Expanded view of ecosystem stability: A grazed grassland case study

Gidon Eshel¹, Yohay Carmel²

¹ Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA, United States of America,
² Dept. of Agricultural Engineering, the Technion, Haifa, Israel

* geshel@gmail.com

Abstract

Analysis of stability under linearized dynamics is central to ecology. We highlight two key limitations of the widely used traditional analysis. First, we note that while stability at fixed points is often the focus, ecological systems may spend less time near fixed points, and more time responding to stochastic environmental forcing by exhibiting wide zero-mean fluctuations about those states. If non-steady, uniquely precarious states along the nonlinear flow are analyzed instead of fixed points, transient growth is possible and indeed common for ecosystems with stable attractive fixed points. Second, we show that in either steady or non-steady states, eigenvalue based analysis can misleadingly suggest stability while eigenvector geometry arising from the non-self-adjointness of the linearized operator can yield large finite-time instabilities. We offer a simple alternative to eigenvalue based stability analysis that naturally and straightforwardly overcome these limitations.

Introduction

Ecosystem stability is vigorously studied, emphasizing both mathematical [1–11] and ecological [12–17] aspects. Ecosystem stability research is broad and diverse, and “ecological stability” can have such diverse meanings as constancy, persistence, robustness or resilience [13]. Here, we (1) highlight two key limitations of ecological stability analysis as widely (but not universally) practiced; (2) demonstrate their characteristic manifestations; and (3) offer simple yet highly effective strategies for overcoming these pitfalls. Taken together, the key message of these three thrusts is simple enough: basing stability decisions on the leading eigenvalue alone —implicitly considering perturbation fates over infinitely long times—can mislead the analyst into expecting stability where perturbations can in fact grow over ecologically relevant time scales.

Traditionally, “stability” and “resilience” are used quite synonymously in the ecological literature, yet with (at least) two different meanings. The first addresses whether or not an ecosystem will return to its original stable state after perturbation, and if so, the rate of perturbation decay (e.g., [18]). The second strives to quantify the magnitude of disturbance required to drive the system toward a different stable state (presumably a distinct attractor in configuration space, e.g., [19], later termed “ecological resilience” by Holling [20]).
Here, by “stability” we mean the rate at which the system returns to a reference state (or dynamic trajectory) after a transient imposed disturbance. Despite their diversity, most analyses of ecosystem stability—including those sharing our stability definition in the context of differential equations based models—have two technical commonalities (with few exceptions highlighted below). First, stability determination is eigenvalue based, thus addressing the time $\to \infty$ limit. Yet analyses of geophysical [21–28] and more recently ecological [29–31] dynamical systems have demonstrated the limitations of eigenvalue based stability determination over finite timescales. These limitations, and the complementary importance of finite time horizons for stability determination are of wide ecological relevance (emphasized by, e.g., Hasting [15] or Cortez [32]). Extending stability analysis beyond eigenvalues, the above papers introduced the key and often ignored role of eigenvector geometry. Yet eigenvector geometry was secondary to the broader thrusts of [29–31], and consequently not emphasized. Moreover, neither strove to provide an introductory, accessible explanation of the mechanisms by which eigenvector geometry contributes to transient growth (or “reactivity” in mathematical ecology [29, 31]). This may explain the observation that most of the hundreds of papers citing the above ecological papers address other elements of the work (notably spatial connectivity).

Appreciation of the unique importance of eigenvector geometry to finite time ecosystem stability analysis is therefore yet to take deep, widespread roots in ecology, despite the potential mechanistic insights it can offer [31]. An explicit, systematic demonstration and explanation of the limitations of eigenvalue based stability analysis (whose precision [30] characterize in their Table 1 as “poor”) is therefore needed and timely. Providing such an introduction is one of our two objectives here.

Second, most ecosystem stability work addresses nonlinearly stable configurations, fixed points or limit cycles (e.g., [1, 29]). The view of the system underlying this focus is that noise—unusual rains and droughts, or N flushes from upstream uplands—amounts to slight, rapidly linearly self correcting departures from the steady state, allowing the nonlinear governing dynamics to nearly always maintain a stable state, which is thus assumed to dominate system history. The key question is then the system’s propensity to linearly damp, or amplify, small perturbations about these nonlinearly stable states, where the state is assumed to nearly universally reside. From an ecological standpoint, this is problematic because even an attractive stable fixed point may be of little relevance if the state trajectory toward the stable states is constantly perturbed by environmental noise. Further, the nonlinear trajectories of the ecosystem may oscillate widely, at times resulting in very low populations at which ecosystem collapse by linear responses to imposed perturbations can occur independently of whether or not the system possesses nonlinearly stable configurations. The fate (growth or decay) of imposed ecological anomalies far from any resting state can thus be of great ecological interest, yet is not often emphasized in the ecological literature. Demonstrating the potentially precarious dynamical balance at rapidly changing states that are the antithesis of fixed points, and comparing their stability properties to those of fixed points, jointly constitute our second main objective here.

