Eigenvalue Attraction

Ramis Movassagh[1,2,*]

[1]*Department of Mathematics, Northeastern University, Boston Massachusetts, 02115*

[2]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge Massachusetts, 02139*

(Dated: March 23, 2015)

We prove that the complex conjugate (c.c.) eigenvalues of a smoothly varying real matrix attract. We derive an equation for this force (Eq. 15) and discuss applications. We then discuss random perturbations of a fixed matrix M. If M is normal, we find that the total expected force on any eigenvalue is only from its c.c. (Eq. 23) and when M is circulant the strength of interaction can be related to the power spectrum of white noise. Next we derive the eigenvalue dynamics of stochastic processes by constructing a smooth family of stochastic processes and explicitly write down the differential equations governing the motion of any eigenvalue. We prove that

$$\mathbb{E}\left(\text{Force of } \bar{\lambda}_i \text{ on } \lambda_i\right) \propto -i\frac{\|\mathbf{u_i}^*\|_2^2}{\text{Im}\left(\lambda_i\right)},$$

where λ_i and $\bar{\lambda}_i$ are c.c. eigenvalues, and $\mathbf{u_i}^*$ is the corresponding left eigenvector. Therefore, c.c. pairs closest to the real axis, or those that are ill-conditioned, attract most strongly and can collide to become exactly real. We quantify the dominance of the c.c. attraction by calculating the variance of other forces. We apply the results to the Hatano-Nelson model and provide other numerical illustrations. Lastly, we discuss the aggregation and low density of the eigenvalues of real random matrices on and near the real axis respectively.

Contents

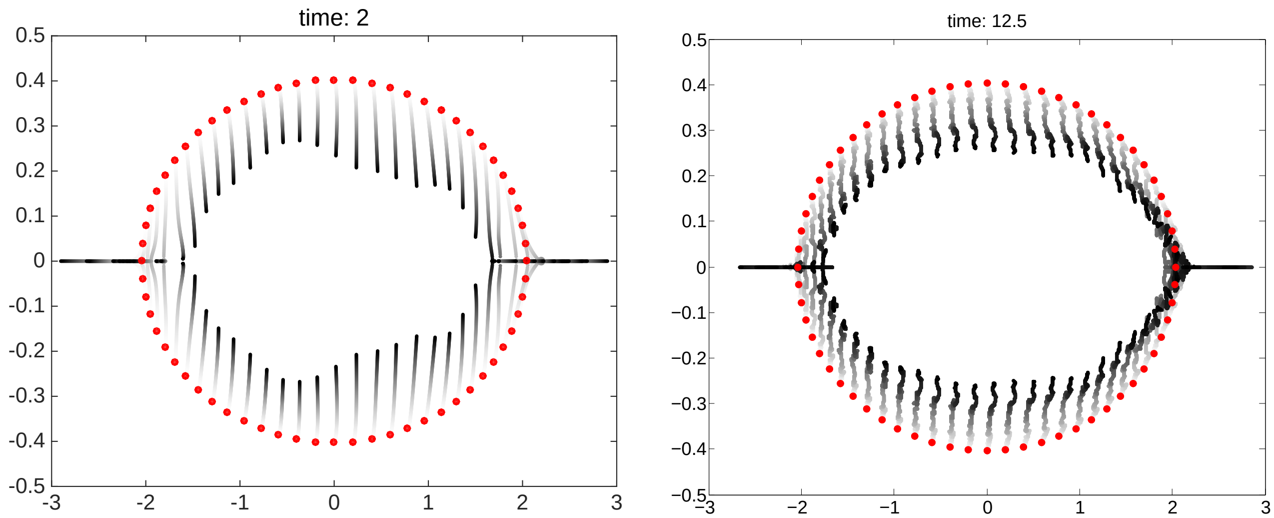


Figure 1: Hatano-Nelson model as described in the text with $g = 0.2$. Left (Demo1): Small perturbation. Right (Demo 2): Stochastic dynamics. In Demo 2 the final time is much larger because of the random impulses imparted at times $0.25i$.

I. BACKGROUND, ILLUSTRATION AND SUMMARY OF MAIN RESULTS

Much work has been devoted to the understanding of the motion of eigenvalues in the presence of randomness. The folklore of random matrix analysis, especially in the case of normal matrices (e.g., Hermitian), suggests that the eigenvalues of a perturbed matrix repel. This has been pointed out previously by various authors [1, 2] and is well known in quantum physics [3, p. 304-305]. More recently, in agreement with the universality conjectures, the level repulsion was proved for the eigenvalues of a Wigner matrix [4].

The stochastic dynamics of the eigenvalues of Hermitian matrices have been vigorously studied in the past [5, recommended]. Most celebrated is Dyson's Brownian motion, which proves that the eigenvalues of a Hermitian matrix undergoing a Wiener process perform a Brownian motion [6].

In physics, one mainly studies Hermitian matrices and operators as their eigenvalues correspond to observable quantities, which clearly need to be real. However, in recent years, non-Hermitian models have gained much attention in the context of pinning of vortices in type II superconductors initiated by Hatano and Nelson [7] and followed up in works on the nature of localized states and eigenvalue distributions [8–11]. Non-hermitian models also come up in fluid mechanics [1, Ref. therein], transport phenomena in photonics [12] and biophysical phenomena [13].

In the Hatano-Nelson model, the eigenvalue distribution gives rise to "wings" of real eigenvalues when the perturbation is sufficiently strong (see for example [1, Section 36] and citations therein as well as Figs. 1 and 6). The wings result from the motion of complex eigenvalues that move in response to the perturbation and ultimately sit on the real axis. Goldsheid et al derived an equation for the shape of "the winged" spectrum [14]. According to [7], these eigenvalues correspond to localized eigenstates.

In investigating the (de)localization of the eigenstates, Feinberg and Zee [10], argued that imaginary eigenvalues near the real axis can attract when perturbed by a Hermitian matrix by providing a 2×2 example of an imaginary diagonal matrix perturbed by a 2×2 Hermitian matrix with zero diagonal entries. Later, Bloch et al [15] considered antisymmetric perturbations of real symmetric matrices in the context of two-color quantum chromodynamics and provided examples that a Hermitian matrix perturbed by a real antisymmetric perturbation can give rise to attraction of eigenvalues. To our knowledge, attraction of the eigenvalues and their eventual aggregation on the real line, in a general setting, has not been proved. Here we prove that *complex conjugate* (c.c.) eigenvalues of a smoothly varying real matrix $M(t)$ attract (i.e., pull on each other). We emphasize that randomness is *not* needed. Before summarizing the main results let us provide an illustration. In Section III G, we provide various other illustrations.

A. An Illustration

Remark 1. **Explanation of the figures:** All the plots were done in MatLab. We take the vertical (horizontal) axis to be the imaginary (real) axis and plot the eigenvalues of $M(t)$. The red dots are the eigenvalues of $M(0)$. To show the dynamics of the eigenvalues as a function of t, we plot the eigenvalues of $M(t)$ in the complex plane in gray scale, where at $t=0$ they are shown in white (coincide with the red dots) and darken as t increases till their final position at $t=t_{max}$ shown in black. The eigenvalues of $M(t)$ at any $0 \le t \le t_{max}$ have the same gray scale color. In MatLab language we use "`hold on`;" to show the eigenvalues for all t.

Demo1: In (Fig. 1, left) we show the spectral dynamics of the Hatano-Nelson model $M(t) = H + \delta t\, P$, where $M(0) \equiv H$ is

$$H = \begin{bmatrix} 0 & e^{g} & & & e^{-g} \\ e^{-g} & 0 & e^{g} & & \\ & e^{-g} & 0 & \ddots & \\ & & \ddots & \ddots & e^{g} \\ e^{g} & & & e^{-g} & 0 \end{bmatrix}, \tag{1}$$

where g governs the non-Hermiticity and P is a real diagonal matrix of random Gaussians. In the figure we took $g=0.2$, which makes $\|H\|_2 = 2.04$. The boundary terms, i.e., $(1,n)$ and $(n,1)$ entries, are crucial for the spectral properties [1, Sec. 36].

Comment: When $g \to 0$, this model coincides with the Anderson model of localization with periodic boundary conditions [16]. However, the latter is a Hermitian model whose properties are quite different from $g \neq 0$ Hatano-Nelson model.

Demo2: In (Fig. 1, right) We took $g = 0.2$ as before and take the time discretization $t_i = 0.25i$ for $i = 0, 1, \cdots, 50$, when at each t_i an independent matrix $P(t_i)$ is introduced according to $P(t_i) = \text{diag}(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)$ where ϵ_j's are drawn independently from a standard normal distribution. Each $P(t_i)$ is normalized to have a unit $2-$norm. A piece-wise linear discrete stochastic process is therefore constructed. We plot the eigenvalues of $M(0.25i + \delta t) = M(0.25i) + \delta t\, P(t_i)$ with $0 \le \delta t \le 0.25$ and time steps $\Delta t = 0.01$. In (Fig. 1, right), we show the evolution of $M(t)$ for the total time $t \in [0, 12.5]$. Note that the eigenvalues move towards the real line as before but make less progress because of the stochastic kicks at times $0.25i$. Below we show that there is an attraction between c.c. pairs that largely governs this behavior of the spectral dynamics in this case. See Section IV for further theoretical discussion.

