Nonlinear analogue of the May–Wigner instability transition

Yan V. Fyodorov and Boris A. Khoruzhenko

School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom

Edited by Srinivasa S. R. Varadhan, Courant Institute of Mathematical Sciences, New York, NY, and approved May 9, 2016 (received for review January 22, 2016)

We study a system of \( N \gg 1 \) degrees of freedom coupled via a smooth homogeneous Gaussian vector field with both gradient and divergence-free components. In the absence of coupling, the system is exponentially relaxing to an equilibrium with rate \( \mu \). We show that, while increasing the ratio of the coupling strength to the relaxation rate, the system experiences an abrupt transition from a topologically trivial phase portrait with a single equilibrium into a topologically nontrivial regime characterized by an exponential number of equilibria, the vast majority of which are expected to be unstable. It is suggested that this picture provides a global view on the nature of the May–Wigner instability transition originally discovered by local linear stability analysis.

The matrix elements \( J_{jk} \) are sampled from a probability distribution with zero mean and a prescribed variance \( \sigma^2 \). This is similar to the approach taken by Wigner in his description of statistics of energy levels of heavy nuclei via eigenvalues of random matrices, which proved to be very fruitful (3). A detailed review of May’s model in the light of recent advances in random matrix theory can be found in ref. 4.

The linear system [1] is stable if, and only if, all of the eigenvalues of \( J \) have real parts less than \( \mu \). Invoking Wigner’s arguments for studying eigenvalues of large random matrices, May claimed that, for large \( N \), the largest real part of the eigenvalues of \( J \) is typically \( \alpha \sqrt{N} \). Obviously, the model’s stability is then controlled by the ratio \( m = \mu / (\alpha \sqrt{N}) \). For large \( N \), system [1] will almost certainly be stable if \( m > 1 \) and unstable if \( m < 1 \), with a sharp transition between the two types of behavior with changing either \( \mu \), \( \alpha \), or \( N \). In particular, for fixed \( \mu \), \( \alpha \), system [1] will almost certainly become unstable for sufficiently large \( N \).

Despite the simplistic character of May’s model (5), his pioneering work gave rise to a long-standing “stability versus diversity” debate, which is not fully settled yet (4, 6–9), and played a fundamental role in theoretical ecology by prompting ecologists to think about special features of real multispecies ecosystems that help such systems to remain stable. Variations of May’s model are still being discussed nowadays in the context of neighborhood stability (see ref. 4 and references therein).

One obvious limitation of the neighborhood stability analysis is that it gives no insight into the model behavior beyond the instability threshold. Hence May’s model has only limited bearing on the dynamics of populations operating out of equilibrium. An instability does not necessarily imply lack of persistence: Populations could coexist thanks to limit cycles or chaotic dynamics of large complex systems.

Significance

Complex systems equipped with stability feedback mechanisms become unstable to small displacements from equilibria as the complexity (as measured by the interaction strength and number of degrees of freedom) increases. This paper takes a global view on this transition from stability to instability. Our examination of a generic dynamical system whereby \( N \) degrees of freedom are coupled randomly shows that the phase portrait of complex multicomponent systems undergoes a sharp transition as the complexity increases. The transition manifests itself in the exponential explosion of the number of equilibria, and the transition threshold is quantitatively similar to the local instability threshold. Our model provides a mathematical framework for studying generic features of the global dynamics of large complex systems.

Author contributions: Y.V.F. and B.A.K. performed research and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence may be addressed. Email: y.fyodorov@qmul.ac.uk or b.khoruzhenko@qmul.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601136113/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1601136113
attractors, which typically originate from unstable equilibrium points. Important questions posing serious challenges then relate to classification of equilibria by stability, studying their basins of attraction, and other features of global dynamics (4). Over the last few years, as computing power grew, nonlinear models have increasingly been used to investigate population dynamics on the global scale by means of numerical integration of the corresponding system of ODEs (10–13). Although such investigations captured a rich variety of types of behavior such as fold bifurcations when points of equilibrium merge/annihilate (14), or various types of chaotic dynamics (15, 16), they provide little analytic insight and are limited to small to medium-sized systems.

In this paper, we attempt to investigate the generic properties of the global dynamics of large complex multispecies systems by retaining in our model only the bare essentials—nonlinearity and stochasticity. Much in the spirit of May’s original approach, the model we propose is simple enough to allow for an analytic treatment yet, at the same time, is rich enough to exhibit a nontrivial behavior. In particular, our model captures an instability transition of the May–Wigner type, but now on the global scale. It also sheds additional light on the nature of this transition by relating it to an exponential explosion in the number of equilibria. Interestingly, despite the nonlinear setting of the problem, the random matrix ideas again play a central role in our analysis.

Similar to the May’s linear model, our toy model is likely to have rather limited practical significance for quantitative description of real ecosystems, but it might provide insight into the generic qualitative features of such systems and beyond, e.g., machine learning (17) or financial ecosystems (18, 19). The idea of destabilization by interaction is of relevance far beyond the mathematical ecology context, as applications of systems of many coupled nonlinear ODEs are vast [e.g., complex gene regulatory networks (20), neural networks (21, 22), or random catalytic reaction networks (23)].

Model

Consider a system of \( N \) coupled nonlinear autonomous ODEs of the form

\[
\frac{dx_i}{dt} = -\mu x_i + \sum_{j=1}^{N} f_j(x_1, \ldots, x_N), \quad i = 1, \ldots, N, \tag{2}
\]

where \( \mu > 0 \) and the components \( f_j(x) \) of the vector field \( f = (f_1, \ldots, f_N) \) are zero mean random functions of the state vector \( x = (x_1, \ldots, x_N) \). To put this model in the context of the discussion in the Introduction, if \( x \) is an equilibrium of [2], i.e., if \( -\mu x + f(x) = 0 \), then, in the immediate neighborhood of \( x \), system [2] reduces to May’s model [1] with \( y = x - x_e \) and \( f_k = (\partial f / \partial x_k)(x) \).