In this paper, we demonstrate both of the above technical limitations of eigenvalue based stability analysis in the context of a canonical ecological system, paying close attention to wide technical accessibility. The paper’s key message is that for high-dimensional systems—systems whose state is given by an $N$-vector ($N > 1$) and not a scalar ($N = 1$)—eigenvalues alone rarely tell the whole stability story, and that for a more complete understanding of the full stability status of the analyzed system, carefully considering eigenvector geometry is essential. We also advance a simple and readily available alternative stability analysis framework [26]. Let us now introduce briefly the limitation of eigenvalue based stability decisions.
Consider an unconsumed (ungrazed) grass population $x$ (in kg ha$^{-1}$) evolving nonlinearly under the standard logistic model (e.g., [33]),

$$\frac{dx}{dt} = \gamma_x(x) = x \left(1 - \frac{x}{K}\right) \tag{1}$$

(where we use “evolving” not in its biological sense of species evolution, but in its dynamical systems sense, the changing system state as it progresses in time under its governing dynamics). In Eq (1), $t$ is time and the leftmost term is thus grass rate of change, $\gamma_x$ is the $x$-dependent rate of grass growth, controlled by the carrying capacity $K$ and by $x$, the intrinsic resource unlimited grass growth when $x \ll K$ (the fastest grass growth possible, attained when resident grass cover is well below the carrying capacity). Initial grass population evolution in time is shown in Fig 1a (where again “evolution” refers not to biological species evolution, but to the changes the system state exhibits in time under the repeated cumulative impact of the system’s governing dynamics).

Now imagine that at some point during the evolution of this ecosystem under its governing equation, an external disturbance arises by, e.g., a dry spell, inducing a population change. Denoting this population perturbation $x_0'$ ($x_0' < 0$ following a drought), the full grass population at any subsequent time is $x = \bar{x} + x'$, where $\bar{x}$ is the expected grass cover based on the
nonlinear dynamics, i.e., \( \frac{d\bar{x}}{dt} = \lambda (1 - \bar{x}/K) \). The stability question addresses the fate of the perturbation at some time of interest \( \tau \) after the perturbation arose. To avoid confusion, throughout this paper we use \( t \) to denote time of nonlinear evolution of the state vector \( \mathbf{x} \) (of which the above \( x \) is a one-dimensional special case) since the initial state \( \mathbf{x}_0 \equiv \mathbf{x}(0) \) at \( t = 0 \), and \( \tau \) to measure time of linearized growth of \( \mathbf{x} \) since its excitation by an externally imposed perturbation as \( \mathbf{x}' \) at \( \tau = 0 \). The stability questions are thus: Will \( x_0 \) linearly decay, allowing \( x \) to return to its unperturbed value \( \bar{x} \) over a reasonably short period of linearized evolution \( \tau \)? Or alternatively, will \( x_0 \) grow, forcing the population onto an altogether different trajectory?

To answer this, we expand the right hand side of Eq (1) in a Taylor series,

\[
\frac{dx}{dt} = \frac{d\bar{x}}{dt} + \frac{dx'}{dt} = \gamma_\lambda(\bar{x} + x') = \gamma_\lambda(\bar{x}) + \frac{d\gamma_\lambda}{dx} \bigg|_{\bar{x}} x' + \frac{1}{2} \frac{d^2\gamma_\lambda}{dx^2} \bigg|_{\bar{x}} x'^2 + \ldots
\]

(2)

where “…” denotes cubic or higher order terms. Because at least near its inception the perturbation is small, \( x'^2 \ll x' \ll \bar{x} \), and we can truncate the series after the first order,

\[
\frac{d\bar{x}}{dt} + \frac{dx'}{dt} \approx \gamma_\lambda(\bar{x}) + \frac{d\gamma_\lambda}{dx} \bigg|_{\bar{x}} x'
\]

(3)

Invoking \( d\bar{x}/dt = \gamma_\lambda(\bar{x}) \) introduced above, this reduces to

\[
\frac{dx'}{dt} = \frac{d\gamma_\lambda}{dx} \bigg|_{\bar{x}} x' = \lambda \left( 1 - \frac{2\bar{x}}{K} \right) x'
\]

(4)

In this one-dimensional case, stability is thus straightforward: stability for \( \lambda (1 - 2\bar{x}/K) < 0 \), instability for \( \lambda (1 - 2\bar{x}/K) > 0 \), as shown in Fig 1b.

The key issue we wish to illustrate in this paper is that this simplicity is lost in higher dimensional systems such as most real ecosystem. That is, while the eigenvalues are always an essential part of the stability story, in most ecosystems of fundamental and practical interest they by no means tell the whole story. Instead, eigenvector geometry is also essential. To be sure, the stability parameter of the above one dimensional example has an \( N \)-dimensional analog, the \( N \) eigenvalues of the linearized operator, the Jacobian (the \( N \)-dimensional analog of the scalar \( dy_\lambda/dx \) above). Nonetheless, tempting though it may be to use eigenvalues as the sole arbiter of stability even for \( N > 1 \) dimensional systems, this can often mislead into expecting stability where linearized perturbation growth is in fact expected. Using the lowest dimensional and thus simplest example necessary to demonstrate the points, the remainder of the paper demonstrates the limitations of eigenvalue based stability analysis in \( N > 1 \) dimensions and the ameliorating potential of careful consideration of eigenvector geometry.

**Methods**

An unambiguous discussion requires an ecologically relevant, suitable testbed model widely familiar to ecologists. We choose one with venerable history (e.g., [34–37]) that is compactly manageable and mechanistically comprehensible, yet that exhibits the relevant characteristics necessary for demonstrating the points.