Comment: The procedure just described does not provide a smooth stochastic process as it is not differentiable at times $0.25i$ (a set of measure zero). This issue persists for piece-wise linear discretization where in each interval a new random $P(t_i)$ is used. In fact a continuous Brownian motion provides an example of a continuous map that, with probability one, is nowhere differentiable [5].

B. Summary of the main results

In this paper we consider the general problem of the interaction of the eigenvalues of a matrix, $M(t)$, whose entries vary with respect to a real parameter, which we think of as time. Although, eigenvalue repulsion is quite a general feature of Hermitian matrices, attraction of eigenvalues is rarely considered. Below we will first define Eigenvalue Attraction (see Def. 1) and give the general equations of the motion of any one of the eigenvalues, which depends on the other eigenvalues and the inertia of the matrix itself. We introduce a new notation more akin to formulations of interacting many-body systems. We will prove that when $M(t) \in \mathbb{R}^{n \times n}$, any non-real eigenvalue attracts its c.c. by a force equal to (see Lemma 1)

$$\left(\text{Force of } \bar{\lambda}_i \text{ on } \lambda_i\right) = -i\, \frac{\left|\mathbf{u}_i^T \dot{M}(t)\, \mathbf{v}_i\right|^2}{\operatorname{Im}(\lambda_i(t))}$$

where $\mathbf{v}_i$ and $\mathbf{u}_i$ are the right (i.e., standard) and left eigenvectors corresponding to the eigenvalue λ_i and complex conjugation is dented by an over bar.

We will apply this to various $M(t)$ such as Hatano-Nelson model or by a convex combination of two deterministic matrices. We then introduce randomness to obtain probabilistic statements. In particular, we consider the line of matrices $M(t) \equiv M + \delta t\, P$, where M is a fixed matrix and P is a random real matrix whose entries have zero mean, bounded moments and are independently and identically distributed (iid). This special case comes up often in applications. In this limit and in order to quantify the dominance of the c.c. attraction, we calculate the expectation value and variance of all other forces excluding the c.c.. We prove that when M is a normal matrix (i.e., unitary diagonalizable), the total expected force on any eigenvalue is only due to its c.c., and when M is circulant and P is diagonal, the force only depends on the eigenvalues and that the strength of interaction is the power spectrum of the diagonal entries of P, which can be a constant independent of the eigenpairs (e.g., the case of white noise). We will prove other results applicable to general circulant matrices and apply them to the Hatano-Nelson model to analyze its spectral dynamics.

The results are extended to $M(t)$'s that are defined by real stochastic processes with independent increments (Eq. 31). The expected force of attraction is always

$$\mathbb{E}\left(\text{Force of } \bar{\lambda}_i \text{ on } \lambda_i\right) = -i\frac{\sum_{m,\ell}\mathbb{E}\left[p_{m\ell}^2\right]\left|u_i^{*,m}\right|^2\left|v_i^{\ell}\right|^2}{2\,\mathrm{Im}\left(\lambda_i\right)} \tag{2}$$

$$\underset{=}{\text{iid}} -i\frac{\mathbb{E}\left[p^2\right]\left\|\mathbf{u_i}^*\right\|_2^2}{2\,\mathrm{Im}\left(\lambda_i\right)} \tag{3}$$

where $\bar{\lambda}_i$ is an eigenvalue that is c.c. to λ_i, $\mathbf{v_i}$ and $\mathbf{u_i^*}$ are the corresponding right and left eigenvectors respectively and the second equality assumes $\mathbb{E}\left[p_{m,\ell}^2\right]$ is the same for all m and ℓ. Clearly, the attraction is strongest near the real line. Since the proportionality constant depends on the $2-$norm of the left eigenvector, the force of attraction can be quite strong for ill-conditioned eigenvalues. When this attractive force is dominant over the force exerted by the rest of the eigenvalues, the c.c. pair approach one another and eventually collide and "scatter" near and ultimately reside at different points on the real line. At this point, the well-known repulsion mechanism takes over and the reality of the matrix ensures that each eigenvalue remains real.

The motion on the real line is not permanent. In most cases, an eigenvalue that moves about on the real line gets close enough to (i.e., collide with) another eigenvalue on the real line, after which they shoot off into the complex plane. By the reality of the matrix, they are forced to form a new c.c. pair. The alternative would be that they would repel and remain on the real line. However, in the majority of cases we investigated, it seems 'energetically' more favorable for them to form a new c.c. pair perhaps because there are more degrees of freedom available away from the real line. In this paper, we will not rigorously investigate this to any depth.

As mentioned above any stochastic process (e.g., Wiener process, Brownian motion) is non-smooth, despite often being continuous. The appearance of a new $P(t_i)$ makes the limits of the derivative from left and right unequal $\dot{M}\left(t_i^>\right) \neq \dot{M}\left(t_i^<\right)$, yet there are powerful tools of matrix calculus that can be utilized if $M(t)$ were differentiable. Moreover, from the applied perspective, nothing is instantaneous.

In the Section IV, we give the basic definitions of discrete stochastic processes and introduce a smoothing construction, that can be used to smoothen any discrete stochastic process with a control over the rapidity of (dis)appearance of every $P(t_i)$ within $[t_i, t_{i+1}]$ (see Eq. 33). The original (non-differentiable) stochastic process is

$$M(t) = \lim_{\epsilon\to 0} M_\epsilon(t) \quad .$$

We will conclude by discussing an open problem pertaining to the sparsity of eigenvalues near the real line for random real matrices and list further open problems.

C. Definition of Attraction

In this paper we take the point of view that the eigenvalues are interacting identical particles whose motions take place in the complex plane and our goal is to better understand their dynamics. The eigenvalues of $M(t)$ are also functions of time and the i^{th} eigenvalue is denoted by $\lambda_i(t)$. Generally the eigenvalues of a continuously varying $M(t)$ are also continuous in t. That is their motion follows a connected path in the complex plane. This follows from the fact that eigenvalues are roots of a characteristic polynomial, which itself is continuous, and a theorem due to Rouché [17, Chapter 4].

Definition 1. (Attraction and Repulsion) At any t we say $\lambda_j(t)$ attracts (repels) $\lambda_i(t)$ if the dependence of $\ddot{\lambda}_i(t)$ on $\lambda_j(t)$ is proportional to $\lambda_i(t) - \lambda_j(t)$ with a negative (positive) constant of proportionality.

Comment: The acceleration, $\ddot{\lambda}_i(t)$, can be thought of as a "force" if we take the eigenvalues to have unit mass. Below we shall use the word force as it provides better intuition.

II. GENERAL DYNAMICS OF EIGENVALUES

Here we follow a derivation similar to that given by T. Tao [2] to obtain the governing dynamical equations for the eigenpairs of a general smoothly varying $M(t)$. We assume that the eigenvalues are simple. The eigenvalue equations are

$$\begin{aligned} M(t)\ \mathbf{v_i}(t) &= \lambda_i(t)\,\mathbf{v_i}(t) \qquad (4)\\ \mathbf{u_i}^*(t)\ M(t) &= \lambda_i(t)\,\mathbf{u_i}^*(t)\,, \qquad (5)\end{aligned}$$

where $\lambda_i(t)$ are the eigenvalues, $\mathbf{v_i}(t)$ the (right) eigenvectors, which we take to be normalized, and $\mathbf{u_i^*}(t)$ are the left eigenvectors dual to $\mathbf{v_i}(t)$. If we consider the matrix of eigenvectors $\mathbf{V}(t) = [\mathbf{v_1}(t)\ \mathbf{v_2}(t), \ldots, \mathbf{v_n}(t)]$, then $\mathbf{u_j}^*(t)$ is the j^{th} row of $\mathbf{V}^{-1}(t)$ and

$$\mathbf{u_j}^*(t)\ \mathbf{v_i}(t) = \delta_{ij} \quad . \qquad (6)$$

Since $\mathbf{v}_1(t), \ldots, \mathbf{v}_n(t)$ form a basis for $\mathbb{C}^n$, any vector $\mathbf{x}$ has the expansion $\mathbf{x} = \sum_{j=1}^{n}\left[\mathbf{u_j}^*(t)\ \mathbf{x}\right]\mathbf{v_j}(t)$. Differentiating Eqs. 4 and 5 with respect to t, gives

$$\begin{aligned} \dot{M}(t)\ \mathbf{v_i}(t) + M(t)\,\dot{\mathbf{v}}_\mathbf{i}(t) &= \dot{\lambda}_i(t)\,\mathbf{v_i}(t) + \lambda_i(t)\,\dot{\mathbf{v}}_\mathbf{i}(t)\\ \mathbf{u_i}^*(t)\ \dot{M}(t) + \dot{\mathbf{u}}_\mathbf{i}^*(t)\ M(t) &= \dot{\lambda}_i(t)\,\mathbf{u_i^*}(t) + \lambda_i(t)\,\dot{\mathbf{u}}_\mathbf{i}^*(t)\end{aligned}$$