The nonlinear system [2] may have multiple equilibria whose number and locations depend on the realization of the random field \( f(x) \). To visualize the global picture, it is helpful to consider first a special case of a gradient descent flow, characterized by the existence of a potential function \( V(x) \) such that \( f = -\nabla V \).

In this case, system [2] can be rewritten as \( dx_i / dt = -\nabla V \), with \( L(x) = \mu |x|^2 / 2 + V(x) \) being the associated Lyapunov function describing the effective landscape. In the domain of \( L \), the state vector \( x(t) \) moves in the direction of the steepest descent, i.e., perpendicularly to the level surfaces \( L(x) = h \) toward ever-smaller values of \( h \). This provides a useful geometric intuition. The term \( \mu |x|^2 / 2 \) represents the globally confining parabolic potential, i.e., a deep well on the surface of \( L(x) \), which does not allow \( x \) to escape to infinity. At the same time, the random potential \( V(x) \) may generate many local minima of \( L(x) \) (shallow wells), which will play the role of attractors for our dynamical system. Moreover, if the confining term is strong enough, then the full landscape will only be a small perturbation of the parabolic well, typically with a single stable equilibrium located close to \( x = 0 \). In the opposite case of a relatively weak confining term, the disorder-dominated landscape will be characterized by a complicated random topology with many points of equilibria, both stable and unstable. Note that, in physics, complicated energy landscapes are a generic feature of glassy systems with intriguingly slow long-time relaxation and nonequilibrium dynamics (see, e.g., ref. 24).

The above picture of a gradient descent flow is, however, only a very special case, because the generic systems of ODEs [2] are not gradient. The latter point can easily be understood in the context of model ecosystems. For, by linearizing a gradient flow in a vicinity of any equilibrium, one always obtains a symmetric community matrix, whereas the community matrices of model ecosystems are, in general, asymmetric. Note also a discussion of an interplay between nongradient dynamics in random environment and glassy behavior in ref. 25.

To allow for a suitable level of generality, we therefore suggest choosing the \( N \)-dimensional vector field \( f(x) \) as a sum of gradient and nongradient (solenoidal) contributions,

\[
f(x) = \frac{\partial V(x)}{\partial x} + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\partial A_i(x)}{\partial x_i}, \quad i = 1, \ldots, N, \tag{3}
\]

where we require the matrix \( A(x) \) to be antisymmetric: \( A_{ij} = -A_{ji} \). The meaning of this decomposition is that vector fields can be generically divided into a conservative irrotational component, sometimes called “longitudinal,” whose gradient connects the attractors or repellers, and a solenoidal curl field, also called “transversal.” As discussed in, e.g., ref. 26, such a representation is closely related to the so-called Hodge decomposition of differential forms and generalizes the well-known Helmholtz decomposition of the three-dimensional vector fields into curl-free and divergence-free parts to higher dimensions. Correspondingly, we will call \( V(x) \) the scalar potential and the matrix \( A(x) \) the vector potential. The normalizing factor \( 1 / \sqrt{N} \) in front of the sum on the right-hand side in Eq. 3 ensures that the transversal and longitudinal parts of \( f(x) \) are of the same order of magnitude for large \( N \).

Finally, to make the model as simple as possible and amenable to a rigorous and detailed mathematical analysis, we choose the scalar potential \( V(x) \) and the components \( A_{ij}(x), \ i < j \), of the vector potential to be statistically independent, zero mean Gaussian random fields, with smooth realizations and the additional assumptions of homogeneity (translational invariance) and isotropy reflected in the covariance structure,

\[
\langle V(x)V(y) \rangle = a^2 \Gamma_V (|x-y|^2); \tag{4}
\]

\[
\langle A_{ij}(x)A_{kl}(y) \rangle = a^2 \Gamma_{A} \left( |x-y|^2 \right) \left( \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk} \right). \tag{5}
\]

Here the angular brackets \( \langle \ldots \rangle \) stand for the ensemble average over all realizations of \( V(x) \) and \( A(x) \), and \( \delta_{ij} \) is the Kronecker delta: \( \delta_{ij} = 1 \) if \( i = n \) and zero otherwise.

For simplicity, we also assume that the functions \( \Gamma_V(r) \) and \( \Gamma_A(r) \) do not depend on \( N \). This implies (27)

\[
\Gamma_V(r) = \int_0^\infty \exp(-sr) \gamma_V(s) ds, \quad \sigma = A, V,
\]

where the radial spectral densities \( \gamma_V(s) \geq 0 \) have finite total mass: \( \int_0^\infty \gamma_V(s) ds < \infty \). We normalize these densities by requiring that \( \int_0^\infty s^2 \gamma_V(s) ds = 1 \). The ratio

\[
\tau = v^2/(v^2+a^2), \quad 0 \leq \tau \leq 1,
\]

is a dimensionless measure of the relative strengths of the longitudinal and transversal components of \( f(x) \): If \( r = 0 \), then \( f(x) \) is divergence-free, and, if \( r = 1 \), it is curl-free.
Results

Determining and classifying all points of equilibria of a dynamical system with many degrees of freedom is a well-known formidable analytical and computational problem. In this paper, we shall focus our investigation on the simplest, yet informative, characteristic of system [2] by counting its total number of equilibria, that is, the total number $N_{\text{tot}}$ of solutions of the simultaneous equations

$$-\mu_i + f_i(x_1, \ldots, x_N) = 0, \quad i = 1, \ldots, N.$$  \[6\]

Certainly, finding $N_{\text{tot}}$ is a good starting point of any phase portrait analysis.

Had we restricted ourselves to the gradient descent flows, $N_{\text{tot}}$ would simply count the number of stationary points (minima, maxima, or saddle points) on the surface of the Lyapunov function $L(x)$. The problem of counting and classifying stationary points of high-dimensional random energy landscapes of various types has attracted considerable interest in recent years (28–33). In particular, works (28, 29) study such energy landscapes generated by a potential equivalent to the above Lyapunov function. One of the main conclusions of that study is that, for large $N$, the topology of the Lyapunov function changes drastically with decrease of the strength of the confining term relative to that of the interaction term in $L(x)$. The change manifests itself in the emergence of a multitude of equilibria, exponential in number. Such a transition is intimately connected to the spin-glass-like restructuring of the Boltzmann–Gibbs measure induced by the Lyapunov function when the latter is treated as an effective energy landscape.