The model describes a grazed grassland comprising an herbivore species subsisting on one grass type in one location. Slightly modified from [38], the governing equation is \( \frac{d\mathbf{x}}{dt} = \left( \begin{array}{c} \gamma_s - \rho_s \\ \gamma_r - \rho_r \end{array} \right) \mathbf{F}(\mathbf{x}) \),

\[
\frac{d\mathbf{x}}{dt} = \left( \begin{array}{c} \gamma_s - \rho_s \\ \gamma_r - \rho_r \end{array} \right) \left( \begin{array}{c} \lambda x(1 - x/K) - axy/(b + x) \\ a \lambda xy/(b + x) - dy \end{array} \right) \]

(5)

\[ \mathbf{F}(\mathbf{x}) \]
the 2D generalization of Eq (1). In Eq (5), $x = (x_y)^T$ is the state vector holding the resident grass and herbivore mass densities in kg ha$^{-1}$, $t$ is time, and $\gamma_x$ and $\gamma_y$ ($\rho_x$ and $\rho_y$) denote x- dependent rates of grass and herbivore growth (removal by grazing and death) respectively. We use three partially overlapping sets of parameter values ($\{x, k, a, b, c, d\}$, reported in S1 Table).

Because all other technical details are standard, we relegate them to appendices, as follows. S1 Appendix offers further details of Eq (5). S2 Appendix describes the solution of the nonlinear equations over $t = [0, 10]$ years, and the nontrivial fixed points are derived in S3 Appendix. Finally, because the crux of the matter is ecosystem stability analysis, in S4 Appendix we derive the Jacobian of Eq (5)’s $F(x)$ (the 2D analog of $dy_x/dx$ in Eq (4)), and then use it to derive the governing equation of perturbation growth (Eq (3) of S4 Appendix, the 2D analog of Eq (4)).

Asymptotic vs. Transient stability: The key role of non-self-adjointness

Traditionally, stability is categorically decided based on a single criterion (e.g., [39]): if the eigenvalue with the largest real part (which we denote $\lambda_{\text{max}}$) satisfies $\Re(\lambda_{\text{max}}) < 0$ (i.e., its real part is negative), the system is said to be stable. Rarely explicitly stated, yet absolutely essential is that this must be modified to “asymptotically stable for $t \to \infty$” [17, 21–31].

To demonstrate the limitations of the exclusive reliance on eigenvalues for finite times $\tau \ll \infty$, we use the linear perturbation equation

$$\frac{dx'}{dt} \approx J_x x' + f(t)$$

derived in S4 Appendix. Because we are interested in the linearized evolution of individual perturbations, we envision the forcing term $f(t)$ as the origin of the perturbations, the mechanism by which the state is locally displaced from the nonlinear flow. Once a specific perturbation has arisen, however, its fate is determined by the locally valid linearized dynamics, not subsequent forcing. Consequently, below we omit $f$, but Farrell and Ioannou [26–28] present a fully developed general theory of the statistical time-integrated response of nonnormal systems to continuous full spectrum stochastic forcing and the variance maintenance consequences of this response.

Taking the time integral of the unforced linearized dynamics (Eq (3) of S4 Appendix with $f = 0$) over $[0, \tau]$, and using eigenvalue/eigenvector decomposition,

$$x'(\tau) = e^{J_x \tau} x_0 = \mathbf{E} \Delta \mathbf{E}^{-1} x_0$$

Here $x_0'$ is the initial perturbation, the initial condition for the linearized dynamics that the relevant $I$ realization has excited at $\tau = 0$, not to be confused with the initial condition for the nonlinear evolution $x_0 \equiv (x_y, y_y)$ at $t = 0$ (S2 Appendix). In Eq (7), $e^{J_x \tau}$ is the propagator [21–23, 26–28], which is expressed in the rightmost product in terms of the eigenvectors $J$ and $e^\lambda$ share (the columns of $E$) and growth factors (the nonzero diagonal elements of the diagonal matrix $\Lambda$, satisfying $\Delta = e^{\lambda \tau}$, where $\lambda_i$ is the ith eigenvalue of $J$). Thus the mapping of $x_0$ onto $x'(\tau)$ by the dynamics is done by a series of modes,

$$x'(\tau) = \mathbf{E} \Delta \mathbf{E}^{-1} x_0 = e_1[(e_{12}^x x_0^x) e^{\lambda_1 \tau}] + e_2[(e_{12}^x x_0^x) e^{\lambda_2 \tau}]$$

where we make explicit use of this system’s specific $N = 2$ (with the trivial generalization of the right hand series in general comprising $N$ terms), where $e_{1,2}$ are the two columns of $E$ (the eigenvectors of both $J$ and its matrix exponential), and where $e_{12}^x$ is the ith row of $E^{-1}$, so that the parenthetical term is a scalar product of the projection of the initial perturbation on the mode $e_x x_0$, and the modal growth $e^{\lambda \tau}$.
This representation is the origin of the focus on the eigenvalues as the deciding stability criterion, because as $\tau \to \infty$, the eigenvectors become unimportant for the evolution of the perturbation magnitude $|x'(\tau)|$. To see why, note that in this limit, the perturbation is essentially unimodal, because with enough time, the disparity in eigenvalues combined with the exponential evolution renders the amplitudes of all trailing modes (not corresponding to $\lambda_{\text{max}}$) negligibly small. Then,

$$
x'(\tau) \approx e_i (e_i^\tau x'_0 e^{i\tau})
$$

and thus $x'(\tau)$ always parallels $e_i$, with magnitude entirely determined by $e^{i\tau}$ (because the other contribution, $e_i^\tau x'_0$, is time-invariant). The $\tau \to \infty$ limit is thus implied despite being rarely explicitly acknowledged by the exclusive reliance on the eigenvalues to determine stability.