Multiplying the first equation on the left by $\mathbf{u_i}^*(t)$ we obtain the "velocity" of $\lambda(t)$ in the complex plane [1]

$$\dot{\lambda}_i(t) = \mathbf{u_i}^*(t)\ \dot{M}(t)\ \mathbf{v_i}(t)\,. \qquad (7)$$

In order to compute the acceleration on any eigenvalue we shall need the derivatives of the left and right eigen*vectors*. They are simple to compute [2],

$$\dot{\mathbf{v}}_\mathbf{i}(t) = \sum_{j\neq i}\frac{\mathbf{u_j^*}(t)\ \dot{M}(t)\ \mathbf{v_i}(t)}{\lambda_i(t)-\lambda_j(t)}\mathbf{v_j}(t) + \eta_i(t)\,\mathbf{v_i}(t) \qquad (8)$$

$$\dot{\mathbf{u}}_\mathbf{i}^*(t) = \sum_{j\neq i}\frac{\mathbf{u_i^*}(t)\ \dot{M}(t)\ \mathbf{v_j}(t)}{\lambda_i(t)-\lambda_j(t)}\mathbf{u_j^*}(t) - \eta_i(t)\,\mathbf{u_i^*}(t) \qquad (9)$$

where $\eta_i(t)$ is a scalar function because a constant multiple of an eigenvector is also an eigenvector. The second derivative, or acceleration, of the eigenvalue $\lambda_i(t)$ is obtained by differentiating Eq. 7 one more time,

$$\ddot{\lambda}_i(t) = \dot{\mathbf{u}}_\mathbf{i}^*(t)\ \dot{M}(t)\ \mathbf{v_i}(t) + \mathbf{u_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) + \mathbf{u_i}^*(t)\ \dot{M}(t)\ \dot{\mathbf{v}}_\mathbf{i}(t)\,.$$

Using Eqs. 8 and 9, the second derivative becomes [2]

$$\begin{aligned} \ddot{\lambda}_i(t) &= \mathbf{u_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) + 2\sum_{j\neq i}\frac{\left[\mathbf{u_i^*}(t)\ \dot{M}(t)\ \mathbf{v_j}(t)\right]\left[\mathbf{u_j^*}(t)\ \dot{M}(t)\ \mathbf{v_i}(t)\right]}{\lambda_i(t)-\lambda_j(t)} \qquad (10)\\ &= \mathbf{v_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) + 2\sum_{j\neq i}\frac{\left|\mathbf{v_i^*}(t)\ \dot{M}(t)\ \mathbf{v_j}(t)\right|^2}{\lambda_i(t)-\lambda_j(t)} \quad \text{Normal matrices} \qquad (11)\end{aligned}$$

---

[1] We remark that the theory of pseudo-spectra [1] quantifies how far an eigenvalue can wander without quantifying the direction of the motion

where for normal matrices one has $\mathbf{u}_i = \mathbf{v}_i$. As pointed out by Tao, the first term can be seen as the inertial force of the matrix and the second the force of interaction of the eigenvalues. The origin of instantaneous repulsive force between eigenvalues of normal matrices (e.g., Hermitian) is easily seen by the second term in Eq. 11 [2]. For example, if $\lambda_i(t) > \lambda_j(t)$, then the force is positive, the effect of $\lambda_j(t)$ in the sum is to push $\lambda_i(t)$ to the right. Similarly, if $\lambda_i(t) < \lambda_j(t)$, the effect of $\lambda_j(t)$ is to exert a negative force on $\lambda_i(t)$. Moreover the strength of the force is inversely proportional to their distance ($1/\left|\lambda_i(t) - \lambda_j(t)\right|$) which is clearly strongest when the eigenvalues are closest [2]. The repulsion is at work for self-adjoint matrices for all t.

Comment: Eqs. (7-11) are essentially the standard first and second order perturbation theory results. See for example, Dirac [18, Section 43] [2], Wilkinson's wonderful exposition [19] and more recently [20].

A. Eigenvalue Attraction

Below we denote the c.c. of an eigenvalue or entry-wise complex conjugation of an eigenvector with an over-bar. To better visualize the kinematics of the eigenvalues in the complex plane, we write

$$\begin{aligned}
\frac{1}{\lambda_i(t) - \lambda_j(t)} &= \frac{\overline{\lambda_i(t)} - \overline{\lambda_j(t)}}{\left|\lambda_i(t) - \lambda_j(t)\right|^2} \equiv \frac{\hat{\mathbf{r}}_{ij}}{\left|\mathbf{r}_{ij}\right|}; \\
\mathbf{r}_{ij} &\equiv \overline{\lambda_i(t)} - \overline{\lambda_j(t)} \\
\hat{\mathbf{r}}_{ij} &\equiv \frac{\overline{\lambda_i(t)} - \overline{\lambda_j(t)}}{\left|\lambda_i(t) - \lambda_j(t)\right|}
\end{aligned} \tag{12}$$

where $\mathbf{r}_{ij}$ is a vector in the complex plane stretching from $\overline{\lambda_j(t)}$ to $\overline{\lambda_i(t)}$ and $\hat{\mathbf{r}}_{ij}$ is the corresponding unit vector. One could further simplify the notation by denoting the complex number

$$c_{ij} \equiv \mathbf{u}_\mathbf{i}^*(t)\ \dot{M}(t)\ \mathbf{v}_\mathbf{j}(t) \quad ,$$

whereby $\dot{\lambda}_i(t) = c_{ii}$. In this paper, we denote complex conjugation of the entries of an eigenvector by an over-bar on the corresponding index. For example, $c_{\bar{i}\,j} = \mathbf{u}_\mathbf{i}^\mathbf{T}(t)\ \dot{M}(t)\ \mathbf{v}_\mathbf{j}(t)$ and since $\dot{M}(t)$ is real, $c_{i\,\bar{j}}\,c_{\bar{i}\,j} = \left|\mathbf{u}_\mathbf{i}^\mathbf{T}(t)\ \dot{M}(t)\ \mathbf{v}_\mathbf{j}(t)\right|^2$ is a real non-negative number.

With this notation Eqs. 10 and 11 read

$$\ddot{\lambda}_i(t) = \mathbf{u_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) + 2\sum_{j\neq i} c_{ij}c_{ji}\frac{\hat{\mathbf{r}}_{ij}}{\left|\mathbf{r}_{ij}\right|} \tag{13}$$

$$\ddot{\lambda}_i(t) = \mathbf{u_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) + 2\sum_{j\neq i} \left|c_{ij}\right|^2\frac{\hat{\mathbf{r}}_{ij}}{\left|\mathbf{r}_{ij}\right|} \quad \text{Normal matrices} \tag{14}$$

We think of Eq. 13 as

$$\begin{aligned}
\ddot{\lambda}_i(t) &= \mathbf{u_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) + 2\sum_{j\neq i} c_{ij}c_{ji}\frac{\hat{\mathbf{r}}_{ij}}{\left|\mathbf{r}_{ij}\right|} \\
&= \{\text{Inertial force of } M\} + \sum_{j\neq i}\{\text{force of } \lambda_j \text{ on } \lambda_i\}
\end{aligned}$$

[2] Strictly speaking Dirac's derivation of Equation 10 in Section 43 of this reference, does not hold in general (e.g., non-Hermitian) as the left eigenvectors are not 'bras' in his notation. The latter is a Hermitian conjugate of a standard (right) eigenvector. In his book, Dirac cites (Born, Heisenberg and Jordan, z.f. Physik 35, 565 (1925)) for these formulas

Remark 2. For non-normal matrices the force between any two eigenvalues is not necessarily repulsive nor attractive. As seen in Eq.13 , the orientation of the force is along the ray $c_{ij}c_{ji}\hat{\mathbf{r}}_{ij}$, rendering generally a non-central force law between the eigenvalues. This contrasts the purely central (and repulsive) nature of the interaction of the eigenvalues for normal matrices (Eq. 14). However, it is generally true that the force law between any two eigenvalues is inversely proportional to their distance.

