **b** frustrated insulator phase and **c** insulator phase. In all cases the approach to equilibrium is obvious, and we expect that for long times a thermodynamic description is justified

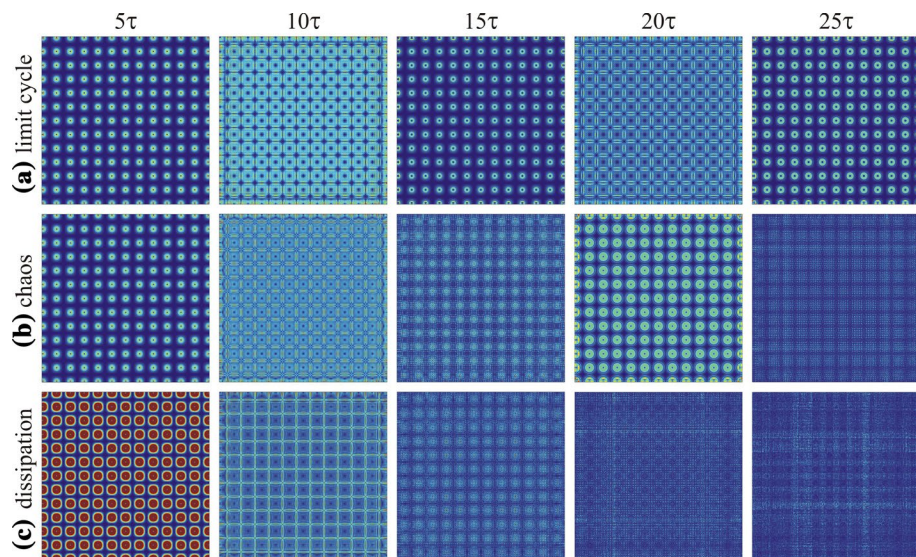


Fig. 6 Time evolution of non-equilibrium patterns. In **a** the limit cycle leads to permanent oscillatory behavior, in **b** wandering along the unstable manifold between the equilibrium points gives rise to chaos and in **c** dissipation wins and dynamics dies out. The parameters are the same as in the previous figure, except that the length L is increased thrice

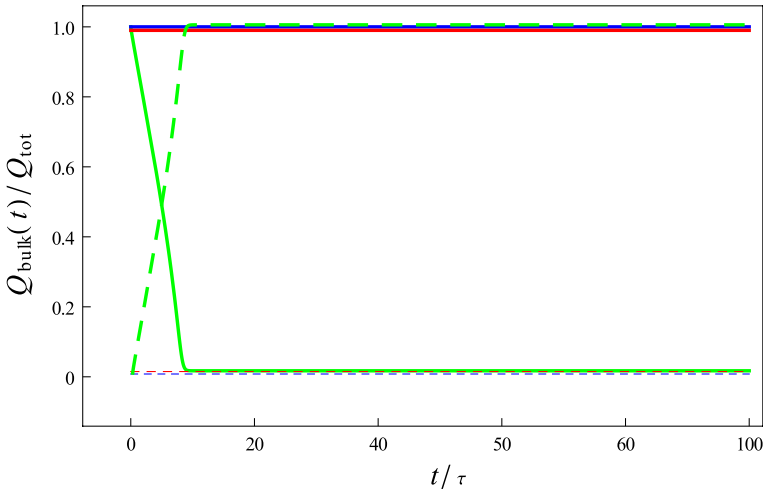


Fig. 7 Same systems as in Fig. 6 but now for the time dependence of the F -vortex charge in the bulk (full lines) and the vortex current flow through their boundary (dashed lines). While the limit cycle (blue) and chaos (red) keep all vortex charge in the bulk, dissipation toward the trivial equilibrium (green) has a systematic vortex flow toward the edges. This is a finite-size effect which would not happen in an infinite field (but it does happen in real-world PR crystals which are, of course, finite)

The other curves in Fig. 4 describe dynamics which does not result in a nontrivial static pattern. The blue curve shows a limit cycle leading to periodic oscillations of the pattern, with half-period about 10τ . The corresponding patterns are seen in Fig. 6a, where we see how the vortex lattice keeps coming back to the original configuration at times $\approx 5\tau, 10\tau, 15\tau$. The red curve corresponds to the chaotic regime with aperiodic dynamics and no relaxation, as in Fig. 6b. Here the pattern keeps changing, wandering among the original lattice (for $t = 5\tau, 20\tau$), the inverse lattice (for $t = 10\tau$) and more or less incoherent patterns (for $t = 15\tau, 25\tau$). Finally, the green curve in Fig. 4 reaches a constant value of r . This corresponds to the pattern which radiates away in Fig. 6c, with total intensity being almost zero for $t > 20\tau$. Here one might wonder what happens to the vortex charge when the initially regular vortex lattice ends up as an incoherent, low-intensity configuration which obviously does not support vortices. The explanation is that the vortex charge flows outward, eventually reaching the edges of the crystal. The finite-size effects then invalidate the vortex charge conservation, as the usual proof that the winding number of the phase is a topological invariant crucially depends on considering the winding at infinity. Vortex charge thus dissipates at the edges. This is demonstrated in Fig. 7, which presents the same systems as in Fig. 6a–c but shows the ratio of the total bulk vortex charge $Q_{\text{bulk}}(t)$ to the total initial vortex charge Q_{tot} . Total initial charge is calculated by definition, as the integrated vorticity of the F -beam, $Q_{\text{tot}} = \int dx \int dy |\omega_F|$, with

$$\omega_F = (\cos \theta_F \partial_x \theta_F, \sin \theta_F \partial_y \theta_F), \tag{11}$$

and all the quantities are taken at $t = 0$. The integral $\int dx \int dy |\omega_F|$ equals precisely the total F -vortex charge summed over all vortices. The bulk charge is computed by subtracting the integrated vorticity flow along the boundary:

$$Q_{\text{bulk}}(t) = Q_{\text{tot}} - \int_0^t dt' \oint d\mathbf{l} \cdot \omega_F(t). \tag{12}$$

Figure 7 shows that the sum of the bulk charge and the vortex current through the boundary is preserved in all cases, including when chaos or dissipation makes the pattern incoherent. In the last case, however, all the vortex charge flows toward the boundaries – this is a finite-size effect which would be absent in infinite field but is observable in realistic PR-crystals which are of finite dimensions. In practice, the matters are even more complicated as the boundary surface carries also new physics (surface polarization etc.), so the starting equations of motion would have to be modified. We believe, however, that the basic picture of vortex charge dissipating at the boundary still remains, because the mapping from the internal $U(1)$ phase onto the loop in the coordinate plane is explicitly broken by the boundary (whatever its detailed physics might be), and the vortex charge nonconservation at the boundary follows from this breaking.

The next task is to consider in more detail the decay of an ordered phase, either to chaos or to a limit cycle (radiating away all intensity is likely a trivial process, fully described by the approximately constant decay rate). We plan to address this problem in further work, and to relate the results to the question of quench dynamics in vortex systems.

Acknowledgements Work at the Institute of Physics is funded by Ministry of Education, Science and Technological Development, under Grants Nos. OI171033 and OI171017. M.P. is also supported by the NPRP 8-028-1-001 project of the Qatar National Research Fund (a member of the Qatar Foundation).

References

- Alexander, T.J., Desyatnikov, A.S., Kivshar, Y.S.: Multivortex solitons in triangular photonic lattices. *Opt. Lett.* **32**, 1293–1295 (2007)
- Anderson, P.W.: Two new vortex liquids. *Nat. Phys.* **3**, 160–162 (2007)
- Antenucci, F., Conti, C., Crisanti, A., Leuzzi, L.: General phase diagram of multimodal ordered and disordered lasers in closed and open cavities. *Phys. Rev. Lett.* **114**, 043901 (2015a)
- Antenucci, F., Crisanti, A., Leuzzi, L.: Complex spherical $2 + 4$ spin glass: a model for nonlinear optics in random media. *Phys. Rev. A* **91**, 053816 (2015b)
- Arnol'd, V.I., Afraimovich, V.S., Il'yashenko, Y.S., Shil'nikov, L.P.: *Bifurcation Theory and Catastrophe Theory*. Springer, Berlin (1994)
- Bagnato, V.S., Frantzeskakis, D.J., Kevrekidis, P.G., Malomed, B.A., Mihalache, D.: Bose–Einstein condensation: twenty years after. *Rom. Rep. Phys.* **67**, 5–50 (2015)
- Berezinsky, V.L.: Destruction of long range order in one-dimensional and two-dimensional systems having a continuous symmetry group. I. Classical systems. *Sov. Phys. JETP* **32**, 493–500 (1971)
- Cardy, J.L., Ostlund, S.: Random symmetry-breaking fields and the XY model. *Phys. Rev. B* **25**, 6899–6909 (1982)
- Castellana, M., Parisi, G.: Non-perturbative effects in spin glasses. *Nat. Sci. Rep.* **5**, 8697 (2015). [arXiv:1503.02103](https://arxiv.org/abs/1503.02103) [cond-mat.dis-nn]
- Cohen, O., Lan, S., Harmon, T., Giordmaine, J.A., Segev, M.: Spatial vector solitons consisting of counter-propagating fields. *Opt. Lett.* **27**, 2013–2015 (2002)
- Cross, M., Hohenberg, P.: Pattern formation outside of equilibrium. *Rev. Mod. Phys.* **65**, 851–1112 (1993)
- Čubrović, M., Petrović, M.: Quantum criticality in photorefractive optics: vortices in laser beams and anti-ferromagnets. *Phys. Rev. A* **96**, 053824 (2017). [arXiv:1701.03451](https://arxiv.org/abs/1701.03451) [physics.optics]
- Denz, C., Schwab, M., Weinau, C.: *Transverse Pattern Formation in Photorefractive Optics*. Springer, Berlin (2003)
- Dreisruh, E., Chervonkov, S., Neshev, D., Paulus, G.G., Walther, H.: Generation of lattice structures of optical vortices. *J. Opt. Soc. Am. B* **19**, 550–556 (2002)
- Fetter, A.L.: Rotating trapped Bose–Einstein condensates. *Rev. Mod. Phys.* **81**, 647–691 (2009)
- Fleischer, J.W., Bartal, G., Cohen, O., Manela, O., Segev, M., Hudock, J., Christodoulides, D.N.: Observation of vortex-ring “discrete” solitons in 2D photonic lattices. *Phys. Rev. Lett.* **92**, 123904 (2004)
- Ghofraniha, N.: Experimental evidence of replica symmetry breaking in random lasers. *Nat. Commun.* **6**, 6058 (2015). arxiv.org/abs/1407.5428 [cond-mat.dis-nn]

- Jović, D.M., Petrović, M.S., Belić, M.R.: Counterpropagating pattern dynamics: from narrow to broad beams. *Opt. Commun.* **281**, 2291–2300 (2008)
- Kosterlitz, J., Thouless, D.: The Kosterlitz–Thouless phase in a hierarchical model. *J. Phys. C* **6**, 1181–1203 (1973)
- Ma, X., Driben, R., Malomed, B., Meoer, T., Schumacher, S.: Two-dimensional symbiotic solitons and vortices in binary condensates with attractive cross-species interaction. *Nat. Sci. Rep.* **6**, 34847 (2016). [arXiv:1606.08579](https://arxiv.org/abs/1606.08579) [physics.optics]
- Malomed, B., Torner, L., Wise, F., Mihalache, D.: On multidimensional solitons and their legacy in contemporary atomic, molecular and optical physics. *J. Phys. B: At. Mol. Opt. Phys.* **49**, 170502 (2016)
- Neshev, D.N., Alexander, T.J., Ostrovskaya, E.A., Kivshar, Y.S.: Observation of discrete vortex solitons in optically induced photonic lattices. *Phys. Rev. Lett.* **92**, 123903 (2004)
- Perret, A., Ristivojevic, Z., Le Doussal, P., Schehr, G., Wiese, K.: Super-rough glassy phase of the random field XY model in two dimensions. *Phys. Rev. Lett.* **109**, 157205 (2012). [arXiv:1204.5685](https://arxiv.org/abs/1204.5685) [cond-mat.dis-nn]
- Petrović, M.S., Belić, M.R., Denz, C., Kivshar, YuS: Counterpropagating optical beams and solitons. *Lasers Photon. Rev.* **5**, 214–233 (2011). [arXiv:0910.4700](https://arxiv.org/abs/0910.4700) [physics.optics]
- Petrović, M., Jović, D., Belić, M., Schröder, J., Jander, P., Denz, C.: Two dimensional counterpropagating spatial solitons in photorefractive crystals. *Phys. Rev. Lett.* **95**, 053901 (2005)
- Rabinovich, M.I., Ezersky, A.B., Weidman, P.D.: *The Dynamics of Patterns*. World Scientific, Singapore (2000)
- Rechtsman, M.C., Zeuner, J.M., Plotnik, Y., Lumer, Y., Podolsky, D., Dreisow, F., Nolte, S., Segev, M., Szameit, A.: Photonic Floquet topological insulators. *Nature* **496**, 196–200 (2013). [arXiv:1212.3146](https://arxiv.org/abs/1212.3146) [physics]
- Sandfuchs, O., Kaiser, F., Belić, M.R.: Self-organization and Fourier selection of optical patterns in a nonlinear photorefractive feedback system. *Phys. Rev. A* **64**, 063809 (2001)
- Silae, M., Babev, E.: Microscopic derivation of two-component Ginzburg–Landau model and conditions of its applicability in two-band systems. *Phys. Rev. B* **85**, 134514 (2012). [arXiv:1110.1593](https://arxiv.org/abs/1110.1593) [cond-mat.supr-con]
- Terhalle, B., Richter, T., Desyatnikov, A.S., Neshev, D.N., Krolikowski, W., Kaiser, F., Denz, C., Kivshar, Y.S.: Observation of multivortex solitons in photonic lattices. *Phys. Rev. Lett.* **101**, 013903 (2008)