We shall prove below that, for large $N$, the general autonomous system defined by Eqs. 2 and 3 exhibits a similar drastic change in the total number of equilibria when the control parameter

$$m = \frac{\mu}{\alpha \sqrt{N}}, \quad \text{where} \quad \alpha = 2\sqrt{\nu^2 + \alpha^2},$$

drops below the threshold value $m_c = 1$. As in the case of gradient systems, the proof involves the Kac–Rice formula as a starting point. However, performing the subsequent steps requires quite different mathematical techniques due to the asymmetry of the Jacobian matrix for nongradient systems.

The Kac–Rice formula (see, e.g., ref. 34) counts solutions of simultaneous algebraic equations. Under our assumptions (homogeneity, isotropy, and Gaussianity of $V$ and $A$), this formula yields the ensemble average of $N_{\text{tot}}$ in terms of that of the modulus of the spectral determinant of the Jacobian matrix $\langle J_i \rangle_{\text{aver}} = \delta_{ji} / \partial x_j$ (see Materials and Methods),

$$\langle N_{\text{tot}} \rangle = \frac{1}{m^N} \left| \langle \det(\mu J + J_i) \rangle \right|,$$  \[7\]

thus bringing the original nonlinear problem into the realms of the random matrix theory.

The probability (ensemble) distribution of the matrix $J$ can easily be determined in closed form. Indeed, the matrix entries of $J$ are zero mean Gaussian variables and their covariance structure, at spatial point $x$, can be obtained from Eqs. 4 and 5 by differentiation,

$$\langle J_{ij} \rangle_{\text{aver}} = \alpha \left[ (1 + \epsilon_N) \delta_{ii} \delta_{jj} + (\tau - \epsilon_N) (\delta_{ij} \delta_{ii} + \delta_{ji} \delta_{jj}) \right],$$

where $\epsilon_N = (1 - \tau) / N$. Thus, to leading order in the limit $N \to \infty$,

$$J_i = \alpha \left( X_i + \sqrt{\tau} \delta_i \xi \right),$$  \[8\]

where $X_i, i, j = 1, \ldots, N$ are zero mean Gaussians with

$$\langle X_i X_j \rangle_{\text{aver}} = \delta_{ii} \delta_{jj} + \tau \delta_{ij} \delta_{ii},$$  \[9\]

and $\xi$ is a standard Gaussian, $\xi \sim N(0,1)$, which is statistically independent of $X = (X_i)$. Note that, for the divergence free fields $f(x)$ (i.e., if $\tau = 0$), the entries of $J$ are statistically independent in the limit $N \to \infty$, exactly as in May’s model. On the other side, if $f(x)$ has a longitudinal component ($\tau > 0$), then this implies positive correlation between the pairs of matrix entries of $J$ symmetric about the main diagonal: $\langle X_i X_j \rangle = \tau$ if $i \neq j$. Such distributions of the community matrix have also been used in the neighborhood stability analysis of model ecosystems (8). Finally, in the limiting case of curl-free fields ($\tau = 1$), the matrix $J$ is real symmetric.

Representation [8] comes in handy, as it allows one to express [7] as a random matrix integral,

$$\langle N_{\text{tot}} \rangle = \frac{N}{m^N} \int_{-\infty}^{0} \langle \det(\mu J - X_i) \rangle_{X_i} \frac{e^{-\sqrt{x}}}{\sqrt{2\pi N}},$$  \[10\]

where $x = \sqrt{N} (m + t \sqrt{\tau})$ and the angle brackets $\langle \ldots \rangle_{X_i}$ stand for averaging over the real elliptic ensemble of random $N \times N$ matrices $X$ defined in Eq. 9; see also Eq. 20. This one-parameter family of random matrices interpolates between the Gaussian Orthogonal Ensemble of real symmetric matrices (GOE, $\tau = 1$) and real Ginibre ensemble of fully asymmetric matrices (rGinE, $\tau = 0$) (see ref. 35 for discussions). Both rGinE and its one-parameter extension defined in Eq. 9 have enjoyed considerable interest in the literature in recent years (36–40).

The matrix $X$ is asymmetric (unless $\tau = 1$) and can have real as well as complex eigenvalues. The latter come in complex conjugate pairs. Their density, in the limit $N \to \infty$, is constant inside the ellipse with the main half-axis $\sqrt{N} (1 \pm \tau)$ and vanishes sharply outside of the ellipse (35, 39, 41). The corresponding theorem is known as the Elliptic Law, and its validity extends beyond the Gaussian matrix distributions (42, 43). However, in the context of our investigation, it is the density of real eigenvalues of $X$ that appears to be most relevant.

Denote by $\rho_N^{(\tau)}(x)$ the density of real eigenvalues of $N \times N$ matrices $X$ [9] averaged over all realizations of $X$. It is convenient to normalize $\rho_N^{(\tau)}(x)$ in such a way that $\int_{-\infty}^{0} \rho_N^{(\tau)}(x) \, dx$ gives the average number of real eigenvalues of $X$ in the interval $[\alpha, \beta]$. A crucial observation is that $\rho_N^{(\tau)}(x)$ is directly related to the averaged value of the modulus of the determinant that appears in Eq. 10. Namely,

$$\langle \det(\mu J - X_i) \rangle_{X_i} = C_N^{(\tau)}(x) e^{-\sqrt{\tau}} P_{N+1}^{(\tau)}(x),$$  \[11\]

where $C_N^{(\tau)}(x) = 2 \sqrt{1 + \tau} (N - 1)! / (N - 2)!$ and $P_{N+1}^{(\tau)}(x)$ is the average density of real eigenvalues of matrices $X$ of size $(N + 1) \times (N + 1)$. For the limiting case $\tau = 0$, this relation appeared originally in ref. 44, and it can be extended to any $\tau \in [0,1]$ without much difficulty (see Supporting Information) for a derivation of Eq. 11 following the approach of ref. 45. In the limiting case of real symmetric matrices $\tau = 1$, all eigenvalues of $X$ are real and relation [11] is also valid (28).