Lemma 1. *(Eigenvalue Attraction) Complex conjugate eigenvalues of $M(t)$ attract (see Def. 1) as long as $M(t)$ is real and $\mathbf{u_i^T}(t)\ \dot{M}(t)\ \mathbf{v_i}(t) \neq 0$.*

Proof. The non-real eigenvalues and eigenvectors of a real matrix come in c.c. pairs [21, Chapter 24]. In Eq. 13, the interaction of $\overline{\lambda_i(t)}$ and $\lambda_i(t)$ is given by the term where $j = \bar{i}$ where $\lambda_i(t) - \overline{\lambda_i(t)} = 2i\ \mathrm{Im}(\lambda_i)$ and

$$\frac{2\, c_{i\,\bar{i}}\, c_{\bar{i}\,i}}{\lambda_i(t) - \overline{\lambda_i(t)}} \;=\; -i\,\frac{|c_{i\,\bar{i}}|^2}{\mathrm{Im}(\lambda_i(t))} \quad . \tag{15}$$

i. $\mathrm{Im}(\lambda_i) > 0$, then the right hand side of Eq. 15 is a negative imaginary number: The effect of $\overline{\lambda_i}$ on λ_i at time t is to push λ_i downwards along the imaginary axis with a magnitude that is inversely proportional to their distance. The constant of proportionality is $|c_{i\,\bar{i}}|^2$– the numerator of Eq. 15.

ii. $\mathrm{Im}(\lambda_i) < 0$, then the right hand side of Eq. 15 is a positive imaginary number: The effect of $\overline{\lambda_i}$ on λ_i at time t is to push λ_i upward along the imaginary axis with a magnitude that is inversely proportional to their distance with the same constant of proportionality. □

Comment: Let M be a fixed non-symmetric real matrix and define $M(t) = M + t\,\mathbb{I}$, where $\mathbb{I}$ is the identity matrix. We expect the eigenvalues of $M(t)$ to simply be the same as those of M shifted by t. No attraction is expected. Indeed, by Eq. 6 $c_{ii} = \mathbf{u_i^T}(t)\,\mathbf{v_i}(t) = 0$. Moreover, from Eq. 7 we have $\dot{\lambda}_i(t) = t$, i.e., a net drift in the complex plane of any eigenvalue with velocity t.

Remark 3. Proving that c.c. eigenvalues attract does *not* imply that in the long run all the eigenvalues will necessarily be real. There are three forces that act on any eigenvalue: 1. The inertial force due to $\ddot{M}(t)$. 2. The attraction of its c.c. 3. The force of the remaining $n-2$ eigenvalues. The governing equation is accordingly written as

$$\ddot{\lambda}_i(t) = \mathbf{u_i}^*(t)\ \ddot{M}(t)\ \mathbf{v_i}(t) - i\,\frac{|c_{i\,\bar{i}}|^2}{\mathrm{Im}(\lambda_i)} + 2\sum_{j\neq\{i,\bar{i}\}} c_{ij}c_{ji}\frac{\hat{\mathbf{r}}_{ij}}{|\mathbf{r}_{ij}|} \quad . \tag{16}$$

Since the force law is inversely proportional to the distance of the eigenvalues, the law of interaction needs to be logarithmic i.e., the potential is $V\left(|\mathbf{r}_{ij}|\right) \propto \log|\mathbf{r}_{ij}|$. The logarithmic interaction on the plane implies a short-range force, where the main contribution to the total force comes from the eigenvalues that are in the vicinity.

A natural question then is: how dominant is the c.c. attraction in determining the motion of any eigenvalue? One cannot give a general answer to this question as the relevance of c.c. attraction depends on the particular $M(t)$ and the particular eigenvalue considered. However, in examining Eq. 16, it is clear that c.c. pairs closest to the real line attract strongest. We will show that ill-conditioning enhances the c.c. attraction as well. This is expected as ill-conditioning generally implies higher sensitivity to perturbations [1, 21]. Below we provide some exact results and demonstrations for certain general sub-manifolds of real matrices and time evolutions that are of interest. Mathematically, whether the second term in Eq. 16 dominates the time evolution depends on its magnitude relative to the first derivative and the other terms on the right hand side (i.e, third term). After some demonstrations of dterministic evolution in the next section, we proceed to calculate the expectation values and variances about that expectation for the first and second derivatives when the evolution is driven by randomness. Further, we derive conditions under which the total expected force is only due to the c.c..

Remark 4. When the dominant force is that of the c.c. attraction, it is interesting that a matrix with a simple spectrum is forced to form degeneracies in its eigenvalues even if the perturbation (or stochastic process below) is generic. This happens momentarily when any λ_i and $\overline{\lambda_i}$ collide on the real line, whereby standard perturbation theory and considerations above breaks down. At this moment the matrix of eigenvalues forms 2×2 Jordan block(s) [19, Chapter 1].

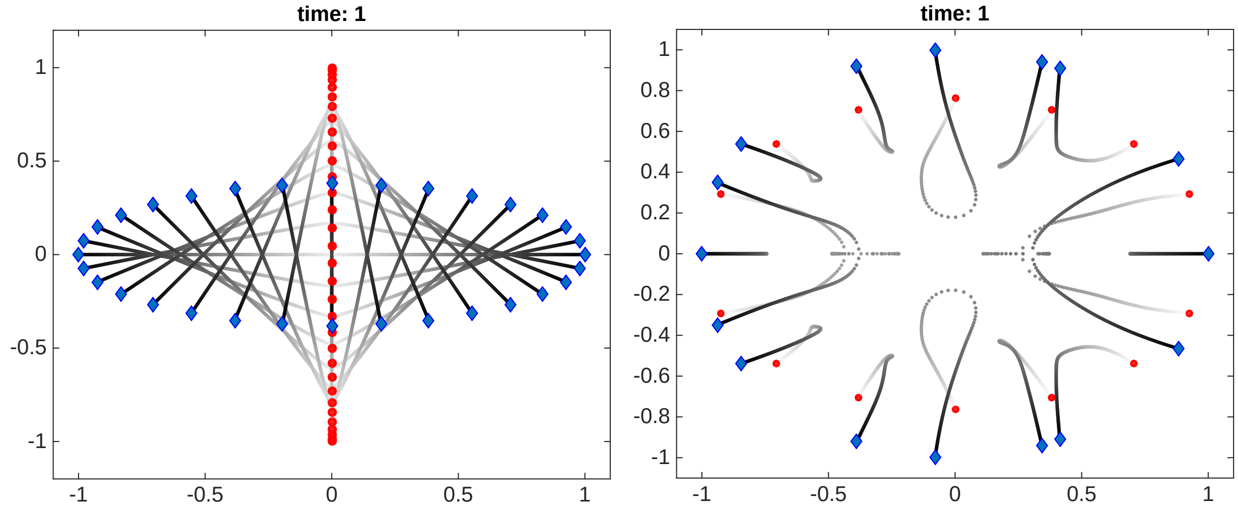


Figure 2: Example 1 (Left): and Example 2 (Right)

III. APPLICATIONS OF EIGENVALUE ATTRACTION

A. Smooth interpolation between fixed matrices

As demonstrations of Lemma 1, let us take $M(t) = (1-t)\,M_1 + t\,M_2$, where M_1 and M_2 are two fixed matrices and $t \in [0,1]$; the eigenvalues of $M(t)$ interpolate between the two. In Figs 2 and 3, as before, the eigenvalues of M_1 are shown in red filled circles and they darken till their final position, which is the eigenvalues of M_2 shown in blue diamonds. We take the size of the matrices to be 16×16 so the trajectories of eigenvalues can be visually traced easier. We normalize all the matrices to have unit 2−norm. The examples shown Figs 2 and 3 are:

Example 1: $M_1 = \begin{bmatrix} 0 & +1 & & & \\ -1 & 0 & +1 & & \\ & -1 & 0 & \ddots & \\ & & \ddots & \ddots & +1 \\ & & & -1 & 0 \end{bmatrix}$ and M_2 is Hatano-Nelson with $g = -0.4$.

Example 2: M_1 is Hatano Nelson with $g = 1$ and M_2 is a random orthogonal matrix.

Example 3: M_1 and M_2 are two real matrices whose entries are iid from the standard normal distribution.

Example 4: M_1 is Hatano Nelson with $g = -0.3$ and M_2 is Hatano-Nelson with $g = +0.3$, note the the spectra of M_1 and M_2 coincide and that every one of the c.c. pairs meet on the real line and walk all the way to the other side.

Example 5: M_1 and M_2 are both random orthogonal matrices.

In all the examples shown here, we expect the c.c. attraction to be more dominant among other forces for pairs closest to the real line. Note that these c.c. pairs first collide and then they move towards their final positions (shown in blue diamonds). A nice corollary is that a real eigenvalue, will remain real in the time evolution. Since the non-real eigenvalues of $M(t)$ come in c.c. pairs, a real eigenvalue cannot move off the real axis. Leaving the real axis can only happen if two eigenvalues collide on the real line and scatter off into the complex plane and remain mirror images (i.e., c.c.). For example such collisions are seen in Examples 2 and 3, where in the latter, one can see a cascade of this. That is a c.c. pair become real and while moving on the real line a new collision takes the eigenvalues off the real line for a short while, but attraction pulls them back in again for a second collision.