P r o c e e d i n g s

of the

10th MATHEMATICAL PHYSICS MEETING:

**School and Conference on
Modern Mathematical Physics**

September 9–14, 2019, Belgrade, Serbia

Editors

B. Dragovich, I. Salom and M. Vojinović

Institute of Physics

Belgrade, 2020

SERBIA

Autor: Grupa autora

Naslov: 10th MATHEMATICAL PHYSICS MEETING: SCHOOL AND CONFERENCE ON MODERN MATHEMATICAL PHYSICS (Deseti naučni skup iz matematičke fizike: škola i konferencija iz savremene matematičke fizike)

Izdavač: Institut za fiziku, Beograd, Srbija

Izdanje: Prvo izdanje
(SFIN year XXXIII Series A: Conferences, No. A1 (2020))

Štampa: Zemunplast, Beograd

Tiraž: 150

ISBN: 978-86-82441-51-9

1. Dragović Branko
Matematička fizika-Zbornici

**CIP – Каталогизација у публикацији
Народна библиотека Србије, Београд**

51-7:53(082)

MATHEMATICAL Physics Meeting: School and Conference on Modern Mathematical Physics (10 ; 2019 ; Beograd)

Proceedings of the 10th Mathematical Physics Meeting: School and Conference on Modern Mathematical Physics, September 9-14, 2019, Belgrade, Serbia / [organizers Institute of Physics, Belgrade ... [et al.]] ; editors B. [Branko] Dragovich, I. [Igor] Salom and M. [Marko] Vojinović. - 1. izd. - Belgrade : Institute of Physics, 2020 (Beograd : Zemunplast). - XII, 383 str. : ilustr. ; 28 cm. - (SFIN ; year 33. Ser. A, Conferences, ISSN 0354-9291 ; n° A1, (2020))

Nasl. u kolofonu: Deseti naučni skup iz matematičke fizike: škola i konferencija iz savremene matematičke fizike. - Tiraž 150. - Str. VII: Preface / editors. - Napomene i bibliografske reference uz radove. - Bibliografija uz svaki rad.

ISBN 978-86-82441-51-9

1. Dragović, Branko, 1945- [urednik] [аутор додатног текста]
а) Математичка физика -- Зборници

COBISS.SR-ID 13561865

CONTENTS

Review and Research Works

D. Benisty, E. Guendelman, E. Nissimov and S. Pacheva Non-Riemannian volume elements dynamically generate inflation	1
F. Bulnes Baryogenesis until dark matter: H-particles proliferation	15
M. Burić and D. Latas Singularity resolution in fuzzy de Sitter cosmology	27
D. J. Cirilo-Lombardo Dynamical symmetries, coherent states and nonlinear realizations: the $SO(2,4)$ case	37
M. Čubrović Fermions, hairy blackholes and hairy wormholes in anti-de Sitter spaces	59
Lj. Davidović, I. Ivanišević and B. Sazdović Courant and Roytenberg bracket and their relation via T-duality	87
Lj. Davidović and B. Sazdović T-duality between effective string theories	97
M. Dimitrijević Ćirić Nonassociative differential geometry and gravity	111
S. Giaccari and L. Modesto Causality in nonlocal gravity	121
J. Leech, M. Šuvakov and V. Dmitrašinović Hyperspherical three-body variables applied to Sakumichi and Suganuma's lattice QCD data	137
N. Manojlović, I. Salom and N. Cirilo António XYZ Gaudin model with boundary terms	143

S. Marjanović and V. Dmitrašinović	
Numerical study of classical motions of two equal-mass opposite-charge ions in a Paul trap	161
A. Miković	
Piecewise flat metrics and quantum gravity	167
D. Minić	
From quantum foundations of quantum field theory, string theory and quantum gravity to dark matter and dark energy	183
M. Mintchev and P. Sorba	
Entropy production in systems with spontaneously broken time-reversal	219
B. Nikolić and D. Obrić	
From 3D torus with H-flux to torus with R-flux and back	233
T. Radenković and M. Vojinović	
Construction and examples of higher gauge theories	251
I. Salom, N. Manojlović and N. Cirilo António	
The spin 1 XXZ Gaudin model with boundary	277
D. Simić	
Velocity memory effect for gravitational waves with torsion	287
O. C. Stoica	
Chiral asymmetry in the weak interaction via Clifford algebras	297
M. Stojanović, M. Milošević, G. Đorđević and D. Dimitrijević	
Holographic inflation with tachyon field as an attractor solution	311
F. Sugino	
Highly entangled quantum spin chains	319
M. Szcząchor	
Two type of contraction of conformal algebra and the gravity limit	331

M. Szydłowski, A. Krawiec and P. Tambor Can information criteria fix the problem of degeneration in cosmology?	339
V. Vanchurin A quantum-classical duality and emergent space-time	347
O. Vaneeva Transformation properties of nonlinear evolution equations in 1+1 dimensions	367
Talks not included in the Proceedings	377
List of participants	381

Fermions, hairy blackholes and hairy wormholes in anti-de Sitter spaces*

Mihailo Čubrović[†]

Center for the Study of Complex Systems,
Institute of Physics Belgrade, Serbia

ABSTRACT

We discuss the existence, properties and construction (analytical and numerical) of hairy black holes with fermionic matter in asymptotically anti-de-Sitter space. The negative cosmological constant makes hairy black holes stable, and the nucleation mechanism can make the formation of hair at the horizon energetically and entropically preferable to conventional black holes. The difficulties intrinsic to fermions at finite density – the Pauli principle and exchange interactions – require some drastic approximations in calculating the stress-energy tensor and geometry. We will consider several methods on the market – Hartree-Fock, WKB, and fluid-mechanical methods, and consider the dual field theories of these constructions. Then we will apply the same methods to the construction of wormholes; fermions are a natural candidate for wormhole source matter as they have a Dirac sea of negative energies, and negative energy-momentum density is the condition for wormhole formation. The field theory interpretation of wormholes is still open but has to do with strongly entangled systems. The paper combines a pedagogical introduction to the basic methods and results (obtained in the last 10+ years) with an account of fresh research results, mainly on the wormhole applications and non-planar black holes.

1. Introduction

AdS black holes are a favorite topic, not only in relation to holography but also in general: AdS space behaves like a potential box, the cosmological constant provides an effective repulsive force at large distances and the existence of a boundary at spatial infinity makes bound states possible. All of

*The author acknowledges funding provided by the Institute of Physics Belgrade through the grant by the Ministry of Education, Science and Technological Development. The author acknowledges the use of the Sci-Hub service. The results described here would never have been possible without the teaching, help and collaboration from Jan Zanen, Koenraad Schalm, Yan Liu, Ya-Wen Sun, Elena Gubankova, Mariya Medvedyeva, Vladan Djukić and Nicolas Chagnet.

[†] e-mail address: mcubrovic@gmail.com

this brings about the famous result that hairy black holes are indisputably possible, and well-studied. In full (global) AdS space, one may have small black holes, which barely see the boundary and radiate like in asymptotically flat space, and large black holes, which reach an equilibrium state with the Hawking radiation at given temperature and remain stable forever (eternal AdS black holes). We will focus on the latter, as they can be treated as (semi)classical stationary systems. Clearly, just like the Hawking radiation, matter and gauge fields can likewise equilibrate between the black hole horizon and AdS boundary, possibly forming hair – by definition, it means nonzero density of some field (and possibly nonzero expectation values of other operators, like charge density, spin, etc) at the horizon itself. This in turn means that the geometry changes as opposed to the no-hair case: the hair itself enters the stress-energy tensor, and the outcome is a hairy black hole geometry, where a horizon still exists but with a different metric. At zero temperature, hair tends to remove extremal black holes in favor of zero-area horizons, with zero Bekenstein-Hawking entropy. We will soon discuss several explicit examples of this phenomenon.

The above story acquires an additional dimension thanks primarily to the AdS/CFT correspondence (gauge/gravity duality) [1, 2, 3] – the fact that the bulk gravity physics is equivalent to a quantum field theory in flat space in one dimension less, whose operators act as boundary sources of the AdS (bulk) fields. The actions in AdS (with field Φ) and in CFT (with field \mathcal{O} , which acts as a boundary source to Φ) are equal:

$$\begin{aligned}
 S_{\text{AdS}} &= S_{\text{CFT}} \\
 S_{\text{AdS}} &= \int \mathcal{D}\Phi \exp \left(- \int_{\text{AdS}} d^{D+1}x \sqrt{-g} \mathcal{L}_{\text{AdS}}(\Phi, \partial_\mu \Phi) + \oint_{\partial} d^D x \sqrt{-h} \mathcal{O} \Phi \right) \\
 S_{\text{CFT}} &= \int \mathcal{D}\mathcal{O} \exp \left(- \int d^D x \mathcal{L}_{\text{CFT}}(\mathcal{O}) \right), \tag{1}
 \end{aligned}$$

where we have denoted by ∂ the boundary of the AdS space, $g_{\mu\nu}$ is the AdS metric and $h_{\mu\nu}$ is the induced metric at the boundary. From now on, integrals over the bulk of AdS will be denoted just by \int , understanding that the integral is over the whole space. At this place we do not intend to explain AdS/CFT and its applications in any detail; suffice to say that one can obtain thermodynamic potentials and correlation functions in field theory, which has found important applications in condensed matter theory, quantum chromodynamics and conformal field theory. Interested readers can consult [4, 5, 6] for reviews. In this work we deal with the gravity side of the correspondence – the formation of a hairy black hole with fermionic matter, which corresponds to a finite electron density phase in field theory. We assume the familiarity with the basic notions of AdS space and quantum field theory in curved spacetime, for example at the level of [7] and [8], respectively.

Mathematically, the topic of this review is the solution of the coupled Einstein-Maxwell-Dirac system with the total action $S_{\text{AdS}} = S_{\text{bulk}} + S_{\partial}$.

The bulk action reads:

$$\begin{aligned}
 S_{\text{bulk}} &= S_{\text{E}} + S_{\text{M}} + S_{\text{Dir}} \\
 S_{\text{E}} &= \int d^4x \sqrt{-g} (R + 6) \\
 S_{\text{M}} &= - \int d^4x \sqrt{-g} \frac{\hat{F}^2}{4} \\
 S_{\text{Dir}} &= - \int d^4x \sqrt{-g} \left(\frac{1}{2} \bar{\Psi} D_\mu e_a^\mu \Gamma^a \Psi + \frac{1}{2} \bar{\Psi} e_a^\mu \Gamma^a \Psi + m \bar{\Psi} \Psi \right). \quad (2)
 \end{aligned}$$

Here, $\hat{F}_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic (EM) field strength tensor, and the cosmological constant in AdS₄ is $6/L^2$, where the AdS radius $L = 1$ is set to unity, as we will mainly work on the Poincare patch of AdS space, so all other dimensionful quantities can be expressed in terms of L . The Dirac bispinor Ψ has mass m and charge q , and the covariant derivative

$$D_\mu = e_\mu^a D_a = \partial_\mu - \frac{i}{8} [\Gamma^a, \Gamma^b] \omega_{ab}^\mu - iq A_\mu \quad (3)$$

depends on the spin connection ω_{ab}^μ and the gauge field A_μ , and the gamma matrices satisfy the usual relations $[\Gamma^a, \Gamma^b] = 2\eta^{ab}$, with the Minkowski metric η . We will be using the mostly plus convention. Obviously, $\Psi = 0$ is a solution, and in this case we get a Schwarzschild black hole if the EM field is also zero, or a charged Reissner-Nordstrom (RN) black hole for nonzero field strength. The question is, are there other solutions, with nonzero profile Ψ ? Such solutions describe hairy black holes at finite temperature: the horizon is typically still there, but the geometry is changed. At zero temperature, the black hole might disappear. Since AdS space has a boundary, there is also a boundary contribution to the action, as in (1), depending on extrinsic curvature K , boundary cosmological constant λ and the boundary values of the fields:

$$S_\partial = \oint_\partial d^3x \sqrt{-h} \left[K - \lambda - \frac{1}{2} n_\mu A_\nu \hat{F}^{\mu\nu} - \frac{1}{2} \bar{\Psi} \Psi \right]. \quad (4)$$

The classical equations of motion do not depend on the boundary action. However, S_∂ is still important (1) to make sure there is a good action principle, i.e., that the on-shell solutions are indeed minima of the action¹ (2) to regularize any UV divergences (3) to get correct thermodynamics. The last point will be particularly important: one way to see that the hairy black hole and not the bald black hole is the true vacuum will be the fact that the action on the hairy solution is lower.