To analyze stability for often more pertinent finite times [15, 31], we again follow, e.g., [26–28] and decompose the propagator using the singular value decomposition, SVD [40, 41]

$$
x'(\tau) = e^{i\tau} x'_0 = UDV^H x'_0
$$

where $U$ and $V$ are $N \times N$ unitary matrices (for readers less familiar with the SVD, chapters 5 & 12 of [42] discuss all elements of the algebraic operation and its applications). The orthogonal unit norm columns of $U$ and $V$ (denoted individually $u_i$ and $v_i$ throughout the paper, unless otherwise explicitly stated, we use “norm” as a shorthand for the $L^2$ or Euclidean norm) are the sets of left and right singular vectors of the propagator, which are two in general distinct orthonormal bases for the algebraic space containing the state vector, here $\mathbb{R}^2$. (Note that some authors, e.g., [17], refer to these sets as the left and right eigenvectors, because they are the eigenvectors of $e^{i\tau} u^H$ and $e^{i\tau} v^H$, respectively. Yet because they are not the eigenvectors of the propagator itself, to clearly distinguish Eqs (7) and (10), here we favor singular.) The $(u_i, v_i)$ pair is the $i$th singular “mode” of the propagator, and the $N$ elements $D_{ii} = \sigma_i$ of the $N \times N$ diagonal matrix $D$ are the singular values of the propagator. While $H$ denotes Hermitian transpose, in most real physical and biological systems such as the grass–herbivore model, $x$, $F$ and $J$ are real, and thus so are all right hand side elements of Eq (10). With this, which we assume below, $H$ reduces to the transpose, i.e., $x'(\tau) = e^{i\tau} x'_0 = UD V^T x'_0$. For such propagators, plotting $\sigma_i(\tau)$ against all $rs$ within a window of interest identifies the global maximum nonnormal growth [26].

The key limitation of Eq (7) ameliorated by Eq (10) stems from the potential for a nonnormal or non-self-adjoint propagator. A square matrix $J$ (rectangular matrices are never self-adjoint) is self-adjoint if $JJ^H = J^H J$, and non-self-adjoint, or nonnormal otherwise. What is key to this distinction is that in the rather restrictive case of a self-adjoint $J$, the eigenvectors are orthonormal. In that case, Eq (7) can be rewritten as $x'(\tau) = E \Delta E^H x'_0$, where we use the fact that the Hermitian transpose of an orthonormal or unitary matrix (here $E$) is also its inverse, $EE^H = E^H E = I$. In the self-adjoint case, as $E \Delta E^H$ maps $x'_0$ onto $x'(\tau)$, the action of each eigenvector is independent, and the eigenvectors play no role in $\|x'\|$ growth, so $J$’s eigenvalues, especially $\Re(\lambda_{\text{max}}(J))$, tell the whole stability story. However, in the more general case of nonnormal Jacobians, this is only true for the $\tau \to \infty$ asymptotic, after trailing modes have already died down, but is emphatically not true for finite times. Before we analyze the $J$ of the grass–herbivore model, we provide a simple explanation.

**Algebra of nonnormal growth**

In this section, we provide simple insights into the algebraic origins of nonnormal growth that may arise when using Eq (7) to analyze the linearized stability of a nonlinear ecological
dynamical system. Yet note that this discussion is obviated by favoring Eq (10) over Eq (7). The only, but crucial, reason this discussion is necessary is the widespread reliance on eigenvalues of the Jacobian as the arbiter of ecological stability, e.g., [43, 44]. The deep roots and current ubiquity of this reliance in ecological discourse mean that ecologists are more likely to intuitively choose Eq (7) over Eq (10), yet ignore the roles E may play. It is this inconsistency that motivates this paper, and necessitates this section.

Algebraically speaking, nonnormal growth has two related sources, corresponding to the two appearances of E in Eq (7). The first and typically dominant is $E^{-1}x'_0$, (where for comparability, $x'_0$ is unit norm), the terms in the round parentheses in Eq (8) for $N = 2$. As the simple expansion of Eq (8) makes clear, the time-$t$ perturbation is a linear combination of E’s columns, with scalar coefficients (in square brackets) that combine multiplicatively eigenvalue based growth or decay $e^{st}$ with the projection of the initial perturbation on the modes. With it, we can better understand the problem with the exclusive focus on the eigenvalues for analyzing growth. This focus, or its accompanying disregard of eigenvector contributions to growth, implies the expectation that because E’s columns and $x'_0$ are unit norm, $\|E^{-1}x'_0\| = 1$. This is guaranteed if $J$ is normal, in which case any part of $x'_0$ that projects on a row of $E^{-1} = E^\dagger$ projects on no other row of $E^\dagger$. With the more general nonnormal $J$, $\|E^{-1}x'_0\| > 1$, because then the eigenvectors are not mutually orthogonal. The effect of this is best appreciated in the extreme case of $e_{1,2}$ being nearly mutually parallel, with a strongly acute angle between them. This means that while they formally span a plane, vectors not along the pair will require extremely large coefficients to express in terms of this defective basis. For example, consider $x'_0 = (-1, 1)^T/\sqrt{2}$ and

$$E = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 - 10^{-5} \\ 1 & 1 \end{pmatrix}.$$  \hspace{1cm} (11)