Combining Eqs. 10 and 11 and changing the variable of integration from $t$ to $\lambda = m + t \sqrt{\tau}$, one can express $\langle N_{\text{tot}} \rangle$ for system [2] with $N$ degrees of freedom in terms of the density of real eigenvalues in the elliptic ensemble of random matrices [9] of size $(N + 1) \times (N + 1)$,

$$\langle N_{\text{tot}} \rangle = \frac{K_N^{(\tau)}}{m^N} \int_{-\infty}^{0} e^{-NS(\lambda)} P_{N+1}(\sqrt{\lambda}) \frac{d\lambda}{\sqrt{2\pi}}$$  \[12\]

where $S(\lambda) = (\lambda - m)^2 / (2(1 + \tau))$ and $K_N^{(\tau)}(\lambda) = N^{\frac{1}{2}+\frac{NS(\lambda)}{\sqrt{\tau}}}$. The importance of this relation is due to the fact that $\rho_N^{(\tau)}(x)$ is known, in closed form, in terms of Hermite polynomials (39). This allows us to carry out an asymptotic evaluation of the integral in [12] and calculate $\langle N_{\text{tot}} \rangle$ in the limit $N \to \infty$. The key finding that emerges from this calculation is that $\langle N_{\text{tot}} \rangle$ changes drastically around $m = 1$. If $m > 1$, then
\[ \lim_{N \to \infty} \langle N_{\text{tot}} \rangle = 1. \]  

On the other hand, if \(0 < m < 1\), then, to leading order in the limit \(N \to \infty\),

\[ \langle N_{\text{tot}} \rangle = \gamma e^{N \sum_{m} \langle m \rangle}, \]  

where \(\sum_{m} \langle m \rangle = (1/2)(m^2 - 1) - \ln m > 0\) for all \(0 < m < 1\). Therefore, if \(m < 1\), then \(\langle N_{\text{tot}} \rangle\) grows exponentially with \(N\). The factor in front of the exponential in Eq. 14 is given by \(\gamma = \sqrt{2/(1 + \tau)}(1 - \tau)\) as long as \(\tau < 1\). The gradient limit \(\tau \to 1\) can be approached by scaling \(x\) with \(N\). Setting \(\tau = 1 - u^2/N\), \(0 \leq u < \infty\), one obtains \(\gamma = 4\sqrt{N/\pi} \int_0^{\infty} e^{-u^2} du\). This regime describes a weakly nongradient flow. The corresponding regime for ensembles of asymmetric matrices was discovered long ago (46, 47).

Close to the phase transition point \(m = 1\), the complexity exponent vanishes quadratically, \(\sum_{m} \langle m \rangle = (1 - m^2)\) as \(m \to 1\), implying that the width of the transition region around \(m = 1\) is \(1/\sqrt{N}\). According to the general lore of phase transitions, for large but finite \(N\), there exists a “critical regime” \(m = 1 + N^{-1/2}\) where the number of equilibria changes smoothly between the two phases [13] and [14]. A quick inspection of Eq. 12 shows that the corresponding crossover profile is determined by the profile of \(\rho_0^i(x)\) in the vicinity of the “spectral edge” \(x = (1 + \tau)/\sqrt{N}\) (see Materials and Methods). After rescaling \(\lambda, \zeta \equiv 1 + \tau + \sqrt{1 - \tau}/\sqrt{N}\), the density \(\rho_0^i(\lambda/N)\) converges to \((1/\sqrt{1 - \tau}) \rho_{\text{edge}}^i(\zeta)\) in the limit \(N \to \infty\), where (39)

\[ \rho_{\text{edge}}^i(\zeta) = \frac{1}{2\sqrt{\pi}} \left[ \text{erf}[\sqrt{2\zeta}] + \frac{1}{2} e^{-\zeta^2} \right], \]  

with \(\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt\). In terms of \(\rho_{\text{edge}}^i(\zeta)\), the critical crossover profile is given by

\[ \langle N_{\text{tot}} \rangle = \gamma e^{\frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \rho_{\text{edge}}^i(c_t + \kappa \tau) \frac{dz}{\sqrt{2\pi}}}, \]  

where \(c_t = \sqrt{1/(1 - \tau)}\). The right-hand side of [16] interpolates smoothly between the two regimes [13] and [14], when parameter \(\kappa\) runs from \(\kappa = -\infty\) to \(\kappa = +\infty\).

Although our investigation is concerned with the ensemble average of the number of equilibria \(N_{\text{tot}}\), we expect that, in the limit \(N \to \infty\), the deviations of \(N_{\text{tot}}\) from its average \(\langle N_{\text{tot}} \rangle\) are relatively small. This is certainly the case above the critical threshold, for \(m > 1\). For, under some additional technical assumptions on the decay of correlations for \(f(x)\), system [2] will almost certainly have at least one stationary point (see ref. 48 for the relevant results about the maxima of homogeneous Gaussian fields). Therefore, \(N_{\text{tot}} \geq 1\), and the established convergence of \(\langle N_{\text{tot}} \rangle\) to 1 in the limit \(N \to \infty\) actually implies that the probability for \(N_{\text{tot}}\) to take other values than 1 is close to zero for large \(N\). The problem of estimating the deviation of \(N_{\text{tot}}\) from its average value in the opposite regime \(0 < m < 1\) is much harder and is an open and interesting question.¹

Discussion

In this paper, we introduced a model describing generic large complex systems and examined the dependence of the total number of equilibria in such systems on the system complexity as measured by the number of degrees of freedom and the interaction strength. The inspiration for our work came from May's pioneering study (1) of the neighborhood stability of large model ecosystems. Our outlook is complementary to that of May's in that it adopts a global point of view, which is not limited to the neighborhood of the presumed equilibrium.