Solving the system (2) is a problem in quantum field theory at finite density. We work with classical general relativity (GR) and classical EM

¹Remember that the (bulk) Euler-Lagrange equations are only a necessary condition for the minimum of the action.

field, but *fermions are never classical*; this is the first important lesson. The Pauli principle always introduces nonlocal correlations which show as the exchange interaction. Another way of saying this is that the pressure of a fermionic gas or fluid always includes the quantum contribution which is absent in both classical and bosonic gas; that is the reason that organized matter such as stars, planets, chairs and notebooks has rigidity and does not collapse onto itself. Therefore, even though we do gravity at $\hbar = 0$, the fermions even at leading order need to be tackled quantum-mechanically. This means calculating the *fermionic determinant*:

$$Z_{\text{Dir}} = \int \mathcal{D}\bar{\Psi}\mathcal{D}\Psi e^{-S_{\text{Dir}}} = [\det(D_\mu e_a^\mu \Gamma^a + e_a^\mu \Gamma^a D_\mu + m)]^{1/2}. \quad (5)$$

We have put the equality sign under quotation marks because the determinant is actually the product of the eigenenergies of all the modes (an infinity of them), which is not only badly divergent (that could be regulated) but is also impossible to calculate because of the *fermion sign problem*, the fact that the fermionic modes enter the path integral with a sign that can be plus or minus. This makes the measure in the path integral (5) non-probabilistic and makes it impossible to expand around a classical solution in a controlled way. Fortunately, the AdS metric turns out to simplify the problem enough that it can be tackled in a way which is tractable and, while of course not exact, can be systematically improved in a perturbative way. This is in fact the motivation behind AdS/CFT modelling of strongly correlated electron systems: the fermion sign problem is fatal for strongly coupled field theories in flat space, but in GR with AdS boundary conditions it transforms into a difficult but doable task.

Is the journey worthwhile? In line with the broad scope of the Belgrade Mathematical Physics Meetings, we have anticipated a broad readership of this paper and thus we have decided to give a very general and perhaps rather dry introduction to the topic of fermionic hairy black holes. This necessarily means that we will not touch upon the many interesting applications: AdS/CFT and its applications to quantum chromodynamics and condensed matter physics, the black hole information problem, the critical phenomena in gravitational collapse and the black hole solutions in string theory. We do discuss one special topic that we currently find very interesting: hairy wormholes generated by fermion matter, where many of the methods used for hairy black holes can be successfully applied. The main task of the paper is to provide a tutorial on the basic methodology and calculation techniques, bringing the reader to the point that he can understand and repeat the calculations from the literature and start doing his own. The existing literature is rather heterogenous and there is no single text to recommend. We will give the references we deem particularly useful throughout the paper, without the pretention of being exhaustive; the choice of references is certainly dictated also by our prejudices and tastes.

Plan of the paper. In Section 2 we first explain the instabilities of AdS space and AdS black holes to a nonzero density profile of fermions, and in-

roduce the basic concepts that will keep appearing throughout the paper: effective potential and the bound states of the fermionic wavefunctions. In Section 3 we first treat the problem in the consistent one-loop (Hartree-Fock) approximation, calculating the determinant (5) by definition, from the individual wavefunctions for different states. We find this job surprisingly difficult – it is still an active research area. But we are able to give a qualitative picture of the outcome and sketch the phase diagram, depending on the chemical potential μ and fermion mass and charge m, q . As we move toward the high-fermion-density corner of the phase diagram, the things simplify. The simplest and "most classical" limit of the problem is the limit of large density. It is a rule of thumb that for fermions, the role of interactions diminishes as the density grows. At high density, WKB approximation works very well. At highest densities, we find semiclassical fluid with an equation of state that takes into account the fermionic pressure, similar in spirit to the Oppenheimer-Volkov equations for neutron stars. In section 4 we apply these methods to a different topic – hairy wormholes instead of black holes. This problem has recently gained notoriety and might carry some important messages for the black hole information problem. The final section sums up the conclusions.

2. Planar AdS black holes and fermion nucleation

In this and the next section we will focus on large planar black holes on the Poincare patch of AdS space. Large black holes can reach equilibrium with the AdS boundary so they do not emit Hawking radiation and can exist eternally. The Poincare patch of AdS₄ space is a coordinate chart with a single boundary on one end and interior on the other end. It does not cover the whole AdS space but is simpler to work with than global AdS and is good enough to describe the instability at the horizon. The metric of pure AdS space without a black hole is given by

$$ds^2 = r^2 (-dt^2 + d\vec{x}^2) + \frac{dr^2}{r^2} = \frac{1}{z^2} (-dt^2 + d\vec{x}^2 + dz^2) \quad (6)$$

where $r = 1/z$ is the radial coordinate, t is time and $\vec{x} = (x, y)$ are the transverse coordinates. The AdS boundary is at $r = \infty$ ($z = 0$), and the interior is at $r = 0$ ($z = \infty$). From now on we will mainly use the z coordinate; we will always specify explicitly if a different radial coordinate is used. In AdS/CFT, the radial coordinate corresponds to the energy scale in field theory: the near-boundary region encodes for the physics at high energies, in the ultraviolet (UV), and the deep interior, with z large, is the infrared (IR). Even though we do not consider the CFT dual here, we will still adopt the UV/IR terminology.

In the presence of a point electric charge e we get a Reissner-Nordstrom (RN) black hole with the horizon at $z_h = 1$, with charge e , mass M and

temperature T :

$$ds^2 = \frac{1}{z^2} \left(-f(z)dt^2 + d\vec{x}^2 + \frac{dz^2}{f(z)} \right), \quad f(z) = 1 - Mz^3 + e^2z^4$$

$$M = z_h^3 + e^2, \quad A = \frac{ez_h}{2\sqrt{\pi}}(1 - z/z_h)dt, \quad T = \frac{3z_h}{4\pi} \left(1 - \frac{e^2}{3z_h^4} \right) \quad (7)$$

For $e = 0$ we get the Schwarzschild AdS black hole, and for $e = \sqrt{3}z_h^2$ the black hole becomes extremal, with temperature $T = 0$. To see this, remember that the black hole temperature is given by $f(z \rightarrow z_H) = 4\pi T(z - z_H) + \dots$, so plugging in f from above we indeed get the correct expression for T . Importantly, the near-horizon region of a black hole is an AdS space [7]. This IR AdS space (near $z = z_h$) has a priori nothing to do with the AdS asymptotics in the UV (near $z = 0$); it is there also for black holes in flat or dS space. At $T = 0$, rescaling $z - e/\sqrt{3} \mapsto 1/6\epsilon\xi$ and expanding in ϵ to lowest order gives the metric

$$ds^2 = \frac{1}{6}(-dt^2 + d\xi^2) + \frac{e^2}{3}d\vec{x}^2, \quad A_t = \frac{1}{\sqrt{6}\xi}. \quad (8)$$

The is $\text{AdS}_2 \times R^2$ geometry, a direct product of AdS with a plane. At finite temperature, a similar rescaling can be worked out, yielding again an AdS_2 throat. Since the throat describes the near-horizon region, instabilities of the black hole can be figured out from possible instabilities of this IR AdS space. Once again, this is *not* the whole AdS_4 , which is always stable far from the horizon, in the UV (otherwise our whole classical gravity approach crumbles down), it is just a region near the horizon, in IR.

In order to write the equations of motion, we have to choose a basis for the gamma matrices and the form of the Dirac bispinor (remember that only two out of four components are really independent degrees of freedom). A convenient representation is

$$\Gamma^0 = \sigma^1 \otimes i\sigma_2, \quad \Gamma^1 = \sigma^1 \otimes \sigma_1, \quad \Gamma^2 = \sigma^1 \otimes \sigma^3, \quad \Gamma^z = \sigma^3 \otimes \hat{1}. \quad (9)$$

so that the Dirac equation in a spherically symmetric metric defined as $\text{diag}(g_{tt}, g_{ii}, g_{ii}, g_{zz})$ gives two equivalent decoupled pairs of equations. Taking the Dirac bispinor in the form $\Psi = (\psi_1, \chi_1, i\chi_2, i\psi_2)^T$, the equations for $\psi_{1,2}$ read [9, 10]:²

$$\partial_z \psi_{1,2} \pm \hat{m} \psi_{1,2} - (\mp \hat{E} + \hat{k}) \psi_{2,1} = 0 \quad (10)$$

$$\hat{m} \equiv m\sqrt{g_{zz}}, \quad \hat{\mu} \equiv \sqrt{\frac{g_{zz}}{-g_{tt}}} A_t, \quad \hat{E} \equiv q\hat{\mu} + E\sqrt{\frac{g_{zz}}{-g_{tt}}}, \quad \hat{k} \equiv \sqrt{\frac{g_{zz}}{g_{ii}}} k. \quad (11)$$

²Since only two components of the Dirac bispinor are independent, the system for $\chi_{1,2}$ yields no new information.

We have Fourier-transformed the derivatives over time and transverse spatial dimensions as $\partial_t = -i\omega$, $\partial_x = ik_x$, $\partial_y = ik_y$, and we have exploited the spherical symmetry to set $k_x = k$, $k_y = 0$. The quantities \hat{E} , \hat{k} , $\hat{\mu}$ can be informally interpreted as "local" values of the energy, momentum and chemical potential, respectively. The "local" values equal E , k , μ at the AdS boundary, grow monotonously toward the horizon and diverge there, a consequence of the infinite redshift seen by a faraway observer. An important idea is to consider the Schrödinger form of the Dirac equation instead, differentiating (10) once with respect to z , decoupling the equations for $\psi_{1,2}$, and eliminating the first derivatives $\psi'_{1,2}$ by introducing the tortoise coordinate s instead of z . The resulting picture is that of a zero-energy Schrödinger equation, of the form $\partial_s^2 \psi_{1,2} - V_{\text{eff}}(s)\psi_{1,2} = 0$, in an effective potential $V_{\text{eff}}(s)$.³ Near the horizon, the potential is constant at leading order [11]:

$$V_{\text{eff}}(s \rightarrow -\infty) = \frac{m^2 + 12k^2/\mu^2 - 2q^2}{(q/\sqrt{2} + k)^2} + \dots \quad (12)$$

It is true that the Schrödinger form is only a consequence of the Dirac equation, not equivalent to it: extra conditions must be imposed on the Schrödinger solution to make it satisfy the Dirac equation. But the effective potential is great for qualitative insights and it contains the basic idea of the black hole instability in a very transparent way. The near-horizon potential can contain bound states if it is negative, hence the instability criterion for a fermionic mode with momentum k is that the numerator of (12) is negative. Fermions fill up the potential well starting from $k = 0$ up to some maximum k for which (12) reaches zero. Therefore, the instability first sets in when V_{eff} is negative for $k \rightarrow 0$, so we get our first rule-of-thumb prediction: the black hole will be surrounded by a gas of fermions and become hairy when

$$m < q\sqrt{2}. \quad (13)$$

But this is just one end of the potential well; what happens at the other end? Plugging in the pure AdS metric (6) into (12) we get

$$V_{\text{AdS}}(s \rightarrow 0) = \frac{m^2 + m + k^2}{(k + \mu)^2} \frac{1}{s^2} + \dots, \quad (14)$$

which is always non-negative, and grows to infinity. This is good – there is never an instability in the far UV, and the fermionic hair can never come arbitrarily close to the AdS boundary. It also means that bound states in the interior will indeed exist whenever (13) is negative. The physical picture is the following: in the presence of EM and gravitational field of the black hole, fermions are pair-created. These pairs are virtual, and

³This is a simple exercise that we will do many times; the reader should be able to do the necessary (straightforward) calculations leading to the expression for $s(z)$ and $V_{\text{eff}}(s)$.

they only have a finite probability of becoming long-living if the external potential energy is large enough. In that case, bound states form, and there is a solution of (2) with nonzero fermion density. In the literature, this is sometimes called fermion nucleation. For scalars, similar logic leads to the Breitenlohner-Freedman bound, which puts a constraint on the scalar mass for the stability of the UV (with fermions, as we have seen, UV is always stable), and in IR it similarly gives a criterion for forming hair [5]. We also see from (14) and Fig. 1 that the potential well becomes shallower as k grows, so the bound states only exist up to some maximum $k = k_F$ which is really the Fermi momentum of the bulk Fermi sea.

From (12,14) we can understand the behavior of the effective potential. In Fig. 1, we give the function $V(s)$ in the whole space, from $z = 0$ ($s = 0$), to $z = z_h$ ($s = -\infty$). The fermionic modes fill the potential well until they reach the energy $E = 0$. From (12), higher modes correspond to higher momentum k . The fermionic density is thus given by a sum over these bound states. The easiest case is in fact an extremely deep well: the energy levels are so dense and so numerous that they can be approximated by a continuum; this is called electron star limit. But the most interesting regime is the one with only a few wavefunctions, which really describes the transition to a hairy solution. This is a much harder nut to crack.