First, and parenthetically, note that the above example is not perfect because $e_2$ is not exactly unit $L^2$ norm, but instead satisfies $1 - \|e_2\| \approx 5 \times 10^{-2}$; this slightly imperfect $E$ nonetheless demonstrates the point simply and convincingly. Second, note that the above $x'_0$ is exactly and nearly orthogonal to the first and second columns of the above $E$ ($e_1$ and $e_2$, respectively), which we choose for powerfully encountering the amplification of initial perturbations by the mutually non-orthogonal eigenvectors of a nonnormal operator. Evaluating $E^{-1}x'_0$ (with $x'_0$ given just above Eq (11)), $x_0 \approx (-2 \times 10^6) e_1 + (2 \times 10^6) e_2$ (where the hat in the first term indicates that $e_1$ is exactly unit $L^2$-norm), with $\|E^{-1}x'_0\| \approx 2.8 \times 10^6$, almost 3-million-fold potential amplification of initial perturbations by eigenvector geometry.

The second source of growth occurring when using Eq (7) for nonnormal systems corresponds to the second appearance of $E$ in the equation. The focus of this, typically secondary, amplification source is the magnitude of $E\hat{g}$, the vector onto which $E$ maps any unit vector $\hat{g}$. That is, whatever growth or decay $g = \Delta E^{-1}x'_0$ embodies, there is still the potential for further growth of $\hat{g} \equiv g/\|g\|$ into $\|E\hat{g}\|$ resulting from the geometry of combining $E$’s columns by the elements of $E$. This amplification is most pronounced for perturbations orthogonal to ones optimally amplified by the first mechanism. Whereas for the first mechanism to dominate the exciting vector $x'_0$ has to be nearly orthogonal to the roughly mutually parallel eigenvectors, the second mechanism dominates when the coefficient vector $\hat{g}$ nearly parallels the defective basis. For example, using again the $E$ in Eq (11) but now with $g = (1 + 10^{-6}) (1)^T/\sqrt{2}$ (with the opposite trivial imperfection of $\|g\| - 1 = 5 \times 10^{-2}$), $\|E\hat{g}\| \approx \sqrt{2}$. Thus even when most pronounced, this second mechanism is indeed secondary, and is active
Fig 2. Demonstration of the two sources of amplification of initial perturbations beyond the eigenvalue based growth or decay. Here, we use the $2 \times 2$ matrix $A$ (Eq (12)) with unit norm columns (recall that unless otherwise explicitly stated, we take “norm” to mean the $L^2$ or Euclidean norm). The vertical axes show the largest singular value of the indicated matrix at the shown $\theta$s (i.e., the curves show the most possible amplification, assuming optimal excitation).

https://doi.org/10.1371/journal.pone.0178235.g002

for perturbations nearly parallel the defective basis, for which the first mechanism is less important.

We conclude the algebraic discussion of the two amplification mechanisms of initial unit perturbations by the linearized dynamical operator by quantifying their impact in the case of a more general, unspecified dynamical system whose eigenvectors are $a_{1,2}$ forming

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & (c + s) \\ 1 & (c - s) \end{pmatrix}$$

(12)

where $A$ is a special case of the general eigenvector matrix $E$, $c = \cos(\theta)$, $s = \sin(\theta)$, and thus $||a_1|| = ||a_2|| = 1$ and

$$a_1^T a_2 = \begin{cases} 0 & \text{for } \theta = 90^\circ \\ 1 & \text{for } \theta = 0^\circ \end{cases}$$

Fig 2a presents the upper bound on the first and typically (but not for the grass–herbivore model) dominant amplification mechanism, due to the action of $A^{-1}$ (with $A$ given by Eq (12)). By using $||A^{-1}|| = \sigma_{\text{max}}(A^{-1})$ as the definition of the matrix norm, the panel considers the optimal initial perturbation, and thus presents the maximum expected magnitude growth, realized when the exciting vector parallels $v_1$, the column of $V$ corresponding to the largest singular value in the SVD of $A$; see [23, 26, 27] for a fundamental discussion of this optimal perturbation. To limit the vertical extent of the plot, we only show $3^\circ \leq \theta \leq 45^\circ$, yet large amplifications are still apparent for small $\theta$s.

The second and typically secondary nonnormal amplification mechanism, due to linearly combining the eigenvectors, is presented in Fig 2b for the same $A$ (Eq (12)). While we show maximum possible amplification by either mechanism, note that in practice both optimals can not by simultaneously realized, as we show in S5 Appendix.

Results

With the parameter sets of S1 Table, the solutions of Eq (5) evolve as presented in Fig 3. To deemphasize the arbitrary initial state, we omit the first year. The system exhibits weakly
damped out-of-phase grass–herbivore oscillations as they evolve toward their respective fixed points (Eq (2) of S3 Appendix) plotted as horizontal lines in Fig 3.