In the context of model ecosystems, our analysis is applicable to complex multispecies communities in which each kind of species on its own becomes extinct and thus interaction between species is key to persistence of the community. The key feature of our analysis is that, in the presence of interactions, as the complexity increases, there is an abrupt change from a simple set of equilibria (typically, a single equilibrium for large number of species \(N \gg 1\)) to a complex set of equilibria, with their total number growing exponentially with \(N\). In the latter regime, we expect the stable equilibria to be only a tiny proportion of all of the multitude of equilibria (see discussion below), which is indicative of long relaxation times and transient nonequilibrium behavior.

We expect this sharp transition in the phase portrait to be shared by other systems of randomly coupled autonomous ODEs with large numbers of degrees of freedom. To that end, it is appropriate to mention that, very recently, a similar explosion in complexity was reported in a model of a neural network consisting of randomly interconnected neural units (22). The model considered in ref. 22 is essentially of form [2] but with the particular choice of \(f_i = \sum_j S(x_j)\) where \(S\) is an odd sigmoid function representing the synaptic nonlinearity and \(J_{ij}\) are independent centered Gaussian variables representing the synaptic connectivity between neuron \(i\) and \(j\). Although Gaussian, the corresponding (non-gradient) vector field is not homogeneous and thus seems rather different from our choice and not easily amenable to a rigorous analysis. Nevertheless, a shrewd semiheuristic analysis of ref. 22 revealed that, close to the critical coupling threshold, the two models actually display very similar behavior, described essentially by the same exponential growth in the total number of equilibria with rate \(\sum_{m} \langle m \rangle\). This fact points toward considerable universality of the transition from [13] to [14] and suggests that the crossover function [16] may be universal as well.

Our model captures an abrupt change in the dynamics of large complex systems on the macroscopic scale. At the same time, zooming in to classify each and every equilibrium point into locally stable or unstable seems a hard task. For, although linearizing the field \(f(x)\) around a given equilibrium is fairly straightforward, with the outcome being the Jacobian matrix \([8]\), conditioning by the positions of equilibria and taking into account all eventualities is a highly nontrivial task. Given the stochastic setup of our model, the question about stability of individual equilibria may even be the wrong question to ask, whereas addressing the statistics of the number of stable equilibria seems very appropriate.

Arguments similar to those in Results yield the ensemble average of the total number of stable equilibria, \(\langle N_{\text{st}} \rangle\), over all realizations of the vector field \(f(x)\) in terms of the random matrix integral (compare with Eq. 10),

\[ \langle N_{\text{st}} \rangle = \frac{N^2}{m^2} \int_{-\infty}^{\infty} \langle \det(x - \delta_0 - X) \chi_{\text{st}}(X) \rangle \chi_{\text{st}}(X) dX, \]  

where \(\chi_{\text{st}}(X) = 1\), if all \(N\) eigenvalues of matrix \(X\) have real parts less than the spectral parameter \(x = \sqrt{N} (m + 1/\sqrt{\tau})\), and \(\chi_{\text{st}}(X) = 0\) otherwise. In the limiting case of a purely gradient dynamics \(\tau = 1\), the rescaled Jacobian matrix \(X\) is real symmetric with all \(N\) eigenvalues real. In this case, the above integral can be related to the probability density of the maximal eigenvalue of the GOE matrix (29, 30), with the latter being a well-studied object in the random matrix theory (see, e.g., ref. 50 and references therein). This observation can then be used to evaluate \(\langle N_{\text{st}} \rangle\) for \(N \gg 1\).

---

¹In this context, we would like to mention the recent work of Subag (33), who proved that the deviations of \(N_{\text{tot}}\) from \(\langle N_{\text{tot}} \rangle\) in the spherical \(p\)-spin model are negligible in the limit of large system size. Although that model is different from ours, it is not dissimilar to the gradient limit of \(\tau = 1\) of our model (22); for instance, the average number of equilibria grows exponentially with \(N\) (38). Thus one might hope to adopt the technique of ref. 33 to our model. Another relevant reference is ref. 49.
One finds (29) that \( \langle N_{e\tau} \rangle \to 1 \) if \( m > 1 \), whereas, if \( 0 < m < 1 \), then, to leading order in \( N_{e\tau} \), \( \langle N_{e\tau} \rangle \propto e^{N_{e\tau}} \), with \( 0 < \Sigma_{e\tau} < \Sigma_{\text{tot}} \). Thus, in the case of purely gradient dynamics, as the complexity increases, large nonlinear autonomous systems assembled at random undergo an abrupt change from a typical phase portrait with a single stable equilibrium to a phase portrait dominated by an exponential number of unstable equilibria with an admixture of a smaller, but still exponential, in \( N_{e\tau} \), number of stable equilibria.

It was suggested to us by J.-P. Bouchaud that, in the general case of nonequilibrium dynamics \( 0 \leq r < 1 \), it would be natural to expect a further phase transition in the plane \((m, r)\) such that, below a certain number \( r_{c}(m)\), stable equilibria are no longer exponentially abundant in the limit \( N_{c} \to \infty \), with further implications for the global dynamics. Unfortunately, for a fixed \( 0 \leq r < 1 \), only a vanishing fraction of order \( N^{1/2} \) of eigenvalues of \( X \) remain real, and the relation of the integral in Eq. 17 to statistics of the largest real eigenvalue in the elliptic ensemble seems to be lost. This fact has prevented us, so far, from reliable counting of stable equilibria for the general case of nonequilibrium flows. In principle, for given values of parameters \( N_{c}, r, m \), one may attempt to evaluate the random matrix ensemble average in the integral in [17] numerically, and then evaluate numerically the integral itself. Although such a procedure seems tractable, its actual implementation with sufficient precision is not straightforward, especially in the limit \( N_{c} \to \infty \), due to the exponentially large values involved. Clarification of the status of the picture suggested by J.-P. Bouchaud and related studies remain an important outstanding issue.