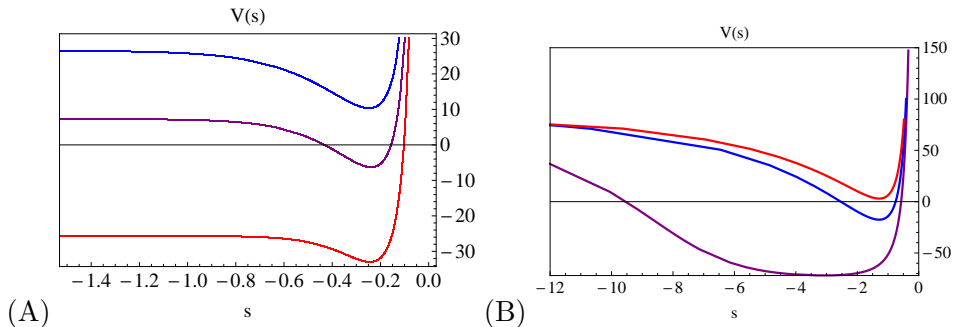


Figure 1: Effective potential V_{eff} , as a function of the tortoise coordinate $s \equiv \int_0^z dz g_{zz}(z)$, in the RN metric (A), and in the Lifshitz metric (B), for $q = 1$, $m = 0.4$, $\mu = 1$, and three momentum values increasing from violet to blue to red: $k = 1, 2, 3$ (A) and $k = 0, 5, 10$ (B). In both cases, the negative potential well becomes shallower and shallower and eventually disappears as k grows, so we fill the bulk Fermi sea up from $k = 0$ to some maximal $k = k_F$. In the black hole background, the potential is flat for $s \rightarrow -\infty$, which corresponds to the AdS_2 near-horizon region and signifies an instability as the bound states extend all the way to the horizon ($s = -\infty$). In the backreacted Lifshitz metric the potential grows for $s \rightarrow -\infty$, suggesting that deep IR is stable: the true vacuum is the Lifshitz geometry, not RN. Taken over from [14].

3. Fermionic hair

Now that we have convinced ourselves that hairy solutions, with finite fermion density, have to exist, we need to solve the full system of Einstein, Maxwell and Dirac equations to find them. Clearly, a more general ansatz for the metric than (7) is needed now, and we will write it as

$$ds^2 = -\frac{F(z)G(z)}{z^2}dt^2 + \frac{1}{z^2}d\vec{x}^2 + \frac{1}{F(z)z^2}dz^2, \quad (15)$$

leading to Einstein-Maxwell equations

$$1 - F + zF'/3 - T_{tt}^{\text{tot}}FG/3z^2 = 0 \quad (16)$$

$$G' + z(T_{tt}^{\text{tot}}/F^2 + T_{zz}^{\text{tot}}G) = 0 \quad (17)$$

$$A_t'' - G'/2GA_t' + qn\sqrt{G}/\sqrt{F}z^3 = 0, \quad (18)$$

where $T_{\mu\nu}^{\text{tot}}$ is the total stress-energy tensor, both from the electric field (which is easy to find) and from the fermions (which is our big problem). A typical situation in hairy problems is that formulating the physically meaningful boundary conditions is not so easy. Notice the Einstein equations are first-order, so we need one boundary condition for each function (F and G), whereas the Maxwell equation is second-order and requires two boundary conditions. Let us now summarize what boundary behavior we expect on physical grounds.

1. The AdS asymptotics for the metric and gauge field require $F(z \rightarrow 0), G(z \rightarrow 0) = 1, A_t(z \rightarrow 0) = \mu$. So far it's all simple.
2. The main puzzle for the IR geometry is – does the horizon disappear or not? At $T = 0$ we do not expect that the degenerate RN horizon can survive. So we do not expect zeros in F, G but we do expect their derivatives to vanish in order to have a smooth solution (finite derivatives at $z \rightarrow \infty$ would likely give divergent curvature). Thus at $T = 0$ we need $F'(z \rightarrow \infty) = G'(z \rightarrow \infty) = 0$ or, in other words, $F(z \rightarrow \infty) = \text{const.} + O(1/z)$ and likewise for G . At finite temperature, general GR arguments suggest there is a horizon at some $z = z_h$ satisfying $F'(z \rightarrow z_h) = 4\pi T$.
3. The IR behavior of the gauge field is related to the question: *is all the charge carried by the fermions, or the charge is shared between the fermions and the horizon?* The Gauss-Ostrogradsky theorem for the AdS space, with a UV boundary and either a horizon or a smooth far-away IR takes the form [12]:

$$\oint_{\partial} d^3x \sqrt{-h}|_{z \rightarrow 0} \star \hat{F} = \int d^4x \sqrt{-g} qn + \oint_{\text{IR}} d^3x \sqrt{-h_{\text{IR}}}|_{z=z_{\text{IR}}} \star \hat{F} \quad (19)$$

Here, $\star \hat{F}$ is the coordinate-invariant flux of the 2-form \hat{F} , and h_{IR} is the induced metric on the surface normal to the radial direction at

$z_{\text{IR}} = z_h$ or $z_{\text{IR}} = \infty$, depending on whether there is a horizon or not. In principle, the IR charge might be shared between the horizon and the fermions. However, we will find that in the semiclassical calculation there are no solutions where the charge is shared – any backreaction will always expell all the charge from the IR.

4. The boundary conditions for the Dirac equation present no problems and are pretty standard in AdS space [13]. In the UV, out of the two branches, we want the subleading one, with the motivation to preserve the AdS asymptotics, i.e., to perturb the space as little as possible in the UV. In particular, the near-boundary expansion of (11) gives

$$\begin{aligned}\psi_1(z \rightarrow 0) &= \frac{E + \mu q - k}{2m - 1} A_2 z^{5/2-m} + B_1 z^{3/2+m} + \dots \\ \psi_2(z \rightarrow 0) &= A_2 z^{3/2-m} + \frac{E + \mu q + k}{2m + 1} B_1 z^{5/2+m} + \dots, \quad (20)\end{aligned}$$

so we pick $A_2 = 0$, as the leading contribution for $z \rightarrow 0$ comes from the $z^{3/2}$ term. In the IR, the metric determines the boundary conditions: if there is a horizon, we need $\Psi(z = z_h) \rightarrow 0$ for stability, if not, then to avoid infinite energy density at large z we require $\partial_z \Psi(z \rightarrow \infty) = 0$, for otherwise a nonconstant density profile would give rise to a diverging curvature. The attentive reader should be alarmed: this means two boundary conditions for each component (one in UV and one in IR), but the equations are only first-order. The resolution is that for given momenta, the energy is not arbitrary but fixed by the dispersion relation $E(k)$; thus solving the Dirac equation in an effective potential well introduces energy quantization, as one would expect.

What remains is to find the fermionic stress tensor. Since spinors couple to the spin connection e_a^μ and not directly to the metric, the stress tensor is expressed as

$$T_{\mu\nu} = \left\langle \frac{1}{4} e_{\mu\alpha} \bar{\Psi} \Gamma^a D_\nu \Psi + (\mu \leftrightarrow \nu) \right\rangle, \quad (21)$$

and the expectation value $\langle \dots \rangle$ reminds us that the fermions are never classical. At zero temperature, the state is pure and can be represented as the sum of (appropriately normalized) radial modes with energies E_ℓ , where ℓ is the radial quantum number, and the energies E_ℓ are all ≤ 0 . At finite temperature, the state is mixed and gets a contribution from both positive and negative energies E_ℓ , with thermal weights $w_\ell = \exp(-\beta E_\ell)/Z$, the partition sum being $Z = \sum_\ell \exp(-\beta E_\ell)$. With this in mind, we can write out (21) as

$$T_{tt} = e_{t0} \sum_{\ell=1}^N w_\ell \int_0^{k_F} \frac{k dk}{(2\pi)^2} \left(\psi_{1;\ell}^\dagger \psi_{1;\ell} + \psi_{2;\ell}^\dagger \psi_{2;\ell} \right) (E_\ell + q A_t)$$

$$\begin{aligned}
 T_{ii} &= e_{i1} \sum_{\ell=1}^N w_{\ell} \int_0^{k_F} \frac{k dk}{(2\pi)^2} \left(\psi_{1;\ell}^{\dagger} \psi_{1;\ell} - \psi_{2;\ell}^{\dagger} \psi_{2;\ell} \right) k \\
 T_{zz} &= e_{z3} \sum_{\ell=1}^N w_{\ell} \int_0^{k_F} \frac{k dk}{(2\pi)^2} \left(\psi_{1;\ell}^{\dagger} \partial_z \psi_{2;\ell} - \psi_{2;\ell}^{\dagger} \partial_z \psi_{1;\ell} \right). \quad (22)
 \end{aligned}$$

For brevity, we write $\psi_{1,2;\ell} \equiv \psi_{1,2}(E_{\ell}, k; z)$. We will consider in detail just the $T = 0$ case, when the weights w_{ℓ} effectively just pick the ground state and cut off all the others, but we will later discuss the results (without details of the calculations) also at finite T . The spectrum is discrete and gapped in the radial direction, so the integral $\int dE/2\pi$ becomes a sum, however in the transverse directions the system remains gapless, filling the whole (spherical) Fermi sea in the k -momentum space, as long as the dispersion relation $E(k) = E_{\ell} \leq 0$ is satisfied for some ℓ . The highest such k , for which $E_{\ell} = 0$, is the Fermi momentum k_F , and the possible momenta are $0 \leq k \leq k_F$. It is this continuous quantum number k that makes our life difficult. Here, indeed, our easy path comes to an end, because a self-consistent calculation of the wavefunctions certainly cannot be done in a closed form. Here we must resort to approximations. The number of occupied levels N is a good guide on the kind of approximation one needs to make. One can rephrase it as the ratio Q/q , where Q is the total fermion charge $\int d^4x \sqrt{-g} q \Psi^{\dagger} \Psi$. The thermodynamic limit, where the number of particles goes to infinity and the charge of an individual fermion to zero so that $N \rightarrow \infty, q \rightarrow 0, Q = Nq = \text{const.}$, is at one extreme. We expect that the problem approaches the classical regime in this case, and it will turn out to be true. The opposite limit is $Q/q = 1$, with just a single excitation, the hairy black hole at birth. We expect this to be likewise a simple limit, however it will turn out not to be quite true. In-between we dial between the quantum mechanics of $N = 1$ and the classical field theory of $N \rightarrow \infty$ [14].

Phase diagram. Before doing that, we can sum up our qualitative knowledge on a phase diagram (Fig. 2). From (12-14), bound states form for small enough m values (panel (A)); if (13) is valid beyond the probe approximation, the borderline is $m = q\sqrt{2}$. Left of this line there is a hairy solution, to the right of it the AdS₂ near-horizon region (and the whole RN black hole) remain. The hairy solutions are best described in different ways depending on the number of filled levels ($N = Q/q$); this is the topic of the rest of this section. One can also plot the situation at finite temperature (panel (B)). The phases remain the same; more precisely, the extremal black hole becomes a finite-temperature black hole, and the hairy solutions also smoothly develop a hairy horizon (thermal horizon with nonzero fermion density $n(z_h)$). What changes is the order of the phase transition: at $T = 0$ it is continuous, and at finite temperature it is discontinuous.

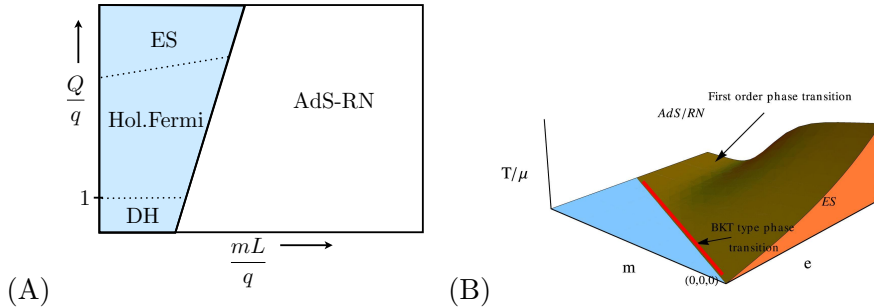


Figure 2: (A) Phase diagram as a function of the total-to-fermion-charge ratio Q/q (y -axis), and the fermion mass (in units of AdS radius L) over charge ratio mL/q (x -axis). For large masses, the effective potential is positive and the ground state is the bald RN black hole, with quantum critical dual field theory. For smaller masses, hair develops, which corresponds to a Fermi liquid in dual field theory. For $Q \sim q$ (few wavefunctions), the single-wavefunction Dirac hair approximation works; for $Q/q \rightarrow \infty$ we approach the semiclassical fluid (electron star) limit; between them there is a smooth crossover with unclear properties, both in AdS and in the holographic dual. Notice different notational conventions for the total charge from the main text (e vs. Q). Taken over from [15]. (B) Adding nonzero temperature as the third axis, we obtain also the thermal phase transitions between the black hole and the hairy solution, which are generically first order, smoothing out to an infinite order (BKT) transition at $T = 0$ – the red line in (B) is the bold black line between the RN and hairy (blue) region in (A).