B. Random perturbations of a matrix

In this section we focus on perturbations of a fixed matrix which come up often in sciences and engineering, where $M(t)$ has the form

$$M(t) = M + \delta t\, P \tag{17}$$

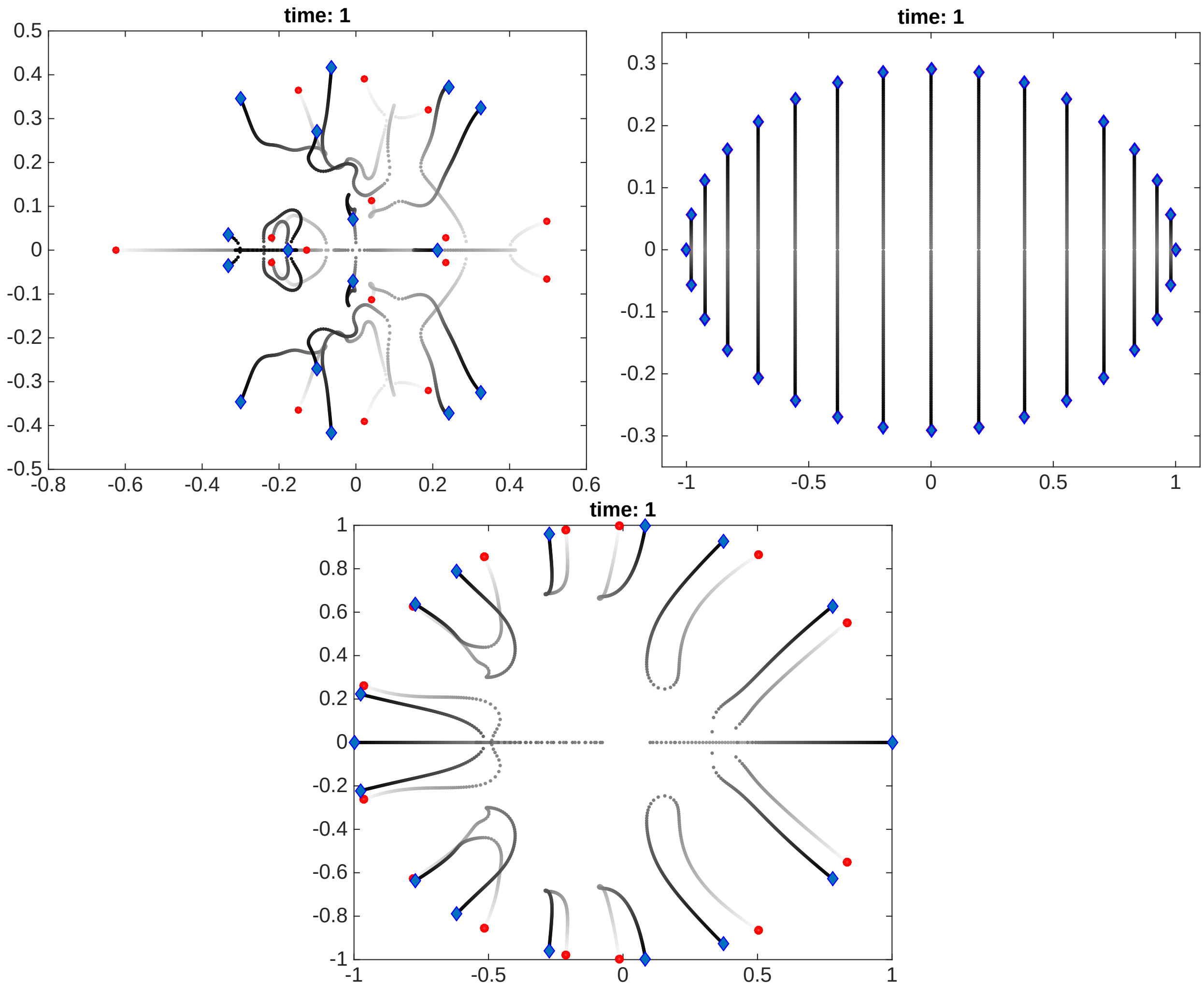


Figure 3: Example 3 (Left). Example 4 (Right) and Example 5 (bottom)

where M and P are $n \times n$ real matrices, M is fixed, P is a random matrix whose entries are independently and identically distributed (iid) with zero mean and bounded moments and δt is a real small parameter such that $|\delta t| \, \|P\| \ll \|M\|$; i.e., $\delta t \; P$ is a perturbation to M. The discussion of this section is important for the formulation and analysis of stochastic dynamics of the eigenvalues in Section IV. Recall that the velocity and acceleration of an eigenvalue are

$$\begin{aligned} \dot{\lambda}_i(t) &= c_{ii}. \\ \ddot{\lambda}_i(t) &= -i \frac{|c_{i\bar{i}}|^2}{\mathrm{Im}(\lambda_i)} + 2 \sum_{j \neq \{i,\bar{i}\}} c_{ij} c_{ji} \frac{\hat{\mathbf{r}}_{ij}}{|\mathbf{r}_{ij}|} \end{aligned}$$

where $c_{ij} \equiv \mathbf{u}_\mathbf{i}^* (t) \; P \; \mathbf{v}_\mathbf{j} (t)$ and $\frac{\hat{\mathbf{r}}_{ij}}{|\mathbf{r}_{ij}|} = \frac{1}{\lambda_i - \lambda_j}$ as before.

C. First variation

The expected first variation with respect to entries of P is

$$\mathbb{E}\left[\dot{\lambda}_i\left(t\right)\right] = \mathbb{E}\left[P_{ab}\right]\ \overline{u_i}^a v_i^b = 0$$

where for notational convenience from now on we drop the dependence of the eigenpairs on t and assume that the repeated indices that label the components of any eigenvector are summed over. We comment that $\mathbb{E}\left[\dot{\lambda}_i\left(t\right)\right] = 0$ even when entries of P are not identically distributed.

The variance is $\mathbb{E}\left[\left|\dot{\lambda}_i\right|^2\right] - \left|\mathbb{E}\left[\dot{\lambda}_i\right]\right|^2$. From above $\left|\dot{\lambda}_i\right|^2 = P_{ab}P_{cd}\bar{u}_i^{\prime a} v_i^b u_i^c \bar{v}_i^{\prime d}$ and $\mathbb{E}\left[\dot{\lambda}_i\right] = 0$. In general, the variance of the first variation denoted by $\sigma_{i,1}^2$ is $\mathbb{E}\left[\left|\dot{\lambda}_i\right|^2\right] = \mathbb{E}\left[p^2\right] \bar{u}_i^{\prime a} v_i^b u_i^a \bar{v}_i^{\prime b} = \mathbb{E}\left[p^2\right] \|\mathbf{u}_i\|_2^2$. Moreover, if M is a normal matrix then $\|\mathbf{u}_i\|_2^2 = \|\mathbf{v}_i\|_2^2 = 1$.

Next suppose that $P = \mathrm{diag}\left(p_1, p_2, \ldots, p_n\right)$ with $\mathbb{E}\left[p_i\right] = 0$. Then $\left|\dot{\lambda}_i\right|^2 = P_a P_b \bar{u}_i^{\prime a} v_i^a u_i^b \bar{v}_i^{\prime b}$. Since $\mathbf{u}_i^* \mathbf{v}_i = 1$, for diagonal perturbation we have $\mathbb{E}\left[\left|\dot{\lambda}_i\right|^2\right] = \mathbb{E}\left[p^2\right]$.

The special cases are worth summarizing

$\sigma_{i,1}^2$	M	P
$\mathbb{E}\left[p^2\right] \|\mathbf{u}_i\|_2^2$	General	General
$\mathbb{E}\left[p^2\right]$	General	Diagonal
$\mathbb{E}\left[p^2\right]$	Normal	General

Now suppose that M is a *circulant* matrix and P is diagonal. In this case (note that below no expectation is taken)

$$\dot{\lambda}_i = \overline{v_i}^{\prime a} P_a v_i^a = P_a \left|v_i^{\prime a}\right|^2 = \frac{1}{n}\sum_a P_a. \qquad \text{Circulant.} \tag{18}$$

This expression is the empirical mean of the diagonal entries of P, which for large n tends to zero– The dynamics of the eigenvalues are primarily governed by the second variation.