Materials and Methods

Kac–Rice Formula. The expected number \( \langle N_{e\tau} \rangle \) of simultaneous solutions to the system of Eq. 6 in \( \mathbb{R}^{d} \) is given by the formula (see, e.g., ref. 34)

\[
\langle N_{e\tau} \rangle = \int_{\mathbb{R}^{d}} \langle \delta(-\mu x + f(x)) \rangle \langle \det(-\rho \partial_{j} + J_{ij}(x)) \rangle \, dx^{N},
\]

where \( \delta(x) \) is the multivariate Dirac \( \delta \)-function, \( dx^{N} \) is the volume element in \( \mathbb{R}^{N_{c}} \), and \( J_{ij}(x) = \partial_{j} f_{i} / \partial_{k} \) are matrix elements of the Jacobian matrix \( J = (J_{ij}) \). By our assumptions, the random field \( f(x) \) is homogenous and isotropic. For such fields, samples of \( f \) and \( J \) taken in one and the same spatial point \( x \) are uncorrelated, \( \langle f_{i} \partial_{j} f_{k} \rangle = 0 \) for all \( i, j, k \). This is well known and can be checked by straightforward differentiation. In addition, the field \( f \) is Gaussian; hence the \( f(x) \) and \( J(x) \) are actually statistically independent. This simplifies the evaluation of the integral in [18] considerably. Indeed, the statistical average in [18] factorizes and, because \( \langle \det(-\rho \partial_{j} + J_{ij}(x)) \rangle \) does not vary with \( x \), the integrand effectively reduces to

\[
\langle \delta(-\mu x + f(x)) \rangle = \int \frac{dx^{N}}{(2\pi)^{d/2}} e^{-\langle k \rangle \cdot k}.
\]

Furthermore, at every spatial point \( x \), the vector \( f(x) \) is Gaussian with uncorrelated and identically distributed components,

\[
\langle f_{i}(x) f_{j}(x) \rangle = \delta_{i j} \sigma^{2}, \quad \sigma^{2} = 2k^{2} \langle \Gamma(0) \rangle = 2k^{2} / \langle \Gamma(0) \rangle = N^{-1} \frac{N}{N}.
\]

Therefore, \( e^{-\langle k \rangle \cdot k} \) is \( e^{-\sigma^{2} \langle k \rangle} / \sqrt{2\pi} \), and, after evaluating the integral on the right-hand side in [19], one arrives at [7].

Real Elliptic Matrices and Asymptotics of \( \langle N_{e\tau} \rangle \). The joint probability density (JPD) function \( P_{N_{e\tau}}(X) \) of the matrix entries in the elliptic ensemble of real Gaussian random matrices \( X \) of size \( N \times N \) is given by

\[
P_{N_{e\tau}}(X) = Z_{N}^{N} \exp \left[ -\frac{1}{2(1+r)^{N}} \text{Tr}(XX^{T} - X^{2}) \right],
\]

where \( Z_{N} \) is the normalization constant and \( r \in (0,1) \). It is straightforward to verify that the covariance of matrix entries \( X_{ij} \) is given by the expression specified in [9]. The mean density of real eigenvalues of \( \rho_{N_{e\tau}}^{(i)}(x) \) in the elliptic ensemble [20] is known, in closed form, in terms of Hermite polynomials (see ref. 39). Assuming, for simplicity, that \( N+1 \) is even, one has \( \rho_{N_{e\tau}}^{(i)}(x) = \rho_{N_{e\tau}}^{(i)}(x) + \rho_{N_{e\tau}}^{(i)}(x) \), where

\[
\rho_{N_{e\tau}}^{(i)}(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{|p_{k}^{(i)}(x)|^{2}}{k!}
\]

and

\[
\rho_{N_{e\tau}}^{(i)}(x) = \frac{1}{\sqrt{2\pi(1+r)^{(N-1)}}} \int_{0}^{x} \psi_{N}^{(i)}(u) \, du.
\]

Here \( \psi_{N}^{(i)}(x) = e^{-x^{2}/2(1+r)} H_{0}^{(i)}(x) \) and \( H_{0}^{(i)}(x) \) are rescaled Hermite polynomials, \( h_{0}^{(i)}(x) = (1/\sqrt{2}) \int_{0}^{x} e^{-x^{2}/2} x^{i} \, dx \). This, together with Eq. 12, allows one to evaluate \( \langle N_{e\tau} \rangle \) in the limit \( N \to \infty \). We shall sketch the corresponding evaluation below.

The asymptotics of \( \rho_{N_{e\tau}}^{(i)}(x) \) in the bulk and at the edge of the support of the distribution of real eigenvalues in the real elliptic ensemble were found in ref. 39, and, outside of the support, it can also be readily extracted using [21] and [22]. In particular, in the bulk, i.e., for \( x < (1 + r)^{-1} N \), the contribution of [21] to \( \rho_{N_{e\tau}}^{(i)}(x) \) is dominant, and, to leading order in \( N \),

\[
\rho_{N_{e\tau}}^{(i)}(x) \sim \frac{1}{2\pi(1+r)^{(N-1)}} \int_{0}^{x} \psi_{N}^{(i)}(u) \, du.
\]

At the same time, for \( x > (1 + r)^{-1} N \), both [21] and [22] yield exponentially small contributions to \( \rho_{N_{e\tau}}^{(i)}(x) \). Allowing one to evaluate \( \langle N_{e\tau} \rangle \) in the parameter range \( 0 < m < 1 \). For \( m > 1 \), the saddle point occurs in the domain \( j_{e} > 1 + r \) so that the analysis requires search for the minimum of \( S(j_{e}) + \Psi(j_{e}) \). After straightforward algebra, we find (d/dx) \( S(j_{e}) + \Psi(j_{e}) \) = \( 1/(2(1+r)) (x^{2} - 4 - 4r) - m\langle \Gamma(0) \rangle \), which is equal to zero at \( j_{e} = m = m + (m+r) = 1 + r \). One also verifies that this is a point of minimum for \( S(j_{e}) + \Psi(j_{e}) \), and a further simple calculation then yields \( S(j_{e}) + \Psi(j_{e}) \) = \( -m\langle \Gamma(0) \rangle \). Calculating the saddle point contribution then yields [39].