3.1. Quantum hairy black holes

A controlled approximation is to solve the problem perturbatively, at one-loop order in fermionic fields. This is nothing but the textbook Hartree-Fock (HF) method, but in curved space. Dynamical spacetime makes a big difference: it introduces an additional strongly nonlinear component of the system, making the solution landscape larger and less predictable, and the UV and IR divergences can appear also in the Einstein equations and need explicit regulators. In fact, this is still an open problem – nobody has yet classified the solutions of the Einstein-Maxwell-Dirac system even in the Hartree-Fock approximation, and we do not know what surprises might lurk in this corner of the phase diagram. The HF electrodynamics contains two diagrams, a vacuum bubble that renormalizes the chemical potential

as $\hat{\mu}(z) \mapsto \hat{\mu}(z) + \delta\hat{\mu}(z)$ (the Hartree term):

$$\delta\hat{\mu}(z) \equiv q \sum_{\ell=1}^N \int \frac{kdk}{2\pi} \left(\psi_1^\dagger(E_\ell, k; z) \psi_1(E_\ell, k; z) + \psi_2^\dagger(E_\ell, k; z) \psi_2(E_\ell, k; z) \right) \quad (23)$$

and the exchange interaction (the Fock term). The explicit z -dependence of the Hartree correction is a gory reminder that the problem is solved in inhomogenous background. This is also the reason why already the Hartree correction is nontrivial: unlike the textbook situation where the shift $\delta\mu$ merely changes the numbers, here it is a radial function $\delta\mu(z)$ and its influence is also qualitative. So far, nobody even tried to do the whole HF calculation, and even just the Hartree term is not easy. We are plagued (1) by the UV divergences introduced by the modes close to $k = k_F$ which, as we have seen, peak most sharply near the boundary and can shatter the AdS space into pieces if not properly renormalized (2) by the IR divergences introduced by the modes with k close to zero, which extend far into large z values and can make the system unstable to forming a naked singularity.

Hard-wall Fermi liquid. The only case which is under good control is the hard-wall model of [12]: the UV divergences are resolved simply by not backreacting on the metric, i.e. solving just the Maxwell-Dirac system in fixed AdS metric (6) even without a black hole, and the IR divergences disappear by cutting off the space at some arbitrary z_0 , so that we simply eliminate the IR region. The approximations are rather drastic, but they allow a complete solution. In pure AdS space, the solutions $\psi_{1,2}$ can be found analytically in terms of Bessel functions, the states form discrete and gapped bands, and we only have to solve the Maxwell equation (18). The outcome is given in Fig. 3. Hard wall acts as an infinite potential barrier, so the wavefunctions should die on it, and the condition $\psi_{1,2}(z_0) = 0$ determines the dispersion relation. The wall should not be charged, so in (19) the second term on the right-hand side equals zero, meaning that $A'_t(z_0) = 0$. The picture is that of a Fermi liquid, nicely filling the Fermi sea at momenta $k \leq k_F$ and having long-living quasiparticles. This model is an important starting point for more complicated setups, and has the advantage of being intuitive, but by itself is too simplistic. Indeed, we want to talk about hairy black holes, and here we don't even have one, as it is hidden behind the hard wall!

An attempt to study a simple setup but with a black hole was made in [16]. In this approach, we are limited to a single energy level, $\ell = 1$. This is justified only when the hair is just starting to form, right at the transition point. There is again no backreaction on metric, but the (fixed) metric is now taken to be the RN black hole. This is actually a big jump in difficulty: the wavefunctions oscillate near the horizon at any nonzero energy (Fig. 4(A)), so they can satisfy the IR boundary condition at any energy and momentum (we can always pick the phase so that $\psi'(z_h) = 0$), and the spectrum is continuous as there is no wall to create a gap. This

is what forces us to consider the single-mode case: with the gapped hard-wall model we could add a *finite* number of modes, but now there is a *continuum* of them, N going to infinity even for arbitrarily small Q/q . The only way out is to *assume* there one mode only and solve the resulting Dirac-Maxwell system. This setup is convenient for understanding the transition itself, which turns out to be discontinuous (first-order) at finite temperature (Fig. 4(B)), and likely infinite-order (Berezinskii-Kosterlitz-Thouless, BKT) at zero temperature, as we shall soon see.

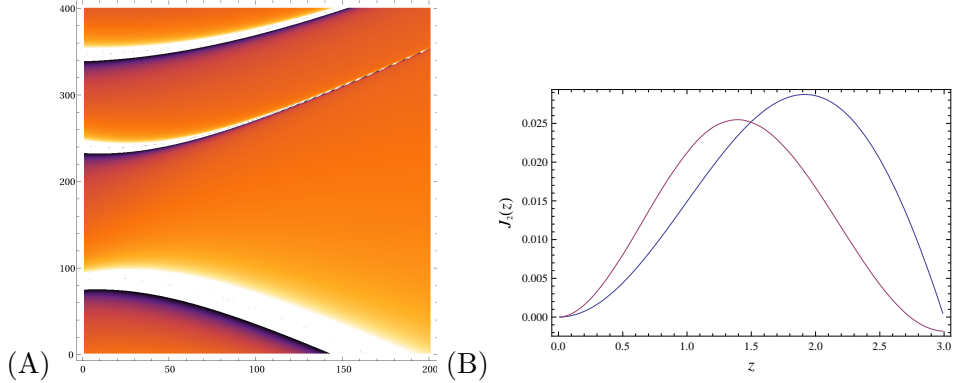


Figure 3: (A) Dispersion relation $E_\ell(k)$ for the first two electron bands $\ell = 1, 2$ in hard-wall AdS space, for $\mu = 1, m = 1, q = 2$. The first band from bottom is the hole band, not an an electron band – its contribution can be absorbed in the redefinition of the parameters and it does not contribute to hair. The colormap shows the resolvent of the Dirac operator, $(D_z \Gamma^z + \vec{D} \cdot \vec{\Gamma} - m - E)^{-1}$, thus the bright white regions show the places where the resolvent diverges and a discrete bound state is formed. The horizontal axis is the momentum and the vertical axis the energy, both in computational units. (B) Wavefunctions $\psi_{1,2}$ (here for $\ell = 1$ and $k = 1$) are smooth everywhere - what happen exactly at the horizon we do not know in this model, as the space is cut off at $z = 3$.

Quantum electron star. The single-mode approach has taught us a lesson: already at the level of the gauge field only, the changes from the finite fermion density are drastic, and the resulting stress tensor is large at the horizon, so a change of the black hole metric is certainly expected. However, when we try to solve the Einstein equations, things become almost intractable. Both UV and IR divergences appear: the former because the currents diverge in continuous space, and the latter because the discrete bands fuse into a continuum in IR. The latter issue is most easily regularized by a hard wall, but a hard wall does not make much sense if we want to backreact on geometry. The regularization of the UV divergences is systematically discussed in [17, 18] and the bottom line is that there is a logarithmic short-distance divergence which can be regularized by point

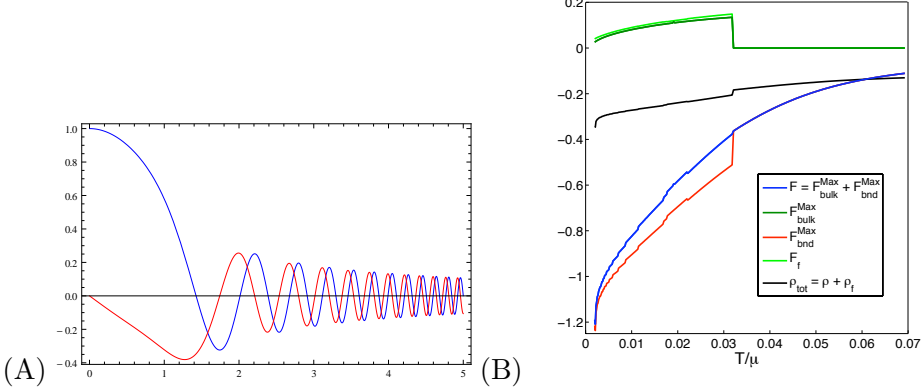


Figure 4: (A) Wavefunctions $\psi_{1,2}$ in RN background, for $\ell = 1$ and $k = 1$, always oscillate and approach an essential singularity at the horizon, which indicates an instability: the metric changes and the degenerate horizon disappears. (B) The bulk action (or free energy F , from AdS/CFT correspondence) of the Maxwell (electric field), in blue, consists of the bulk and boundary contribution (dark green and red), the former practically identical to the contribution from fermions. All these are computed from the action (2-4). While the *total* free energy is continuous, it has a cusp, made manifest by the slight jump in density (black), a sign of first-order hair-forming transition.

splitting; in this procedure the cosmological constant becomes renormalized. This is not a drastic change: it will just change the numbers but not qualitative behavior. The IR problem is still unsolved. The approach of [18] is to put the system in global AdS space⁴ whose radial slices are spheres, not planes, so the AdS radius provides a regulator. A perhaps more physical approach, motivated by consistent truncations from string theory, is to introduce a non-minimally coupled scalar, i.e., a dilaton that introduces a soft wall and suppresses the IR degrees of freedom in a continuous way, without an abrupt cutoff at some z_0 , so the total bulk action is now

$$\begin{aligned}
 S_{\text{bulk}} = & \int d^4x \sqrt{-g} \left[R - V(\Phi) - \frac{1}{2} (\partial\Phi)^2 - \frac{Z(\Phi)}{4} \hat{F}^2 \right] - \\
 & - \int d^4x \sqrt{-g} \bar{\Psi} \left(\frac{1}{2} D_a \Gamma^a e^\Phi + \frac{1}{2} e^\Phi D_a \Gamma^a + m \right) \Psi, \quad (24)
 \end{aligned}$$

where the dilaton potential reproduces the AdS cosmological constant near the boundary, i.e., $\Phi(z \rightarrow 0) = 0$ and $V(\Phi \rightarrow 0) = 6$, $Z(\Phi \rightarrow 0) = 1$. It is not clear if one can ever remove the IR regulator. That is precisely the

⁴Dual field theory then lives on a sphere instead of a plane.

reason that we regard the dilaton regulator as more physical, since string theory constructions as a rule contain non-minimally coupled scalars, and the action (24) can be obtained by consistent truncation; whereas global AdS is essentially an ad hoc solution, though a very interesting one, with possible applications in AdS/condensed matter duality, where systems that live on surfaces (such as a sphere) appear naturally.

While this is still very much a work in progress,⁵ preliminary results suggest that the RN-to-hairy-black-hole transition at zero temperature is an infinite-order (BKT) transition, where all derivatives of S remain smooth (Fig. 5). This is the point where the potential just starts deviating very slightly from the flat IR behavior in Fig. 1(A). At the end of this section we will try to understand this (still conjectural) numerical finding analytically.

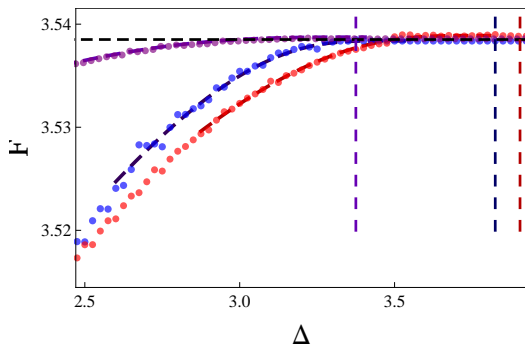


Figure 5: The bulk action (here denoted as free energy F , from AdS/CFT correspondence) as a function of the fermion mass (here denoted as $\Delta = 3/2 + m$) is very well fit by the BKT function $\exp(-c/\sqrt{\Delta_c - \Delta})$. The parameter c is determined by the chemical potential (we plot for three values $mu = 1.0, 1.5, 2.0$ in violet, blue, green). To the right of the transition point the action is independent of m as there is no hair, fermion density is zero, and so nothing depends on the fermion parameters. To the left of the transition point, the fermions form hair of nonzero density. Nobody knows yet how the near-horizon metric changes.

3.2. WKB star and electron star

WKB approach. We have followed the logical chain of reasoning from the point where the hair starts growing, having $Q/q \sim 1$ and deforming the black hole just a slight bit, towards larger and larger hair, eventually reaching the regime $Q/q \gg 1$. But this last regime is the easiest to approach, as the fermions become as close to classical as they can possibly be. A good starting point is the controlled expansion in \hbar , where we solve the

⁵With N. Chagnet, V. Djukić and K. Schalm.

Dirac equation in the eikonal approximation or, in other words, the WKB approach [15]. We express the wavefunction as

$$\psi_{1,2} = e^{i\theta_{\pm}}/\sqrt{p}, \quad p \equiv \sqrt{\hat{E}^2 - \hat{m}^2 - \hat{k}^2}, \quad (25)$$

where p has the role of the canonical momentum. The wavefunction is nonzero between the turning points z_{\pm} , determined by the equation $p(z_{\pm}) = 0$. The explicit form of the phase θ_{\pm} as well as higher-order corrections to the phase can be found in [15], but the reader should in fact have no difficulty in deriving them, following the usual WKB procedure (though for the Dirac equation instead of the Schrödinger equation). Now the density and pressure are found by inserting the solution (25) into (22). The procedure can be iterated to obtain self-consistent solutions, but now we solve the whole system including the Einstein equations. It is instructive to plot the total on-shell action (2) as a function of temperature (remember that finite temperature is imposed through the corresponding boundary condition for the metric function F).⁶ Fig. 6 plots the dependence $\mathcal{F}(T)$ in the vicinity of the transition value T_c : the derivative $\partial\mathcal{F}/\partial T$ undergoes a jump which is nothing but the entropy $\mathcal{S} \equiv \partial\mathcal{F}/\partial T$. We thus find a *first-order phase transition* at the point when Fermi hair starts forming. Of course, don't forget that the WKB approach is in fact *not* to be trusted very near the transition point: at the transition N changes from 0 to 1, which is far from the regime $N \ll 1$. But the qualitative insight that at finite temperature the system undergoes a non-symmetry-breaking transition is likely robust and we expect to prove it also within the more rigorous fully quantum-mechanical approach of the previous subsection. It is a hairy version of the celebrated Hawking-Page transition [19], and confirms the intuition that the high-temperature phase is always a black hole; but now, the low-temperature phase is not simply a gas, but a dense fluid in AdS.