D. Second variation

When $M\left(t\right)$ is the line of matrices given by Eq. 17, $\ddot{M}\left(t\right) = 0$, $\dot{M}\left(t\right) = P$ and

$$\ddot{\lambda}_i = -i\frac{\left|\mathbf{u_i^T}\ P\ \mathbf{v_i}\right|^2}{2\ \mathrm{Im}\left(\lambda_i\right)} + \sum_{j\neq\{i,\bar{i}\}} \frac{\left[\mathbf{u_i^*}\ P\ \mathbf{v_j}\right]\left[\mathbf{u_j^*}\ P\ \mathbf{v_i}\right]}{\lambda_i - \lambda_j}. \tag{19}$$

To quantify the dominance of the attractive force, we calculate the expectation and variance of in Eq. 16

$$\begin{aligned}\mathbb{E}\sum_{j\neq\{i,\bar{i}\}} \frac{\left[\mathbf{u_i^*}\ P\ \mathbf{v_j}\right]\left[\mathbf{u_j^*}\ P\ \mathbf{v_i}\right]}{\lambda_i - \lambda_j} &= \mathbb{E}\sum_{j\neq\{i,\bar{i}\}} \frac{\left[\overline{u_i}^a P_{ab} v_j^b\right]\left[\overline{u_j}^c P_{cd} v_i^d\right]}{\lambda_i - \lambda_j} = \sum_{j\neq\{i,\bar{i}\}} \frac{\mathbb{E}\left[P_{ab}P_{cd}\right]\left[\overline{u_i}^a v_j^b \overline{u_j}^c v_i^d\right]}{\lambda_i - \lambda_j} \\ &= \mathbb{E}\left[p^2\right]\sum_{j\neq\{i,\bar{i}\}} \frac{\left(\mathbf{v_i^T}\mathbf{v_j}\right)\left(\mathbf{u_i^*}\overline{\mathbf{u_j}}\right)}{\lambda_i - \lambda_j},\end{aligned} \tag{20}$$

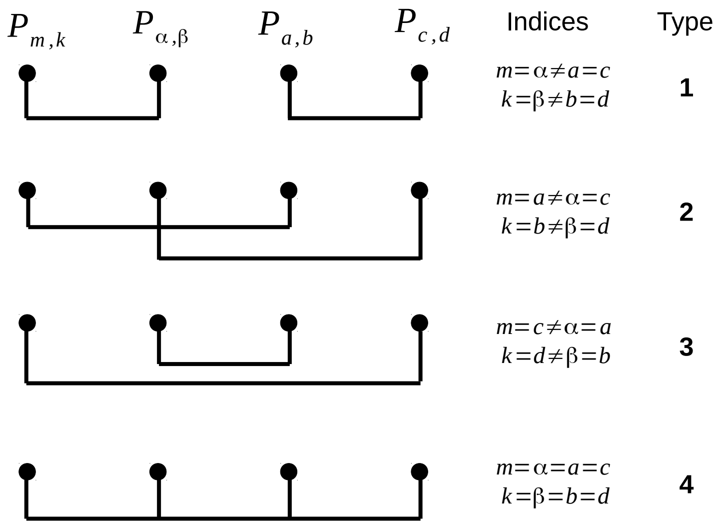


Figure 4: Non-zero contribution in $\mathbb{E}\left\{P_{mk}P_{\alpha\beta}P_{ab}P_{cd}\right\}$

where p denotes any entry of P and $\mathbb{E}\left[P_{ab}P_{cd}\right] = \mathbb{E}\left[p^2\right]\delta_{ac}\delta_{bd}$. This is the general form of the expected force of the remaining $n-2$ eigenvalues on λ_i (excluding its c.c.). Although, the attraction is always present, we remark that similar treatment of the first term in Eq. 19 leads to

$$\mathbb{E}\left[\ddot{\lambda}_i\right] = -i\frac{\mathbb{E}\left[p^2\right]\|\mathbf{u_i}^*\|_2^2}{\mathrm{Im}\left(\lambda_i\right)} + \mathbb{E}\left[p^2\right]\sum_{j\neq\{i,\bar{i}\}}\frac{\left(\mathbf{v_i^T v_j}\right)\left(\mathbf{u_i^*}\overline{\mathbf{u_j}}\right)}{\lambda_i-\lambda_j} \quad .$$

We now calculate the variance of the second term, denoted by $\sigma_{i,2}^2$, in Eq. 19

$$\begin{aligned}
\mathbb{E}\left[\sigma_{i,2}^2\right] &= \mathbb{E}\left\{\left|\sum_{j\neq\{i,\bar{i}\}}\frac{\left[\mathbf{u_i^*}\ P\ \mathbf{v_j}\right]\left[\mathbf{u_j^*}\ P\ \mathbf{v_i}\right]}{\lambda_i-\lambda_j}\right|^2\right\} - \left|\mathbb{E}\sum_{j\neq\{i,\bar{i}\}}\frac{\left[\mathbf{u_i^*}\ P\ \mathbf{v_j}\right]\left[\mathbf{u_j^*}\ P\ \mathbf{v_i}\right]}{\lambda_i-\lambda_j}\right|^2 \\
&= \sum_{\ell,j\neq\{i,\bar{i}\}}\frac{\mathbb{E}\left[P_{mk}P_{\alpha\beta}P_{ab}P_{cd}\right]\bar{u}_{i'}{}^{m} v_j^k \bar{u}_j^{\alpha} v_i^{\beta} u_i^a \bar{v}_\ell^b u_\ell^c \bar{v}_i^d}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} - \mathbb{E}^2\left[p^2\right]\left|\sum_{j\neq\{i,\bar{i}\}}\frac{\left(\mathbf{v_i^T v_j}\right)\left(\mathbf{u_i^*}\overline{\mathbf{u_j}}\right)}{\lambda_i-\lambda_j}\right|^2 . \qquad (21)
\end{aligned}$$

There are potentially four nonzero contribution to the first sum shown in Fig. 4. We proceed to calculate them one by one.

Type 1

The contribution of this type to Eq. 21 is

$$\begin{aligned}
\mathbb{E}^2\left[p^2\right]\sum_{j,\ell\neq(i,\bar{i})}\frac{\bar{u}_{i'}{}^{m} v_j^k \bar{u}_j^m v_i^k u_i^a \bar{v}_\ell^b u_\ell^a \bar{v}_i^b}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} &= \mathbb{E}^2\left[p^2\right]\sum_{j,\ell\neq(i,\bar{i})}\frac{\left(\mathbf{u}_i^*\bar{\mathbf{u}}_j\right)\left(\mathbf{v}_j^T\mathbf{v}_i\right)\left(\mathbf{u}_i^T\mathbf{u}_\ell\right)\left(\mathbf{v}_\ell^*\bar{\mathbf{v}}_i\right)}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} \\
&= \mathbb{E}^2\left[p^2\right]\left|\sum_{\ell\neq(i,\bar{i})}\frac{\left(\mathbf{v}_\ell^T\mathbf{v}_i\right)\left(\mathbf{u}_i^*\bar{\mathbf{u}}_\ell\right)}{\lambda_i-\lambda_\ell}\right|^2 \quad .
\end{aligned}$$

where the sum over j and ℓ are complex conjugates. Note that this term cancels the second term in Eq. 21.

Type 2

Since $\|\mathbf{v}_i\|_2^2 = 1$, the contribution of this type to Eq. 21 is

$$\mathbb{E}^2\left[p^2\right]\sum_{j,\ell\neq(i,\bar{i})}\frac{\bar{u}_i{}'^a v_j^b \bar{u}_j^c v_i^d u_i^a \bar{v}_\ell^b u_\ell^c \bar{v}_i^d}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} = \mathbb{E}^2\left[p^2\right]\|\mathbf{u}_i\|_2^2\sum_{j,\ell\neq(i,\bar{i})}\frac{\left(\mathbf{v}_\ell^*\mathbf{v}_j\right)\left(\mathbf{u}_j^*\mathbf{u}_\ell\right)}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} \quad .$$

Type 3

Similarly to Type 1, the contribution of this type to Eq. 21 is calculated to be

$$\mathbb{E}^2\left[p^2\right]\sum_{\ell,j\neq\{i,\bar{i}\}}\frac{\left(\mathbf{u}_j^*\mathbf{u}_i\right)\left(\mathbf{u}_i^*\mathbf{u}_\ell\right)\left(\mathbf{v}_\ell^*\mathbf{v}_i\right)\left(\mathbf{v}_i^*\mathbf{v}_j\right)}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} \;=\; \mathbb{E}^2\left[p^2\right]\left|\sum_{\ell\neq\{i,\bar{i}\}}\frac{\left(\mathbf{u}_\ell^*\mathbf{u}_i\right)\left(\mathbf{v}_i^*\mathbf{v}_\ell\right)}{\left(\lambda_i-\lambda_\ell\right)}\right|^2 \quad .$$

Type 4

The contribution of this type to Eq. 21 is

$$\mathbb{E}\left[p^4\right]\sum_{\ell,j\neq\{i,\bar{i}\}}\frac{\bar{u}_i{}'^a v_j^b \bar{u}_j^a v_i^b u_i^a \bar{v}_\ell^b u_\ell^a \bar{v}_i^b}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} = \mathbb{E}\left[p^4\right]\sum_{\ell,j\neq\{i,\bar{i}\}}\frac{\left|u_i^a\right|^2\bar{u}_j^a u_\ell^a\left|v_i^b\right|^2 v_j^b\bar{v}_\ell^b}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} \quad .$$

We conclude that Eq. 21

$$\begin{aligned}\mathbb{E}\left[\sigma_{i,2}^2\right] \;=\; & \mathbb{E}^2\left[p^2\right]\left\{\|\mathbf{u}_i\|_2^2\sum_{j,\ell\neq(i,\bar{i})}\frac{\left(\mathbf{v}_\ell^*\mathbf{v}_j\right)\left(\mathbf{u}_j^*\mathbf{u}_\ell\right)}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)}+\left|\sum_{\ell\neq\{i,\bar{i}\}}\frac{\left(\mathbf{u}_\ell^*\mathbf{u}_i\right)\left(\mathbf{v}_i^*\mathbf{v}_\ell\right)}{\left(\lambda_i-\lambda_\ell\right)}\right|^2\right\} \\ & +\;\mathbb{E}\left[p^4\right]\sum_{\ell,j\neq\{i,\bar{i}\}}\frac{\left|u_i^a\right|^2\bar{u}_j^a u_\ell^a\left|v_i^b\right|^2 v_j^b\bar{v}_\ell^b}{\left(\lambda_i-\lambda_j\right)\left(\bar{\lambda}_i-\bar{\lambda}_\ell\right)} \quad .\end{aligned} \tag{22}$$

The derivations of this section are purely theoretical. In practice, actual estimates of Eq. 22 would depend on the particular $M(t)$; i.e., P and the eigenpairs of M. There are universality results on the eigenvectors of generic Hermitian matrices such as generalized Wigner matrices [22] but not nearly as much is known for the generic matrices. Once an estimate is found, convergence to the expected value of the other forces can be proved using standard techniques such as Markov's inequality. Below we discuss special cases that are of theoretical and applied interest.