The above asymptotic analysis assumes that \( 0 \leq r < 1 \). Let us now discuss the modifications required to study the scaling regime of weakly non-gradient flow \( r \to 1 \) for \( 0 < m < 1 \). We only need to evaluate the leading contribution to \( \rho_{N_{e\tau}}^{(i)}(x) \) given by [21]. By making use of the above integral representation for the scaled Hermite polynomials \( H_{0}^{(i)}(x) \) and applying the scaling \( x = u^{1/2} / \sqrt{r} \), we can write

\[
\rho_{N_{e\tau}}^{(i)}(z) = \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-z)^{n}}{n!} \Phi_{n}(z) / \sqrt{n!} / \sqrt{n!} / \sqrt{n!},
\]

where \( \Phi_{n}(z) = e^{-z^{2}/2} \int_{0}^{z} e^{-v^{2}/2} v^{n} \, dv \). Recalling that the limit of \( \Phi_{n}(z) \) as \( N \to \infty \) is \( 1 / (2\pi)^{n/2} \), one obtains

\[
\rho_{N_{e\tau}}^{(i)}(z) \sim \frac{1}{\sqrt{2\pi}} \int_{0}^{z} e^{-u^{2}/2} \, du.
\]

Substituting this expression into the integrand in [12] and evaluating the integral in the limit \( N \to \infty \) (hence, \( r \to 1 \)) by the Laplace method then yields \( \langle N_{e\tau} \rangle \) in the weakly nongradient regime.

Finally, our calculation of the profile \( \langle N_{e\tau} \rangle \) in the transitional region \( m = 1 + \epsilon N^{-1/2} \) uses the fact that, in such a regime, the main contribution to
the integral [12] comes from the neighborhood of the spectral edge, \( x = 1 + r + \sqrt{1 - r^2} / N \), where we have, to the leading order in \( N \),

\[
\frac{e^{-\beta[\log\xi^2 - 2\tau N]} e^{-\beta [z^2 + c^2] + \sqrt{1 - r^2} / \tau}}{m^N} = \exp \left[ -2r \left( x^2 + c^2 \right) + \sqrt{1 - r^2} / \tau \right].
\]

This, together with [26] and [15], converts [12] to [16].

### ACKNOWLEDGMENTS

The authors thank David Arrowsmith, Matthew Evans, and Vito Latora for comments on various versions of this paper and J.-P. Bouchaud for his comments on the number of stable equilibria in nongradient systems. The research of Y.V.F. was supported by EPSRC Grant EP/K002763/1, “Insights into Disordered Landscapes via Random Matrix Theory and Statistical Mechanics.” B.A.K. thanks the Isaac Newton Institute for Mathematical Sciences for its hospitality during the program Periodic and Ergodic Spectral Problems, supported by EPSRC Grant EP/K032208/1.
Supporting Information

Fyodorov and Khoruzhenko 10.1073/pnas.1601136113

SI Text

In this supporting information, we express the mean density of real eigenvalues in the real elliptic ensemble (see Eq. 9) of \( N \times N \) matrices \( X_N \) in terms of the modulus of the spectral determinant of real elliptic matrices of size \( (N-1) \times (N-1) \). Our derivation is based on the method suggested in ref. 44; however, our Jacobian computation differs from the one given in ref. 44 and may be of interest on its own. For a similar calculation in the context of complex matrices, see ref. 45.

Houssholder Reflections and Partial Triangularization of Real Matrices

The key idea of ref. 44 is based on using Houssholder reflections described by matrices

\[
P_v = I_N - 2v v^T, \quad v^T v = 1,
\]

where \( I_N \) is the \( N \times N \) identity matrix, and \( v \) is a column vector in \( \mathbb{R}^N \) of unit length. Its transpose, \( v^T \), is a row vector, so that the Kronecker product \( v \otimes v^T \) is a matrix, \( P_v \) describes the reflection about the hyperplane with normal \( v \) and passing through the origin. Obviously, \( P_v \) is symmetric and orthogonal: \( P_v^2 = P_v \) and \( P_v^T = I_N \).

Let \( e_1 \) be the first vector of the standard basis in \( \mathbb{R}^N \), i.e., \( e_1^T = (1,0, \ldots, 0) \). For any unit vector \( x, |x| = 1 \), define

\[
v = \frac{x + e_1}{\sqrt{2(1 + x_1)}}
\]

and consider the Houssholder reflection \( P_v \) built from the above \( v \) according to Eq. S1. Then it is easy to check that \( P_v x = e_1 \). This implies that, for any nonzero vector \( x \), there exists a Houssholder reflection such that \( P_v x = e_1 \), with \( k = |x| \).

Let \( \lambda \) be a real eigenvalue of the \( N \times N \) matrix \( A^{(N-1)} \) with real entries \( A^{(N-1)} \), i.e., \( A^{(N-1)} x = \lambda x \) for some column \( N \)-vector \( x \) of unit length. Our goal is to demonstrate that it is always possible to represent that matrix as

\[
A^{(N-1)} = P \begin{pmatrix} \lambda & w^T \\ 0 & A^{(N-1)} \end{pmatrix} P^T
\]

for some real \((N-1)\)-component vector \( w \) and a real matrix \( A^{(N-1)} \) of size \((N-1) \times (N-1)\). To verify this, we start with the left-hand side as \( P_A^{(N-1)} P v x = kP_A^{(N-1)} P e_1 \), which, after denoting \( A = P A^{(N-1)} P_0 \), implies \( A e_1 = \lambda e_1 \). Using the definition of \( e_1 \), we see that \( A_{11} = \lambda \) and \( A_{1j} = 0 \) for all \( j = 2, \ldots, N \). Therefore, \( A \) can be written as \( \hat{A} = \begin{pmatrix} \lambda & w^T \\ 0 & A^{(N-1)} \end{pmatrix} \), and, as \( A^{(N)} = P A \), the relation [S3] follows.