Plotting the density and pressure in Fig. 7(A), one finds that for high values of N they tend to a constant value in deep interior. This motivates the fluid ansatz taken in the electron star limit, now to be considered.

Electron star. Electron star is a charged, AdS version of the neutron stars, described as perfect fluid by the Oppenheimer-Volkov equations. The idea is to *assume* that the fermionic matter is a perfect fluid, and then express the energy density ρ , pressure p and charge density n in terms of *integrals* over energy and momenta (i.e., assume that the bound states are infinitely close, and the gaps between them vanish). The fluid approximation thus becomes exact in the limit of $N \rightarrow \infty$, as we expect from a semiclassical approximation. Anticipating the current and stress tensor of the form

$$T_{\mu\nu} = (\rho + p)u_{\mu}u_{\nu} + pg_{\mu\nu}, \quad N_{\mu} = nu_{\mu}, \quad (26)$$

⁶In AdS/CFT, the bulk on-shell action \mathcal{S} precisely equals the free energy \mathcal{F} of the CFT side. But even without considering the details of the CFT, we can still make use of this interpretation to detect a phase transition in the system.

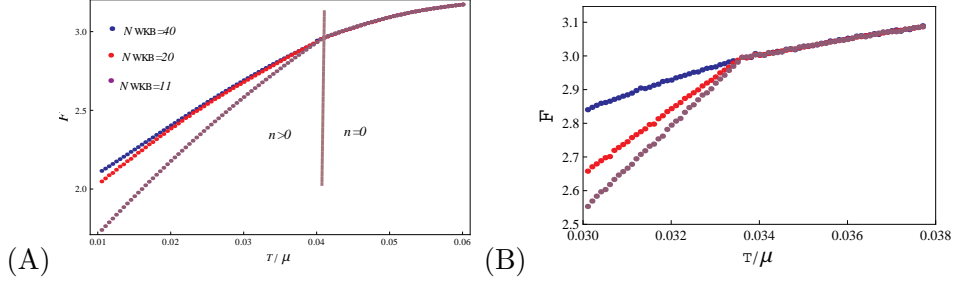


Figure 6: (A) The on-shell action or free energy as a function of temperature, in the presence of fermions. For low temperatures, the fermion density is finite and the derivative $\partial\mathcal{F}/\partial T$ jumps at $T = T_c$, a sign of first-order transition with the development of the hair. This is in line with the Dirac hair result in the previous figure, and indeed for the lowest number of levels N_{WKB} the transition is the sharpest. In (B) we zoom in into the transition region.

we can write the density starting from (22) and making use the optical theorem to relate it to the imaginary part of the Feynmann propagator G_F . This spells out as

$$\begin{aligned}
 \rho &= \int_0^{\hat{E}^2 - k^2} \frac{dE}{2\pi} \int_0^{k_F} \frac{d^3k}{(2\pi)^3} \hat{E} \Im \text{Tr} \Gamma^0 G_F(E, k) \\
 &= \int_0^{\hat{E}^2 - k^2} dE \int \frac{k^2 dk}{4\pi^3} \frac{1}{2} \left(1 - \tanh \left(\frac{\beta}{2} \hat{E} \right) \right) \text{Tr} (i\Gamma^0)^2 \delta \left(\hat{E} - \sqrt{k^2 + m^2} \right) \\
 &= \frac{1}{\pi^2} \int_m^{\hat{\mu}} dE E^2 \sqrt{E^2 - m^2}. \tag{27}
 \end{aligned}$$

We similarly find the number density n , whereas the pressure need not be computed explicitly: since we work with an isotropic free Fermi fluid, its equation of state has to be $p = \rho - qn\hat{\mu}$. It is here that the approximate nature of the electron star with respect to the WKB star becomes obvious (Fig. 7): in WKB star there is an extra term in the pressure, coming from the nodes of the WKB wavefunction. One can check that the integral in (27) indeed approaches a constant as we go into deep interior. On the other hand, at some z_* when $\hat{\mu}(z_*) = m$ the density falls to zero: the star is a classical object and has a sharp border. So for $0 < z < z_*$ we continue the metric to the RN metric (the metric outside a charged isotropic object).

Since we can express n, ρ, p explicitly, we get a nice system of local ordinary differential equations in F, G, A_t , with all quantum expectation values pulled under the rug. This completes the circle, and brings another universal message: *due to Pauli principle, fermionic operators are never local, except in two extreme cases: when only one state is occupied (so the format of the Slater determinant is 1×1 , i.e., it contains a single state), or*

when infinitely many states are occupied, so the Slater determinant turns into a classical, continuous probability density. In Fig. 7 we can see how the WKB solution captures the quantum "tails" near the turning points, which the electron star does not have. It is also instructive to compare this solution to the Oppenheimer-Volkov equations in flat space: in the latter case, $\hat{m} \sim 1/\sqrt{F}$ is always larger than $\hat{\mu} \sim 1/F\sqrt{G}$, unlike in AdS where $\hat{m} \sim 1/z\sqrt{F}$ and for $z > z_*$ it becomes smaller than the local chemical potential, so the integral in (27) has a nonzero range. This is because AdS acts like a potential box that can hold the charged fermions together against electrostatic repulsion. In flat space that does not happen, and we have only neutron stars, not electron stars.

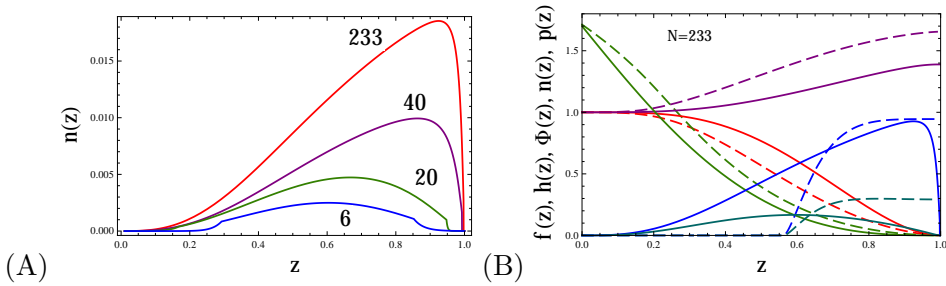


Figure 7: (A) Density of the finite temperature WKB star at various fillings N_{WKB} ; besides the classically allowed region, there are also exponentially decaying tails in the classically forbidden region, where $V_{\text{eff}} > 0$. (B) In the electron star (fluid limit), there are no such tails and the star has a sharp border. Taken over from [15]. (B) Comparison of the WKB solution (full lines) and the electron star solution (dashed lines) at the same chemical potential, fermion charge and mass. We plot the metric functions f, h (F, G in the main text) in red and violet, the gauge field Φ (A_t in the main text) in green, and density and pressure n, p in blue and dark green. The metric solutions do not differ much, despite the long quantum WKB tails, absent in the electron star.

3.3. Lifshitz metric, BKT transition and the missing pieces

In the framework of the electron star model, the Einstein-Maxwell equations can be solved analytically, thanks to the fact that, in deep IR, $n, \rho, p = \text{const.}$ and we can employ a scaling ansatz for the metric. The idea is to match the IR expansion around the scaling solution to the UV expansion around pure AdS. With ansatz of the form $g_{tt} \propto -1/z^\alpha$, $g_{ii} \propto 1/z^\beta$ and $g_{zz} = 1/z^2$ (one metric component we can fix at will as it amounts to picking the gauge for the metric), equations of motion give the IR solution

$$ds^2 = -\frac{1}{z^{2\zeta}} dt^2 + \frac{1}{z^2} d\vec{x}^2 + \frac{1}{z^\zeta} dz^2$$

$$A_t = \frac{1}{z\zeta}, \quad \mathcal{L} = \frac{2}{z^2\zeta} + 6 - \partial_z A_t^2 - nA_t^2 - p_\perp, \quad (28)$$

where in the expression for the total Lagrangian density in the second line, we have inserted in the action (2) the solutions for the metric and the gauge field, as well as the constant (z -independent) solution obtained for ρ in (27) and similarly for n, p . Three important conclusions can be drawn: (1) the IR metric is scale-invariant, with anisotropic scaling of time and space, so that the scaling transformation has the form $t \mapsto \lambda t$, $\vec{x} \mapsto \vec{x}\lambda^{1/\zeta}$ (2) the on-shell Lagrangian density effectively describes a *massive vector field*, with mass squared equal to fermion density n (3) the fermionic contribution to the action equals the pressure. The second point agrees with the known result that Lifshitz black holes are generated by Proca fields [21], and what happens is the Abelian-Higgs mechanism: fermion density acquires a finite expectation value which in turn breaks the $U(1)$ symmetry, giving the photon a mass. The third point is expected within a fluid model, since the action of an ideal Lorentz-invariant (semi)classical fluid equals its pressure [7]. In the fluid limit we can also understand the first-order transition at finite temperature, because it is just a van der Waals-type liquid-gas transition.

We have seen that the thermal transition from RN to a Lifshitz black hole is of first order, and that the $T = 0$ transition is apparently a BKT (infinite order) transition. The latter is not quite clear yet because, as we have emphasized, nobody has yet managed to peek into the deep IR, it remains hidden behind the hard wall. But if we tentatively accept the numerical evidence for the infinite-order transition, can we understand it theoretically? The key lies in understanding how the AdS_2 throat disappears. The conformality-breaking mechanism of [22, 23] gives an idea, though the details are still missing. The crucial moment is that the near-horizon geometry is AdS_2 . Right at the horizon ($s \rightarrow -\infty$) the potential is approximately constant. In the UV of the AdS_2 throat, which is around some finite value s_0 , the potential behaves as $-c/(s - s_0)^2$. This inverse-square potential is known to describe conformal quantum mechanics when $c > -1/4$. For $c = -1/4$ the conformal invariance breaks. discrete states appear and the effective potential is not consistent unless regularized as

$$V_{\text{eff}} = \frac{c}{(s - s_0)^2} - v\delta(s - s_0), \quad (29)$$

and the solution of the effective Schrödinger equation is

$$\psi(r) = c_+(s - s_0)^{\alpha_+} + c_-(s - s_0)^{\alpha_-}, \quad \alpha_\pm = \frac{1}{2} \pm \sqrt{c + \frac{1}{4}}, \quad (30)$$

and the ratio c_+/c_- is given in terms of Bessel functions $J_{1/2}$ and $J_{-1/2}$:

$$\frac{c_+}{c_-} = -\epsilon^{\alpha_- - \alpha_+} \frac{\gamma + \alpha_-}{\gamma + \alpha_+}, \quad \gamma = \sqrt{v} \frac{J_{1/2}(\sqrt{v})}{J_{-1/2}(\sqrt{v})} \quad (31)$$

The solution (30) diverges at $s = s_0$ unless we introduce a cutoff at some distance ϵ from s_0 . Imposing the renormalization condition that c_+/c_- remains independent of ϵ , we get the β -function of the renormalization group as (ℓ being the RG scale):

$$\beta \equiv \frac{d\gamma}{d\ell} = (c + 1/4) - (\gamma + 1/2)^2. \quad (32)$$

And we're done: the fixed points of the above flow equation are easily found to be $-\alpha_{\mp}$. For $\gamma = -\alpha_{\mp}$ we get the solution for ψ from (30) with $c_{\pm} = 0$ respectively. The free energy scaling is obtained as $S_{\text{on-shell}} = \mathcal{F} \propto \int d\ell/\beta$, which gives just the form found in Fig. 5. However, the presence of both a hard-wall cutoff in z and the soft-wall dilaton, completely unaccounted for in the above analysis, clearly suggest more work is needed for everything to click together.

4. Wormholes with fermion hair

The lengthy review we have given so far is meant to be self-contained and helpful for those interested in understanding and contributing to the problem of black hole instabilities with fermionic matter. As we have seen, it contains some puzzling questions and is of more than technical interest (after all, the whole field has been active mostly for the last fifteen years or so). But we also want to point out that with the methodological powerhouse of the HF, WKB and fluid methods, one can tackle new problems. A recent issue where fermions at finite density seem very relevant is the search for traversable wormholes.

The motivation for this story lies mainly in the celebrated black hole information paradox: as far as we know, the Hawking radiation is thermalized, meaning that the information content of the matter falling into the black hole is lost. A possible way out or, at least, a way to better understand the issue, is to consider the maximally extended Carter-Penrose diagram of a black hole, which contains two horizons and two spacetimes. If transport between the two were possible, one could imagine that the information is not lost because the matter falling into one horizon is entangled with the matter on the opposite side. This is the idea of the ER=EPR conjecture [24]. In order to build a traversable wormhole, one needs negative that the stress-energy tensor averaged over a geodesic be negative, thus violating the so-called averaged negative energy condition (ANEC) [25, 26]. This will never happen with conventional classical matter. One needs either exotic fields or quantum corrections. Recently however, a few traversable wormholes have been realized with only standard-model matter. The most "conservative" is the setup of [27] which creates negative energy by considering a particle-hole symmetric spectrum of massless fermions in a magnetic monopole field: because of the negative Landau levels, the net energy is negative. The starting point is thus a pair of magnetically charged RN black holes with magnetic charges H and $-H$, with the hope that the

negative energy Landau levels will push the averaged stress tensor to large enough absolute values to open up a wormhole. In this way, [27] constructs a quasi-stationary (long-living) wormhole in *asymptotically flat* space. In AdS, negative energy density can easily be constructed by coupling the two boundaries nonlocally: in this way temporary wormholes, opening up for the finite duration of the perturbation, can be constructed [28], and even eternal wormholes are possible but at the cost of much more exotic boundary CFTs and their couplings [29, 30, 31]. Here we are interested in making a wormhole in a more "down-to-earth" manner, by growing negative-energy fermion levels as in [27]. The task is to make such wormholes more stable, and to see if they survive at higher fermion density rather than just a single wavefunction as in [27]. Here the previously developed methods can help us.