E. Normal and Circulant matrices

Now let us confine to the case where M is a normal matrix , i.e., unitary diagonalizable [21].

Corollary 1. *Suppose M is a normal matrix and P is a random matrix with entries that are iid and have zero mean. Then the total expected force on any eigenvalue is only due to its complex conjugate.*

Proof. If M is normal, its eigenvectors form an orthonormal basis for $\mathbb{C}^n$ and $\mathbf{u}_i = \mathbf{v}_i$ with $\mathbf{v_i^*}\mathbf{v_j} = \delta_{ij}$. Eq. 20 gives the expected force on λ_i by all other eigenvalues excluding its c.c., which now reads

$$\mathbb{E}\left[p^2\right]\sum_{j\neq\{i,\bar{i}\}}\frac{\left|\mathbf{v_i^T v_j}\right|^{\mathbf{2}}}{\lambda_i-\lambda_j} = 0$$

since the sum over j excludes $\bar{i}$, by orthonormality of the eigenvectors $\mathbf{v_i^T v_j} = 0$. Hence,

$$\mathbb{E}\left[\ddot{\lambda}_i\right] = -i\frac{\mathbb{E}\left[p^2\right]}{\mathrm{Im}\left(\lambda_i\right)} \qquad \text{Normal .} \tag{23}$$

The *total* expected force is only due to the c.c. □

The variance takes on a simpler expression as well. In Eq. 22 the first and third sums are zero because $\ell \neq \{i, \bar{i}\}$ and therefore $\mathbf{v}_\ell^T \mathbf{v}_i = 0$ and $\mathbf{v}_i^* \mathbf{v}_\ell = 0$. Moreover, in the sum in Eq. 22, the sum over j collapses because $\mathbf{v}_\ell^* \mathbf{v}_j = \delta_{\ell j}$ and $\|\mathbf{u}_i\|_2^2 = 1$.

Type 4 can be simplified as follows

$$\begin{aligned}\sum_{j\neq(i,\bar{i})}\sum_{\ell\neq(i,\bar{i})}\frac{\mathbb{E}\left[P_{ab}^4\right]\ \bar{v}_i'^a v_j'^b \bar{v}_j^a v_i'^b\ \bar{v}_i'^a v_\ell'^b \overline{v}_\ell'^a v_i'^b}{\left(\lambda_i-\lambda_j\right)\overline{\left(\lambda_i-\lambda_\ell\right)}} &= \mathbb{E}\left[p^4\right]\sum_{j\neq(i,\bar{i})}\sum_{\ell\neq(i,\bar{i})}\frac{\left|v_i'^a\right|^2\left|v_i'^b\right|^2 \bar{v}_j^a v_\ell'^a\ v_j'^b\ \overline{v}_\ell'^b}{\left(\lambda_i-\lambda_j\right)\overline{\left(\lambda_i-\lambda_\ell\right)}} \\ &\leq \mathbb{E}\left[p^4\right]\left|\sum_{\ell\neq(i,\bar{i})}\frac{1}{\lambda_i-\lambda_\ell}\right|^2 .\end{aligned} \tag{24}$$

because $\sum_{ab}\left|v_i'^a\right|^2\left|v_i'^b\right|^2 \bar{v}_j^a v_\ell'^a\ v_j'^b\ \overline{v}_\ell'^b = \left|\sum_a\left|v_i'^a\right|^2 \bar{v}_j^a v_\ell'^a\right|^2 \leq \sum_a\left|\bar{v}_j^a v_\ell'^a\right|^2 \leq 1.$

Adding the contribution of the four Types in the case where *the matrix is normal*, we get

$$\mathbb{E}\left[\sigma_{i,2}^2\right] \leq \mathbb{E}\left[p^2\right]\sum_{\ell\neq(i,\bar{i})}\frac{1}{\left|\lambda_i-\lambda_\ell\right|^2} + \mathbb{E}\left[p^4\right]\left|\sum_{\ell\neq(i,\bar{i})}\frac{1}{\lambda_i-\lambda_\ell}\right|^2 . \tag{25}$$

where the inequality comes about from the estimate of the Type 4 terms.

How are we to visualize the force between any two eigenvalues? In general one can consider any eigenvalue, λ_i as a point in the complex plane with coordinates $(\mathrm{Re}\lambda_i, \mathrm{Im}\lambda_i)$ and the unit vector $\hat{\mathbf{r}}_{ij}$, given by Eq. 12, as one proportional to $\left(\mathrm{Re}\left(\lambda_i-\lambda_j\right), -\mathrm{Im}\left(\lambda_i-\lambda_j\right)\right)$. Looking at the Eq. 13, we note that the force exerted on λ_i from λ_j is in the direction of $c_{ij}c_{ji}\hat{\mathbf{r}_{ij}}$ (recall that $c_{ij} = \mathbf{u_i^*}(t)\ P\ \mathbf{v_j}(t)$). However, in the case of normal matrices $c_{ji} = \overline{c_{ij}}$ and the direction of the force is solely determined by $\hat{\mathbf{r}}_{ij}$. Even in the case of normal matrices, unlike Hermitian matrices, this force is not in general central (attractive or repulsive) as the imaginary part needs to be negated.

Remark 5. For the sake of argument let us consider the force on any eigenvalue of a normal matrix denoted by $\lambda_k \equiv a + ib$ by $\lambda_j = c + id$ and let $x \equiv a - c$ and $y \equiv d - b$ [3]. The direction of the force is in the direction of $\overline{\lambda_k} - \overline{\lambda_j} = (a-c) + i(d-b)$ as seen in Eq. 14. By examining the sign of the real/imaginary parts, we conclude that *any eigenvalue of a normal matrix experiences a repulsion in the real direction and an attraction in the imaginary direction from any other eigenvalue.* The total force is of course complicated to predict , which depends on relative position of the eigenvalues as well as the distances between them.

Next take $P = \mathrm{diag}\left(p_1, p_2, \ldots, p_n\right)$ and let M be a circulant matrix; note that now P does *not* have to be random. For circulant matrices $\mathbf{v}_j^T = \frac{1}{\sqrt{n}}\left[1, \omega_j, \omega_j^2, \ldots, \omega_j^{n-1}\right]$ with $\omega_j = \exp\left(2\pi i j/n\right)$ and for the force we get

$$\begin{aligned}\ddot{\lambda}_k &= \sum_{j\neq k}\frac{\left|\mathbf{v_k^*}\ P\ \mathbf{v_j}\right|^2}{\lambda_k-\lambda_j} = -i\frac{\left|v_k'^a p_a v_k'^a\right|^2}{2\ \mathrm{Im}\left(\lambda_k\right)} + \sum_{j\neq\{k,\bar{k}\}}\frac{\left|\bar{v}_k'^a p_a v_j'^a\right|^2}{\lambda_k-\lambda_j} \\ &= -i\frac{\frac{1}{n^2}\left|\sum_{a=1}^n p_a\omega_{2k}^{(a-1)}\right|^2}{2\ \mathrm{Im}\left(\lambda_k\right)} + \sum_{j\neq\{k,\bar{k}\}}\frac{\frac{1}{n^2}\left|\sum_{a=1}^n p_a\omega_{j-k}^{(a-1)}\right|^2}{\lambda_k-\lambda_j},\end{aligned} \tag{26}$$

[3] We denote the eigenvalue by λ_k instead of λ_i above because $i \equiv \sqrt{-1}$ appears more in the following discussion.

since $P^T = P$. The quantify $\frac{1}{n}\sum_{a=0}^{n-1} p_a \omega_{j-k}^{a-1}$ is the discrete Fourier transformation of the diagonal elements of P.