Considering now the volume element \( dA^{(N-1)} = \prod_{n=m} dA^{(N-1)} \), our next goal is to write it down in terms of \( N^2 \) independent variables parametrizing the right-hand side of [S3], that is, \((N-1)^2\) variables parametrizing \( A^{(N-1)} \), \( N-1 \) components of \( w \), one parameter for \( \lambda \) and the remaining \( N - 1 \) parameters for representing the matrix \( P \). A convenient parametrization for \( P \) comes from using [S2] for the vector \( v \), which shows that the last \( N-1 \) components of that vector \( (v_2, \ldots, v_N)^T \) \( \equiv q \) can be used as independent variables, whereas normalization fixes the first component. Writing \( v = (\sqrt{1 - q_1 q_1}, q_1) \) with \( |q| < 1 \) and using [S1] yields an explicit parametrization,

\[
P = \begin{pmatrix} 2q^T q - 1 & -2q^T \sqrt{1 - q_1 q_1} \\ -2q \sqrt{1 - q_1 q_1} & I_{N-1} - 2q \otimes q^T \end{pmatrix}.
\]

The problem therefore amounts to calculating the Jacobian of the transformation \( A^{(N)} \rightarrow (\lambda, w, q, A^{(N-1)}) \). To that end, we start with differentiating [S3], which gives

\[
dA^{(N)} = P \left\{ \begin{pmatrix} P \left[ \frac{\partial v}{\partial q} \right]_{v^T} \\ 0 \end{pmatrix} \right\} + \begin{pmatrix} \partial \lambda & dwd^T \\ 0 & dP \end{pmatrix} \left\{ \begin{pmatrix} \partial \lambda & dwd^T \\ 0 & dP \end{pmatrix} \right\} P^T,
\]

where we made use of \( dP = -P dP \) and also used the notation \( [A, B] = AB - BA \) for the matrix commutator. A direct calculation using [S4] shows that the matrix \( (P dP) \) can be symbolically written as

\[
(P dP) = \begin{pmatrix} 0 & -db^T \\ db & df \end{pmatrix}.
\]

Substituting [S5] into the expression for \( dA^{(N)} \), we find that

\[
dA^{(N)} = P \begin{pmatrix} \partial \lambda - \frac{w^T}{\lambda - A^{(N-1)}} & dwd^T \\ \frac{w^T}{\lambda - A^{(N-1)}} & dP \end{pmatrix} \left\{ \begin{pmatrix} \partial \lambda - \frac{w^T}{\lambda - A^{(N-1)}} & dwd^T \\ \frac{w^T}{\lambda - A^{(N-1)}} & dP \end{pmatrix} \right\} P^T.
\]

From this expression, we easily read off the required Jacobian to be given by

\[
\text{Jacobian} = \left| \text{det} \left( \frac{\partial\lambda}{\partial q_i} \right) \right| \left| \frac{\partial \lambda}{\partial q_i} \right|,
\]

where the last factor symbolically denotes the part of the Jacobian coming from the transformation \( db \rightarrow dq \) described in [S6]. A straightforward calculation shows that

\[
\frac{\partial \lambda}{\partial q_i} = 2N(1 - q^T q)^{1/2} - 1,
\]

so that finally we arrive at the change-of-variables formula

\[
dA^{(N)} = 2N(1 - q^T q)^{1/2} \left| \text{det} \left( \frac{\partial\lambda}{\partial q_i} \right) \right| dq^{N-1} dA^{(N-1)}.
\]

Elliptic Ensemble of Gaussian Random Matrices

The JPD of the elliptic ensemble of Gaussian random matrices \( X_N \) of size \( N \times N \) whose entries have covariance \( \langle x_{ij} x_{im} \rangle = \delta_{ij} \delta_{im} + \tau \delta_{jm} \delta_{im} \) is given by

\[
P(X_N) = Z_N^{-1} \exp - \frac{1}{2(1 - \tau)} \left[ \text{Tr}(X_N X_N^T) - \tau \text{Tr}(X_N^T X_N) \right].
\]
where $Z_N$ is the corresponding normalization factor

$$Z_N = 2^{N/2} \pi^{N(N+1)/2} (1 + \tau)^{N(N+1)/4} (1 - \tau)^{N(N-1)/4}.$$ 

Our goal is to find the JPD of variables $\lambda, w, q, X_{N-1}$ used to perform the partial triangulation of $X_N$ via $[S3]$, with the role of $A^{(N)}$ played now by $X_N$. We have

$$\text{Tr}(X_N X_N^T) = \lambda^2 + w^T w + \text{Tr}(X_{N-1} X_{N-1}^T), \quad \text{Tr}(X_N^2) = \lambda^2 + \text{Tr}(X_{N-1}^2).$$

Taking into account the change-of-variables formula $S8$, we see that the corresponding JPD can be written as

$$P(\lambda, w, q, X_{N-1}) = 2^N (1 - q^T q) \frac{Z_N}{Z_{N-1}} e^{-\frac{1}{2} \tau^T w^T w}$$

By definition, the density of real eigenvalues $\rho^{(r)}_N(\lambda)$ is obtained by integrating the above JPD over variables $|q| < 1, -\infty < w < \infty$ and finally over $X_{N-1}$. After performing the integrals over $q$ and $w$, we immediately arrive at the relation

$$\rho^{(r)}_N(\lambda) = \frac{1}{C_{N-1}(\tau)} e^{-\frac{1}{2} \tau \tau^T} |\text{det}(\lambda I_N - X_{N-1})|_{X_{N-1}},$$

where $C_{N-1}(\tau)$ is a certain constant that can be found explicitly. This is precisely equivalent to Eq. 11, with the obvious change of notations: $\lambda \to x$ and $N \to N + 1$. 

Fyodorov and Khoruzhenko www.pnas.org/cgi/content/short/1601136113 2 of 2