Stability at the nontrivial fixed points

At the nontrivial fixed point, the Jacobian of the grass–herbivore model with parameter set 1 has a conjugate pair of eigenvalues, \( \lambda_{1,2}(J) \approx 10^{-8}(-10.8 \pm 74.3i) \). Because \( \Re(\lambda_{1,2}(J)) < 0 \) (which also holds for parameter sets 2 and 3, not shown), the eigenvalue-based stability criterion suggests that at its nontrivial fixed point, small perturbations decay at a rate of \( e^{|\Re(\lambda_{\text{max}})|\tau} \) (solid in Fig 4), ensuring that the system returns to the attractive fixed point. At \( \tau = 400 \) days, \( e^{|\Re(\lambda_{\text{max}})|\tau} \approx 0.6 \). To decay to one half its original magnitude, a perturbation requires \( \tau \approx \log(0.5)/(-10.8 \times 10^{-8}) \approx 643 \) days, or over 21 months (this \( \tau \) is not shown in Fig 4, which only shows \( 0 \leq \tau \leq 400 \) days). To decay to 10% the original magnitude, \( \|x'(\tau)\| = 0.1 \|x_0\| \), requires \( \tau \approx 71 \) months (with similar results for parameter sets 2 and 3). This demonstrates a key limitation of the eigenvalue based \( \tau \rightarrow \infty \) analysis; it is unrealistic to expect an ecological system to be nonlinearly frozen in place for years, allowing perturbations to stably decay.

Instead, with more realistic inclusion of geometrical considerations permitted by using \( \sigma_{\text{max}} = D_{11} \) based on Eq (10) as the growth criterion (dashed in Fig 4), perturbations not only fail to decay, but in fact double in magnitude at \( \tau \approx 180 \) days, varying slightly among parameter sets yet telling the same story. Thus in this rather typical situation, while incomplete application of Eq (7) suggests slow decay of small perturbations about the fixed point, some perturbations may in fact grow for quite some time, because of the geometry of the eigenvectors, before succumbing to the inexorable eigenvalue based modal decay.
By searching over a range of $r$s within a window of interest, we find that the initial perturbation that maximizes finite time growth is the leading right singular vector of $\mathbf{J}^{184}$, where 184 is the time of maximum linearized growth in Fig 4b, $v_1^{\text{day}184} \approx (10^{-4} \, 1)^T$. That is, the most linearized growth on $\tau \leq 200$ day timescales is excited by perturbations nearly entirely in herbivore mass, with virtually no changes in grass, which can be easily realized with, e.g., intensified hunting or herbivore epidemic. Thus at its nontrivial fixed point, the grass–herbivore model is not stable and decaying, as the eigenvalues suggest, but in fact sustains large, prolonged perturbation growth.

**Stability at ecologically critical points**

As Fig 3 shows, over $t = [18, 26]$ months grass cover and herbivore population decline perceptibly. Because the growth terms in Eq (5) both vanish for $x = y = 0$, system collapse is of particular concern then. It is easy to imagine that if at that vulnerable point a fire or a rainsless stretch further undermines the system, collapse may follow independently of the existence of an attractive fixed point. It therefore makes sense to analyze the system stability in such times as well. We thus next examine the linear stability of Eq (5) at $t = 800$ days (when herbivore population is minimized), using Eqs (7) and (10).

Using parameter set 1, $\mathbf{x}_{800} \approx (271 \, 686)^T$, yielding

$$\mathbf{J} \approx 10^{-4} \begin{pmatrix} -13 & -180 \\ 31 & 0 \end{pmatrix}$$

(Eq (1) of S1 Appendix) and $\lambda_{1,2}(J) \approx 10^{-4}(-6.6 \pm 73.8i)$, i.e., $\mathbb{R}[^{}\lambda_{\text{max}}(J)] < 0$. That is, as the system evolves under its governing nonlinear dynamics, it fluctuates widely, at times attaining such states as $\mathbf{x}_{800}$ where—because of dangerously low population(s)—ecosystem collapse by stochastic anomalies with linearly growing amplitudes becomes possible. Should the asymptotic decay indicated by the above $\mathbb{R}(\lambda_{\text{max}}) < 0$ lay this concern to rest? Several observations suggest that this alone does not settle the stability issue.

First, as was true above for the fixed point, at $t = 800$ days the decay timescale—with perturbation half life of $\tau = \log(0.5)/(-6.6 \times 10^{-4}) \approx 3$ years—is also too long for reassurance; too much is changing in most ecosystems over 3 years to expect the linearization to remain valid.
Fig 5. Evolution of initial perturbations $x_0$ about the near catastrophe day 800 state of the simple grass–herbivore model with parameter set 1. As in Fig 4, solid/squares present $e^{Wt_{	ext{max}}}$, the growth assuming $t \to \infty$. Dashed/triangles present the growth measure appropriate over finite times, $\max \|\sigma_t \exp(1J^t)\|$, the largest singular value of the propagator. The maximum growth factor and the lead time in days at which it is realized are indicated. Because of the near neutrality of the leading eigenvalue, the thin horizontal line separating the growth and decay regions is barely discernible.