Magnetic electron star. The crucial consequence of the magnetic field is the Landau quantization. The motion along the x -coordinate is quantized into discrete levels, whereas the motion along y is not quantized and introduces degeneracy. The quantization along x -axis makes our life somewhat easier – even without any IR cutoff the ground state wavefunction now has a discrete quantum number, the Landau level m_j . The magnetic field breaks the spherical symmetry of the wavefunctions down to cylindrical, so it is convenient to introduce the polar angles θ, ϕ :

$$ds^2 = -A(z)dt^2 + B(z)dz^2 + C(z) \left(d\theta^2 + \sin^2 \theta d\phi^2 \right) \quad (33)$$

and to pick a different gamma matrix basis: $\Gamma^0 = \iota\sigma_1 \otimes \hat{1}$, $\Gamma^1 = \sigma_2 \otimes \hat{1}$, $\Gamma^2 = \sigma_3 \otimes \sigma_1$, $\Gamma^3 = \sigma_3 \otimes \sigma_2$. Separating the variables and representing the wavefunction as

$$\Psi = \sum_{m_j=-j}^j (\psi_+(m_j; z), \psi_-(m_j; z)) \otimes (\eta_1(m_j; \theta), \eta_3(m_j; \theta)) e^{im_j\phi}, \quad (34)$$

where j is the total number of Landau levels $j = (H - 1)/2$, we get the fully spin-polarized solution ($\eta_2 = 0$) for zero fermion mass:

$$\psi_{\pm}(m_j; z) = \exp \left(\pm \iota E(m_j) \int_0^z dz' \sqrt{\frac{B(z')}{A(z')}} \right), \quad (35)$$

$$\eta_1(m_j; \theta) = \frac{e^{\iota H \sin \theta/2}}{\sqrt{\sin \theta}} \left(\tan \frac{\theta}{2} \right)^{m_j}.$$

For nonzero mass, we can perform a Foldy-Wouthuysen transform starting from the above solution. Unlike the massless case considered in [27], the resulting stress-energy tensor will not be traceless, but that is precisely what will give us extra stability. The reason this is consistent is the Landau quantization: the levels for different m_j are gapped from each other and

each Landau level can be treated as a single-particle solution which does not mix with other Landau levels. This results in the stress tensor

$$\langle T_{zz} \rangle = \frac{E_n}{(1+z^2)^2} (\sin 2\alpha - \cos 2\alpha), \quad \tan \alpha = -m/E(m_j). \quad (36)$$

Fig. 8(A) shows the radial pressure T_{rr} as a function of energy, the outcome being that *positive* stress energy tensor is produced for $0 > E > -m$. In order to avoid this positive contribution, the Landau level spacing has to be large enough, i.e., larger than the mass gap (at zero mass this condition is trivially satisfied, as it simply means that any finite $E(m_j = 1)$ will do; this is the case studied in [27]). The simplest gapping mechanism we can think of is the chemical potential, i.e. an electrostatic field in addition to the magnetostatic one. The black hole thus has to become dyonic, with magnetic charge H and electric charge e . Assuming we have ensured the negativity of (36), we can write it in the form $T_{zz} = -\tau/(1+z^2)^2$, with τ a positive constant. Its magnitude roughly determines the size of the wormhole opening.

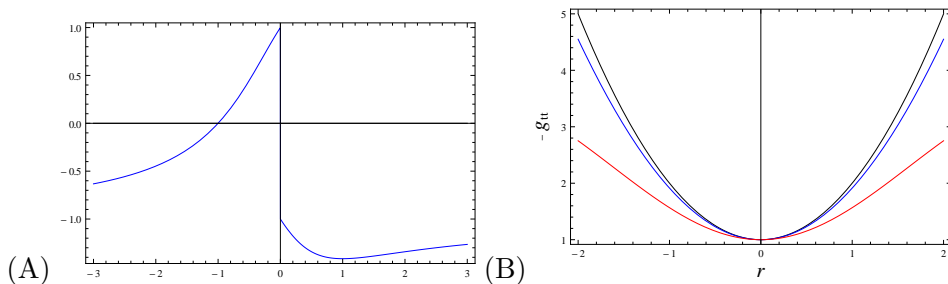


Figure 8: (A) Radial component of the stress-energy tensor $\langle T_{rr} \rangle$ as a function of the (discrete) fermion energy E . Positive contribution only comes when $> E(m_j) > -m$. In order to avoid this range of energies we need a nonzero chemical potential (i.e., electric field, resulting in a dyonic black hole) to stabilize the wormhole with massive fermionic hair. (B) The solution for the metric component g_{tt} in the intermediate region, as a function of the radial coordinate r , for $\tau = 0, 0.05, 0.10$ (black, blue, red). Wormhole solutions (blue, red) are quantitatively very close to the unperturbed black hole (black) but qualitatively different as there is no zero anymore.

Wormhole solution and matching. Having computed the stress-energy tensor (36), we can solve the Einstein equations. The strategy is again matching the expansions, but now we have three regions: the far region which is asymptotically AdS or even flat (we have mentioned that in the presence of magnetic field discrete bound states can form even in absence of AdS boundary), the intermediate region is a slightly perturbed near-horizon AdS₂ region of our magnetic RN geometry, and the inner region, the wormhole throat that opens up, turns out to be a *global* AdS₂ at leading order,

so it has a spherical boundary continuing onto the intermediate regions. The inner, near-global-AdS₂ metric in the form (33) at leading order reads

$$\begin{aligned} A(z) &= R_0^2 \left[1 + z^2 - 8\pi\tau \left(z^2 + (3z + z^3) \arctan z - \log(1 + z^2) \right) \right] \\ B(z) &= R_0^4/A(z), \quad C(z) = R_0^2 [1 + 8\pi\tau (1 + z \arctan z)]. \end{aligned} \quad (37)$$

This solution is to be matched to the intermediate-region solution. Now large z corresponds to the wormhole mouth, i.e., the matching is to be done at large z , where small z is the "center" of the wormhole throat. The solution to match onto is the RN black hole metric:

$$\begin{aligned} ds^2 &= -l^2 f(r) dt^2 + \frac{dr^2}{f(r)} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \\ l &= \frac{R_0}{2\pi^2\tau}, \quad R = \frac{r - \sqrt{\pi}\sqrt{e^2 + H^2}}{2\pi^2\tau}. \end{aligned} \quad (38)$$

The solution thus exists for any choice of e and H . But for large e (in other words, for a large chemical potential), the density of the hair will increase significantly and we should repeat the WKB star or electron star approach. In AdS this is a simple matter, proving the stability of the configuration even at high densities. The interesting question is, can it work also in asymptotically flat space? In absence of magnetic field, the answer is certainly no – without an AdS boundary, there is nothing to equilibrate the electrostatic repulsion of electrons. But in the presence of magnetic field, one might obtain a stable charged hairy wormhole if the change in the near-horizon geometry is sufficient to effectively decrease the electrostatic energy density. This is the logical immediate task for future work.

We finish this short review of our work in progress on hairy wormholes with a somewhat more ambitious task. The dyonic wormhole model considered here is obviously quite simplistic and artificial. A much more realistic model is to start from a pair of Kerr black holes and see if these can open up a wormhole in a manner analogous to the scenario we have considered. In this case the magnetic field would be generated self-consistently by the (rotating) fermionic hair, removing the need for the magnetic monopole charge. Such an object would come much closer to realistic astrophysical matter.

5. Instead of a conclusion

We have given a crack and practical review of the insights and technologies needed to describe and understand hairy black holes in anti-de Sitter space. The phase diagram in the presence of nonzero fermion density is quite rich, and it involves two deep and universal phenomena. First, the finite-temperature hairy black holes develop through a discontinuous phase

transition akin to the Hawking-Page transition (indeed, it is precisely the Hawking-Page transition but at finite density). The standard lore that at high enough temperatures black holes will always form is confirmed. Notice this is true at any fermion mass and charge, and thus at any occupation number, from a single wavefunction to the fluid limit, so the finding is definitely robust. Second, at zero temperature the transition is driven by the fermionic charge and/or chemical potential, i.e., electric charge of the black hole. In this case the black hole vanishes infinitely slowly, in a BKT transition that can be understood as the breaking of the one-dimensional conformal symmetry of the wavefunctions in the effective inverse-square potential well. This is solely the consequence of the near-horizon physics, independent of the AdS boundary. Similar conformality-breaking infinite-order transitions are known in various backgrounds in string theory. Maybe one could relate the case described here to some consistent top-down model.

As mentioned in the Introduction, we have deliberately left out extensions and applications of the formalism described, for reasons of space and also generality of discussion. The field of applications closest to our experience is the AdS/CFT correspondence. Electrically charged black holes are dual to field theories at finite $U(1)$ density. The transition from a bald black hole to a hairy black hole is thus a transition between two phases at equal chemical potential. How do they differ then? We know that a black hole is dual to the Coulomb (deconfined) phase of some non-Abelian finite-temperature gauge theory [1, 4]; in the simplest setup coming from type IIB string theory, it is the $\mathcal{N} = 4$ supersymmetric $SU(N)$ theory. Coulomb phase means that the $U(1)$ charge is carried by $SU(N)$ -gauge-charged operators, in our case fermions ("mesinos") and thus not visible to low-energy probes, since at low energies all operators are likely $SU(N)$ -gauge-neutral. The hairy phase describes a dual field theory where the charge is carried by gauge-neutral operators ("baryons") and thus visible to probes such as a photon. This viewpoint was tried and confirmed in [11, 12, 15, 23]. It has realizations in condensed matter systems such as strange metals and heavy fermion materials. In this case, the gauge fields are emergent and arise from the spin-charge separation, and the transition between a black hole and a hairy geometry is a transition between a non-Fermi liquid, where most of the charge is carried by complicated excitations that are not directly seen in the spectrum, and a Fermi liquid where the fundamental degrees of freedom are just renormalized electrons. In QCD, this picture describes the phase diagram at intermediate energy scales and finite densities, where a black hole describes quark-gluon plasma, and a hairy solution describes either the color condensate or conventional barionic matter depending on the details of the model. One can learn a lot on AdS/condensed matter and AdS/QCD from [5, 6].

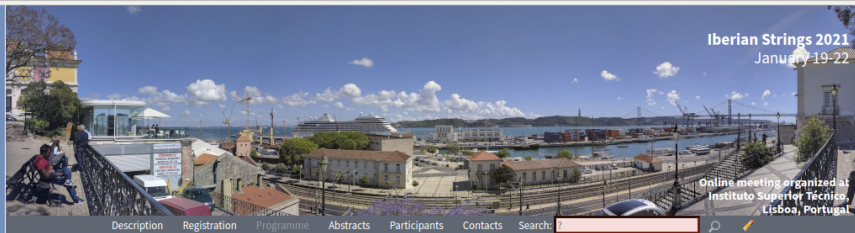
Finally, the search for wormhole solutions and how fermionic hair might stabilize them is likely to become very important in the future, in connection to the quantum information theory and the firewall, ER=EPR and other approaches to the black hole information problem. One can use much

of the formalism developed for hairy black holes, but the interpretation is still challenging. It is also unclear how realistic the wormhole proposal is if we work with only conventional, standard model matter, i.e. is it just an important proof of concept or a realistic model?