https://doi.org/10.1371/journal.pone.0178235.g005

Just as important is the nonnormal growth that may arise during that time. This is quantified by favoring Eq (10) over Eq (7), which for the $N=2$ grass–herbivore model is simply

$$\dot{x}(t) = [\sigma_1 v_i^* x_0]u_1 + [\sigma_2 v_i^* x_0]u_2$$

(14)

where now the SVD is of the propagator, so that $u_{1,2}$ and $v_{1,2}$ are the left and right singular vectors of $e^T$. Because $\sigma_1 > \sigma_2$, the most growth occurs when $x_0$ projects most on $v_1$. Taking this to the limit, the optimal $x_0$ is $v_1$ itself, with which

$$\dot{x}(t) = [\sigma_1 v_1^* v_1]u_1 = \sigma_1 u_1$$

(15)

As Fig 5 shows, these optimal initial perturbations grow to 213% of their initial unit norm by day 196, likely a more ecologically relevant timescale than the 3 years calculated above for asymptotic decay. At the precarious $t = 800$ day state, Eq (10) with $\tau = 196$ days yields $v_1 \approx (0 1)^T$, like at the fixed point. This means that at the point in which the nonlinear dynamics minimize herbivore population, optimal seasonal scale linearized growth occurs when the initial perturbation comprises only perturbed herbivore biomass, with essentially intact grass cover. And after 196 days, this $v_1$ becomes $\sigma_1 u_1 \approx 2.1(0 1 0.1)^T$ by the linearized dynamics.

Put simply: if a grass–herbivore ecosystem is at a nonlinear state like $x_{\text{agg}}$, and some external perturbation then forces the grazer community away from the 686 kg ha$^{-1}$ expected based on the nonlinear dynamics, then for every kg herbivore biomass ha$^{-1}$ added, 196 days later the grass cover will be 2.1 kg ha$^{-1}$ lower than what is expected based on the nonlinear dynamics, with the initial herbivore surplus reduced by 90%. That is, the more excess herbivore now, the less grass later, with the perturbation amplitude approximately doubling over roughly 200 days due to energy "leakage" during trophic cascading. This basic logic is straightforward and compelling, yet the leading eigenvalue leads to expecting perturbation decline (Fig 5, solid).
A reasonable concern about the above calculation is the validity of the linearization over a timescale that is a nontrivial fraction of the $\approx 2.5$ y period of the oscillations the nonlinear trajectory exhibits (Fig 3). That is, because the Jacobian is a strong function of the state, it is possible that the linearization is used over timespans exceeding its expected utility. We test this possibility in S5 Appendix, and find the linearization adequate.

Pursuing the ultimate goal of modeling, mechanistic understanding of system dynamics [31], we now highlight an important role the eigenvalues do play despite their limitations as a stability criterion that is the main thrust of this paper. The dash-dotted curve in Fig 6 shows that throughout $t \in [1, 5]$ years (and beyond, not shown), $1.6 \leq ||E^{-1} v_1|| \leq 2.1$. That is, all nonlinear states considered here are capable of ample linearized nonnormal growth by the first mechanism, projection of initial perturbations on $E^{-1}$. Yet any linearized nonnormal growth mechanism can only amplify initial perturbations to the extent permitted by the potentially unyielding eigenvalue-based modal decay. That is, even a hefty $||E^{-1} v_1|| = 10^6$ will result in overall damping in the face of, say, $e^{\lambda_{\text{max}} t} = 10^{-12}$. Indeed, despite the afore-mentioned ample growth by $E^{-1} v_1$ (dash-dotted), this potential amplification is damped by modal decay to a variable degree, so that whereas $\max(||E^{-1} v_1||) \approx 2.1$ (at $t = 992$ d), $\max(||\Delta E^{-1} v_1||) \approx 1.7$ (at $t = 708$ d; dashed). Yet further premultiplying $\Delta E^{-1} v_1$ by $E$ (the second nonnormal growth mechanism; the solid curve) more than offsets the modal loss, increasing the maximum overall amplification to 2.2 (at $t = 720$ d). This highlights the importance of the balance between nonnormal amplification on the one hand, and modal damping on the other.

Thus the stability of the modeled grass–herbivore ecosystem is qualitatively similar at the system’s nontrivial fixed point and when the ecosystem is most vulnerable due to unusually low populations. In these situations, the ecosystem is not only stable and decaying, as the eigenvalues suggest, but may also exhibit large, prolonged perturbation growth. And this difference may well translate to the difference between smooth recovery and system collapse, as follows.

Fig 6. The various contributions to the overall growth by the propagator. We use parameter set 1, $t \in [1, 5]$ years, and $\tau = 200$ days of linearized growth throughout. For all curves, $v_1$ is the same, and is the unit norm perturbation that optimally excites the full propagator $e^{\Sigma_{200}}$. The solid curve shows the full impact of the propagator on the magnitude of the initial perturbation, $||E\Delta E^{-1} v_1||$. The contributions to this impact by the first discussed mechanism of nonnormal growth, $||\Delta E^{-1} v_1||$, and the modifications of this by the eigenvalues, $||\Delta E^{-1} v_1||$, are given by the dash-dotted and dashed curves, respectively. The vertical solid and dashed lines show the time of grass and herbivore global minima ($t = 584$ and 800 d, respectively).

https://doi.org/10.1371/journal.pone.0178235.g006
Let us define “collapse” as zero grass, \( \bar{x} + x' = 0 \), and derive the initial perturbation \( \mathbf{x}_0' \) required to arise at \( t = 800 \) d to guarantee such collapse after the optimal \( \tau = 196 \) days. Starting from Eq (10), \( \mathbf{x}_0' = \mathbf{UDV}^\top \mathbf{v}_1 \). We also make use of the fact that \( \mathbf{x}_0' = w \mathbf{v}_1 \), i.e., that the optimal initial perturbation parallels the leading right singular vector \( \mathbf{v}_1 \) with some weight \( w \) to be determined shortly. Therefore,

\[
\mathbf{x}_0' = w \mathbf{UDV}^\top \mathbf{v}_1 = w \sigma_1 \mathbf{u}_1
\]

but because we only strive to bring grass to zero, we focus exclusively on the first row,

\[
-\bar{x}_{196} \approx 271 = w \sigma_1 \mathbf{u}_{1,1} \quad \Rightarrow \quad w = -\bar{x}_{196}/(\sigma_1 \mathbf{u}_{1,1}) \approx 128
\]

where \( \mathbf{u}_{1,1} \) is the first element of \( \mathbf{u}_1 \). With this, \( \mathbf{x}_0' \approx (0 \quad 128)^\top \), yielding \( \mathbf{UDV}^\top \mathbf{x}_0' + \bar{x}_{800} = (0 \quad 709)^\top \), i.e. zero grass, by construction, and a slightly enlarged herbivore population (recall that \( \bar{y}_{800} \approx 686 \)). With no grass for subsistence and more mouths to feed, widespread herbivores starvations or desperate exodus follows, with complete collapse of the modeled ecosystem shortly thereafter. The important detail here is that at that precarious \( t = 800 \) d state, all that is required for complete ecosystem collapse is 128 added kg herbivore ha\(^{-1}\), hardly a fanciful eventuality. Thus whereas the eigenvalue based stability criterion promises safe ecosystem perpetuation, a modest (128/686 < 19%) herbivore addition—by, e.g., inward migration or hunting of herbivore consumers—can readily usher in ecosystem collapse.

How unique are these states in which large potential perturbation growth can occur when modal stability is firmly in place? This question is answered by Fig 7, showing modal damping (a) and nonnormal growth (b) as functions of \( t \) (horizontal axes). Clearly, throughout the

Fig 7. Modal (a) and nonnormal (b) growth throughout \( t \in [1, 10] \) y. Because the system is at times near neutral, panel a presents \( 10^2\mathbf{R}(\lambda_{\text{max}}) \), with least damping for \( \tau = 200 \) d indicated. Panel b shows nonnormal growth (right colorbar) as a function of both \( t \) (horizontal) and \( \tau \) (vertical).

https://doi.org/10.1371/journal.pone.0178235.g007
shown range, the system is modally stable, yet can support magnitude doubling because of nonnormality for $150 \leq \tau \leq 250$ days. The answer, that is, is “not unique at all; the rule rather than the exception”.

Discussion and conclusions

The primary message of this paper is that while eigenvalues are important and revealing, they by no means tell the full stability story. Yet the least damped eigenvalue is the stability criterion most likely to be used in ecology [45, 46] (and elsewhere, see applications of this critique to, e.g., astrophysical magnetohydrodynamics [47, 48], chemistry [49], or the geodynamo [50]). Far from merely an algebraic curiosity, this can determine the fate of an ecosystem in the opposite direction to what the leading eigenvalue suggests. Take, e.g., the $t = 800$ state of critically low populations in the grass–herbivore model used in this paper. As the solid-squares curve of Fig 5 makes clear, the lead eigenvalue is damped, predicting stability or gradual decay of any initial perturbation at that tricky point. Yet ecosystem collapse is possible, as demonstrated above.

Potentially suggesting safe stability where ecosystem collapse is distinctly possible, stability determination based exclusively on the least damped eigenvalue does not serve ecosystem science well. By simply favoring the inclusive, more complete Eq (10) over the customary yet restrictive partial application of Eq (7), this potential pitfall can be easily avoided. That these two stability analyses may yield diametrically opposite results may have broad implications to both fundamental ecosystem science and to applied conservation science.

Supporting information

S1 Fig. The assumed functions.
(TEX)

S2 Fig. Comparison of the nonlinear (top) and linearized (bottom) flow throughout the relevant state range.
(TEX)

S1 Table. The three parameter sets we use for the grass–herbivore model.
(TEX)

S1 Appendix. Governing equations.
(TEX)

S2 Appendix. Nonlinear solution and initial state.
(TEX)

S3 Appendix. Fixed points.
(TEX)

S4 Appendix. Linearization.
(TEX)

S5 Appendix. On the non-simultaneity of optimally exciting $E$ and $E^{-1}$.
(TEX)

S6 Appendix. Linear vs. Nonlinear flow throughout the state space.
(TEX)
Author Contributions
Conceptualization: GE.
Data curation: GE YC.
Formal analysis: GE.
Investigation: GE YC.
Methodology: GE.
Project administration: GE.
Resources: GE.
Software: GE.
Supervision: GE.
Validation: GE.
Visualization: GE.
Writing – original draft: GE.
Writing – review & editing: GE YC.

References

PLOS ONE | https://doi.org/10.1371/journal.pone.0178235  June 7, 2017  15 / 17