References

- [1] J. Maldacena, *The large N limit of superconformal field theories and super-gravity*, Adv. Math. Theor. Phys. **2**, 231 (1998). [arXiv:hep-th/971120]
- [2] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, *Gauge theory correlators from non-critical string theory*, Phys. Lett. B **428**, 105 (1998). [arXiv:hep-th/9802109]
- [3] E. Witten, *Anti de Sitter space and holography*, Adv. Math. Theor. Phys. **2**, 253 (1998). [arXiv:hep-th/9802150]
- [4] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri and Y. Oz, *Large N field theories, string theory and gravity*, Phys. Rep. **323**, 183 (2000). [arXiv:hep-th/9905111]
- [5] J. Zaanen, Y.-W. Sun, Y. Liu and K. Schalm, *Holographic duality in condensed matter physics*, Cambridge University Press, 2015.
- [6] M. Ammon and J. Erdmenger, *Gauge/gravity duality: foundations and applications*, Cambridge University Press, 2015.
- [7] S. W. Hawking and G. F. R. Ellis, *The large-scale structure of space-time*, Cambridge University Press, 2010.
- [8] N. D. Birrel and P. C. W. Davies, *Quantum fields in curved space*, Cambridge University Press, 1982.
- [9] H. Liu, J. McGreevy, D. Vegh, "Non-Fermi liquids from holography", Phys. Rev. D **83**, 065029 (2011). [arXiv:0903.2477[hep-th]]
- [10] M. Čubrović, J. Zaanen and K. Schalm, "String Theory, Quantum Phase Transitions and the Emergent Fermi-Liquid", Science **325**, 439 (2009). [arXiv:0904.1993[hep-th]]
- [11] T. Faulkner, H. Liu, J. McGreevy and D. Vegh, *Emergent quantum criticality, Fermi surfaces, and AdS₂*, Phys. Rev. D **83**, 125002 (2011). [arXiv:0907.2694 [hep-th]].
- [12] S. Sachdev, *A model of a Fermi liquid using gauge-gravity duality*, Phys. Rev. D **84**, 066009 (2011). [arXiv:1107.5321 [hep-th]]
- [13] W. Mück and K. S. Viswanathan, *Conformal field theory correlators from classical field theory on anti-de Sitter space: Vector and spinor fields*, Phys. Rev. **D58**, 106006 (1998). [arXiv:hep-th/9805145]
- [14] M. Čubrović, Y. Liu, K. Schalm, Y.-W. Sun and J. Zaanen, *Spectral probes of the holographic Fermi groundstate: Dialing between the electron star and AdS Dirac hair*, Phys. Rev. **D84**, 086002 (2011). [arXiv:1106.1798 [hep-th]]
- [15] M. Medvedyeva, E. Gubankova, M. Cubrovic, K. Schalm, J. Zaanen, *Quantum corrected phase diagram of holographic fermions*, JHEP **2013**, 025 (2013). [arXiv:1302.5149[hep-th]].
- [16] M. Čubrović, J. Zaanen, K. Schalm, *Constructing the AdS dual of a Fermi liquid: AdS Black holes with Dirac hair*, JHEP **2011**, 17 (2011). [arXiv:1012.5681[hep-th]]
- [17] A. Allais, J. McGreevy and X. Josephine Suh, *Quantum electron star*, Phys. Rev. Lett **108**, 231602 (2012). [arXiv:1202.5308[hep-th]]
- [18] A. Allais and J. McGreevy, *How to construct a gravitating quantum electron star*, Phys. Rev. **D88**, 066006 (2013). [arXiv:1306.6075[hep-th]]

- [19] S. Hawking and D. Page, *Thermodynamics of black holes in anti de Sitter space*, Comm. Math. Phys. **87**, 577 (1977).
- [20] S. A. Hartnoll, A. Tavanfar, *Electron stars for holographic metallic criticality*, Phys. Rev. D **83**, 046003 (2011). [arXiv:1008.2828[hep-th]]
- [21] K. Balasubramanian and J. McGreevy, *An analytic Lifshitz black hole*, Phys. Rev. D **80**, 104039 (2009). [arXiv:0909.0263[hep-th]]
- [22] D. B. Kaplan, J.-W. Lee, D. T. Son, M. A. Stephanov, *Conformality Lost*, Phys. Rev. D **80**, 125005 (2009). [arXiv:0905.4752[hep-th]].
- [23] M. Čubrović, *Confinement/deconfinement transition from symmetry breaking in gauge/gravity duality*, JHEP **2016**, 102 (2016). [arXiv:1605.07849[hep-th]]
- [24] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, Fortschr. Phys. **61** 781 (2013). [arXiv:1306.0533[hep-th]]
- [25] M. S. Morris and K. Thorne, *Wormholes in spacetime and their use for interstellar travel: A tool for teaching general relativity* Am. J. Phys **56**, 395 (1988).
- [26] D. Hochberg and M. Visser, *Null energy condition in dynamic wormholes*, Phys. Rev. Lett **81**, 746 (1998).
- [27] J. Maldacena, A. Milekhin and F. Popov, *Traversable wormholes in four dimensions*, (2018). [arXiv:1807.04726[hep-th]]
- [28] P. Gao, D. L. Jafferis and A. C. Wall, *Traversable wormholes via a double trace deformation*, JHEP **12**, (2017) 151. [arXiv:1608.05687[hep-th]]
- [29] Z. Fu, B. Grado-White and D. Marolf, *A perturbative perspective on self-supporting wormholes*, Class. Quant. Grav. **36**, (2019) 045006. [arXiv:1807.07917[hep-th]]
- [30] D. Marolf and S. McBride, *Simple perturbatively traversable wormholes from bulk fermions*, (2019). [arXiv:1908.03998[hep-th]]
- [31] J. Maldacena and X.-L. Qi, *Eternal traversable wormhole*, (2018). [arXiv:1804.00491[hep-th]]



Programme

Move the mouse over the schedule to see start and end times and complete clipped titles.

	Tue, 19 Jan 2021	Wed, 20 Jan 2021	Thu, 21 Jan 2021	Fri, 22 Jan 2021
9-11	Malcom Perry Soft Charges, Soft Hair and Black Hole Entropy	Johanna Erdmenger Information geometry and QFT	Jerome Gauntlett Geometric Extremization for AdS/CFT and Black Hole Entropy	Lárus Thórlaclaus Entanglement islands
	David Mateos Strong-coupling dynamics and entanglement in de Sitter space	Martin Sasietta A Momentum/Complexity Correspondence	Anayeli Ramirez 1/4 BPS AdS ₂ /CFT ₂	Roberto Emparan Quantum BTZ black hole
	Coffee break	Coffee break	Coffee break	Coffee break
11-13	Matteo Baggioli How small hydrodynamics can go	Mihailo Cubrovic Lyapunov spectra in traversable wormholes and their holographic duals	Salvatore Baldino Seiberg-Witten theory, string theory and WKB analysis	Marija Tomasevic Multi-mouth traversable wormholes
	Pablo Cano Novel higher-curvature variations of \mathbb{R}^2 inflation	Kevin Nguyen Slow scrambling in extremal BTZ and microstate geometries	Gabriel Larios Kaluza-Klein spectra and consistent truncations	Tomas Andrade Violations of weak cosmic censorship in black hole collisions
	Mikel Sanchez Garitaonandia Bubble Mergers from Holography	Johannes Lahnsteiner A Non-Relativistic Limit of NS-NS Gravity	Mattia Cesaro Kaluza-Klein fermion mass matrices from Exceptional Field Theory and $N = 1$ spectra	Marti Roselló Arithmetic of decay walls through continued fractions: a new exact dyon counting solution in $\mathcal{N} = 4$ CHL models
	Yago Bea New insights from Real-time Dynamics	Edvard Musaev Non-abelian U-duality of M-theory background	Mario Herrero Amplitudes Positivity and IR-UV connections in Graviton Exchange	Pablo Emmanuel León Torres Massive supermembrane in ten non-compact dimensions
13-15	Lunch	Lunch	Lunch	Lunch

Mihailo Cubrovic University of Belgrade

Lyapunov spectra in traversable wormholes and their holographic duals

2021/01/20, 11:00 – 11:30

We study the decay of out-of-time-ordered correlators (OTOC) in an AdS traversable wormhole and its gravity dual, two coupled Sachdev-Ye-Kitaev models ("left" and "right" subsystem). The gravity calculation of OTOC involves perturbative equations more involved than for a black hole, as the perturbation has complex kinematics and can bounce back and forth through the wormhole many times. The outcome is a phase diagram with three regions. One is black-hole like with uniform exponential growth and the Lyapunov exponent $\lambda = 2\pi T$ ("the chaos bound"). The intermediate phase has OTOCs with a spectrum of different exponents for different operator modes, all below the maximal chaos bound. The third phase has exponentially small Lyapunov exponents, behaving as $\exp(-1/T)$, in accordance with a recent field-theory calculation in the literature. The Lyapunov spectrum carries more information than just the maximum exponent: it can be related e.g. to teleportation fidelity from left to right subsystem.

Frank Eisenhauer Max Planck Institute for extraterrestrial Physics

The Discovery of the Massive Black Hole in the Center of the Galaxy. Outreach colloquium on the occasion of The Nobel Prize in Physics 2020

2021/01/20, 17:30 – 18:30

Outreach colloquium on the discovery of a supermassive compact object at the centre of our galaxy.

Black Holes are among the most mysterious objects in the Universe. They are so massive and compact that nothing - not even light - can escape their gravity. The 2020 Nobel Prize in Physics was awarded to Roger Penrose for showing that these exotic objects are a direct consequence of Einstein's general theory of relativity, and to Reinhard Genzel and Andrea Ghez for the discovery of such a monster in the center of our Galaxy. Our presentation will portrait the 40 year journey from the first indications to the overwhelming observational evidence for a extremely heavy and compact object in the Galactic Center, for which a supermassive black hole is the only known explanation. Using the world's largest telescopes and most advanced optics technology, astronomers can now follow the stars orbiting the central object, precisely measure its mass, and detect the stunning effects of general relativity. In our talk we will present both the spectacular observations and the technology behind.

Zachary Elgood Universidad Autónoma de Madrid

The first law of heterotic stringy black hole mechanics at zeroth order in α'

2021/01/22, 16:30 – 17:00

We re-derive the first law of black hole mechanics in the context of the Heterotic Superstring effective action compactified on a torus to leading order in α' , using Wald's formalism, covariant Lie derivatives and momentum maps. The Kalb-Ramond field strength of this theory has Abelian Chern-Simons terms which induce Nicolai-Townsend transformations of the Kalb-Ramond field. We show how to deal with all these gauge symmetries deriving the first law in terms of manifestly gauge-invariant quantities. In the presence of Chern-Simons terms, several definitions of the conserved charges exist, but the formalism picks up only one of them to play a role in the first law. The derivation of a first law is a necessary step towards the derivation of a manifestly gauge-invariant entropy formula which is still lacking in the literature. In its turn, this entropy formula is needed to compare unambiguously macroscopic and microscopic black hole entropies.

Activities Google Chrome Tue 13:52

www2.yukawa.kyoto-u.ac.jp/~qft.web/2021/

Leonard Susskind (Stanford)
Does the Holographic Principle Apply to de Sitter space, and if so, How?

Topics Covered

There have been annual workshops in Kyoto on string theory and quantum field theory for a long time, which are definitely the largest in this field in Japan. This year marks the 25th anniversary. [The contents of the past conferences are archived here.](#)

This year we invite a few speakers from abroad and all the talks will be in English. We plan to have a few parallel sessions for short talks if we receive a large amount of applicants.

Registration

The registration has been closed. The registration information can be changed [here](#).

There is no registration fee.

The workshop will be held purely online, due to the worsening COVID situation in Kyoto.

The deadline of the registration **for giving talks** is **Thursday, July 1st**.
The deadline of the registration **requesting financial support** is also **Thursday, July 1st**.
The deadline of the registration **for participation** is **Monday, August 9th**.

The workshop consists of oral sessions and poster sessions. At the registration, a short oral talk or a presentation at virtual poster sessions can be chosen. We plan to have a few parallel sessions for short talks if we receive a large amount of applicants. We will try to accommodate your preference, but we may need to limit the number of talks. The abstract submitted at the registration will be announced at the web site later.

Slides in oral sessions and files in poster sessions will be uploaded at this web site and submitted to [the online Japanese journal 素粒子論研究](#) (translated as "Research in high-energy physics"). We will not separately publish proceedings for this workshop. Please contact the organizers if you do not want to disclose your slides or poster files. All the talks will be broadcast online.

The organizers

Masashi Hamanaka (Nagoya University)
Koji Hashimoto (Kyoto University)
Yasuaki Hikida (Kyoto University)
Masazumi Honda (Kyoto University)
Kazuo Hosomichi (National Defense Academy)
Hiroshi Kunitomo (Kyoto University)
Kazunobu Maruyoshi (Seikei University, chair)
Takeshi Morita (Shizuoka University)
Keia Nii (Kyoto University)
Toshifumi Noumi (Kobe University)
Tadakatsu Sakai (Nagoya University/KMI)
Makoto Sakamoto (Kobe University)
Shigeki Sugimoto (Kyoto University)
Yuji Tachikawa (IPMU)
Tadashi Takayanagi (Kyoto University)
Satoshi Yamaguchi (Osaka University)
Kentaro Yoshida (Kyoto University)

If you have any questions not covered in this webpage, please send an email to maruyoshi@st.seikei.ac.jp.

Activities Google Chrome Tue 15:54

www2.yukawa.kyoto-u.ac.jp/~qft.web/2021/program.html

Shuichi Yokoyama	Imperial College London	Topologically protected phase factor, spin Chern-Simons theory and duality on lens space
Zhenghao Zhong	Imperial College London	3d Mirrors of U & SU Quivers
Parallel Slot B Aug 25 08:10 CEST - Aug 25 09:10 CEST		
Satoshi Kanno	University of Tsukuba	Various generalizations of the matrix regularization
Mihailo Cubrovic	Institute of Physics Belgrade	Chaos and replica wormholes in the IIB matrix model
<p>We study the dynamics and statistics of the Ishibashi-Kawai-Kitazawa-Tsuchiya (IKKT) IIB matrix model. We first consider the spectrum and detect signs of level repulsion and quantum chaos. Then we expand perturbatively the partition function of the quenched model and find that, with an appropriately chosen quenching, the expansion contains wormhole terms akin to those in the time-frozen SYK model. We conjecture that such terms are crucial in describing the chaotic aspects of the IIB model.</p>		
Invited Aug 25 09:40 CEST - Aug 25 10:50 CEST		
Sven Krippendorff	LMU München	Theoretical Particle Physics and Machine Learning