If we assume that the nonzero entries of P are standard Gaussians, Eq. 26 is simply the discrete Fourier transform of the white noise and $\frac{1}{n^2}\left|\sum_{a=0}^{n-1} p_a \omega_{j-k}^{a-1}\right|^2$ is the power intensity which is a real positive constant independent of the frequency $j-k$ [23], denoted here by κ^2 . So the total force is

$$\ddot{\lambda}_k = \kappa^2 \left\{ \frac{-i}{2\,\mathrm{Im}\,(\lambda_k)} + \sum_{j \neq \{k,\bar{k}\}} \frac{1}{\lambda_k - \lambda_j} \right\} \quad . \tag{27}$$

Remark 6. The foregoing equation and the following arguments can apply to the more general cases. For example, in Eq. 14 if P is taken to be diagonal to ensure reality of the numerator and that the numerators have similar magnitudes the following analysis applies.

Consider the case where b is small. The contribution of the terms from j and $\bar{j}$ to the sum in Eq. 27 is

$$\mathrm{Re}\left\{ \frac{1}{\lambda_k - \lambda_j} + \frac{1}{\lambda_k - \bar{\lambda}_j} \right\} = x \left\{ \frac{1}{x^2+y^2} + \frac{1}{x^2+y^2\,(1+2b/y)^2} \right\} \tag{28}$$

$$\mathrm{Im}\left\{ \frac{1}{\lambda_k - \lambda_j} + \frac{1}{\lambda_k - \bar{\lambda}_j} \right\} = y \left\{ \frac{1}{x^2+y^2} - \frac{1+2b/y}{x^2+y^2\,(1+2b/y)^2} \right\} \tag{29}$$

Evidently and as shown above any eigenvalue λ_j *repels* λ_k *along the real axis* since the sign of the force follows the sign of x (See the right figure in Fig.6). Let us suppose that $2b/y \equiv \epsilon \ll 1$ (e.g., force of bulk eigenvalues on a λ_k near the real line), then Eq. 29 becomes

$$\mathrm{Im}\left\{ \frac{1}{\lambda_k - \lambda_j} + \frac{1}{\lambda_k - \bar{\lambda}_j} \right\} = -\frac{y\epsilon}{x^2+y^2}\left\{1 - \frac{2y^2}{x^2+y^2}\right\}$$

to interpret this equation for now suppose that the imaginary parts of λ_k and λ_j are positive, then $\epsilon, y > 0$ and the effect of the imaginary part of the pair $\lambda_j, \bar{\lambda}_j$ on λ_k is to compress it towards the real axis as long as $x^2 > y^2$. Moreover the imaginary part of the force, unlike the real part, is $\mathcal{O}\,(\epsilon)$. Entirely a similar argument applies to the case where the imaginary parts of λ_k and λ_j are negative. Hence as long as the difference of the real parts is larger than that of the imaginary parts, *there is a compressive push towards the real line from any c.c. pair of eigenvalues on λ_k with a small net magnitude*. In many examples of circulant matrices the pair of eigenvalues closest to the real line appear on the "edges" of the spectrum and the assumptions made above are applicable. We conclude that in such cases, the eigenvalue closest to the real line will become real.

F. Hatano-Nelson Model

Let $H\,(t) = H + \delta t\, P$, where H is a circulant matrix given by Eq. 1 and P is a real diagonal matrix with iid and mean zero entries. The eigenpairs of H are easy to calculate

$$\begin{aligned} \lambda_k &= 2\left\{\cosh g\, \cos\left(2\pi k/n\right) + i\, \sinh g\, \sin\left(2\pi k/n\right)\right\} \\ \mathbf{v}_k^T &= \frac{1}{\sqrt{n}}\left[1, \omega_k, \omega_k^2, \ldots, \omega_k^{n-1}\right], \end{aligned}$$

where $\omega_k = \exp\left(2\pi i k/n\right)$. See Fig. 5 for the effect of the asymmetry parameter g on the spectrum of H.

Previously we showed that $\dot{\lambda}_i \approx 0$ for sufficiently large n (Eq. 18). Therefore the dynamics is governed by the acceleration $\ddot{\lambda}_k$.

Even though, H is normal, $H\,(t)$ is not. However, we observe that for small g, the eigenvalues rush to the real line almost undeflected. Why is the Hatano-Nelson model dominated by attraction towards the real line for relatively large t when g is small?

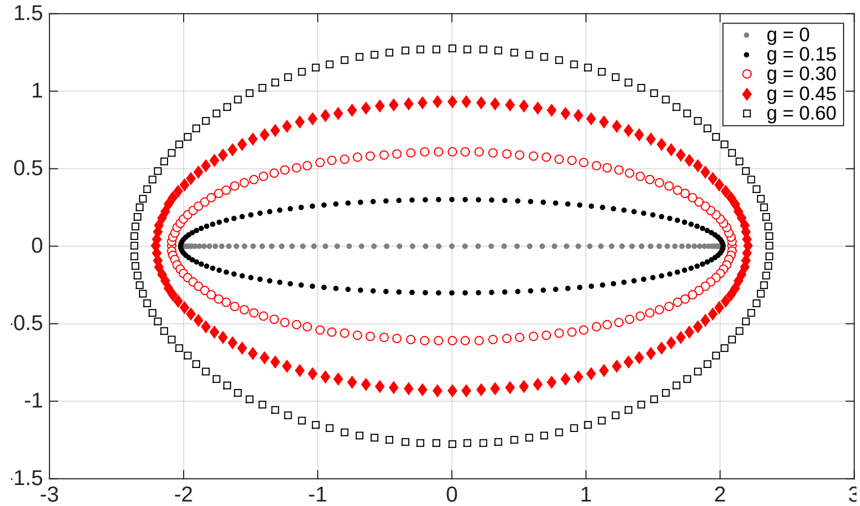


Figure 5: The spectrum of Hatano-Nelson model (Eq. 1) as a function of g. The spectrum is quite flat in the bulk for small g.

When $0 < g \ll 1$, $\lambda_k \approx 2\left\{\cos\left(2\pi k/n\right) + ig \sin\left(2\pi k/n\right)\right\}$ and Eq. 27 becomes

$$\ddot{\lambda}_k \approx \kappa^2 \left\{ \frac{-i}{2\,\mathrm{Im}\left(\lambda_k\right)} + \sum_{j \neq \{k,\bar{k}\}} \frac{1}{2\left[\cos\left(2\pi k/n\right) - \cos\left(2\pi j/n\right)\right]} - i\frac{g\,\left[\sin\left(2\pi k/n\right) - \sin\left(2\pi j/n\right)\right]}{2\left[\cos\left(2\pi k/n\right) - \cos\left(2\pi j/n\right)\right]^2} \right\}$$

In the foregoing sum we calculate the contribution of the sum of two terms given by j and $\bar{j}$. Ignoring $\mathcal{O}\left(g^2\right)$, we get

$$\frac{1}{\left[\cos\left(2\pi k/n\right) - \cos\left(2\pi j/n\right)\right]} - i\frac{g\,\sin\left(2\pi k/n\right)}{\left[\cos\left(2\pi k/n\right) - \cos\left(2\pi j/n\right)\right]^2} \quad .$$

Since $\mathrm{Im}\left(\lambda_k\right) \approx g\,\sin\left(2\pi k/n\right)$, the total force is

$$\frac{\ddot{\lambda}_k}{\kappa^2} \approx \sum_{\substack{j \neq \{k,\bar{k}\} \\ \mathrm{Im}\left(\lambda_j\right) > 0}} \left\{ \frac{1}{\left[\cos\left(2\pi k/n\right) - \cos\left(2\pi j/n\right)\right]} \right\} - i \left\{ \frac{1}{2\,\mathrm{Im}\left(\lambda_k\right)} + \sum_{\substack{j \neq \{k,\bar{k}\} \\ \mathrm{Im}\left(\lambda_j\right) > 0}} \frac{\mathrm{Im}\left(\lambda_k\right)}{\left[\cos\left(2\pi k/n\right) - \cos\left(2\pi j/n\right)\right]^2} \right\}. \tag{30}$$

We conclude that there is a net pull of any eigenvalue towards the real line as $\mathrm{Im}\left(\lambda_k\right) > 0$ gives a force in $-i$ direction and $\mathrm{Im}\left(\lambda_k\right) < 0$ gives a force in $+i$ direction. Moreover, the real part of the force of any eigenvalue on λ_k is repulsive as can easily be seen by analyzing the sign of the denominator of the first sum in the foregoing equation. The eigenvalues rush towards the real line and flow outwards away from the origin (See Fig. 6).

Comment: Take $k = n/4$, where $\lambda_{n/4} = i\,2g$, it is easy to see that $\mathrm{Re}\left(\ddot{\lambda}_{\frac{n}{4}}\right) = 0$ by rewriting the real part of the sum (Eq. 30) as

$$\mathrm{Re}\left(\frac{\ddot{\lambda}_{n/4}}{\kappa^2}\right) = \sum_{\substack{j \neq \{k,\bar{k}\} \\ \mathrm{Im}\left(\lambda_j\right) > 0}} \frac{-1}{\cos\left(2\pi j/n\right)} = \sum_{\ell=1}^{n/4-1} \frac{-1}{\cos\left(2\pi\left(n/4+\ell\right)/n\right)} - \frac{1}{\cos\left(2\pi\left(n/4-\ell\right)/n\right)} = 0 \quad .$$

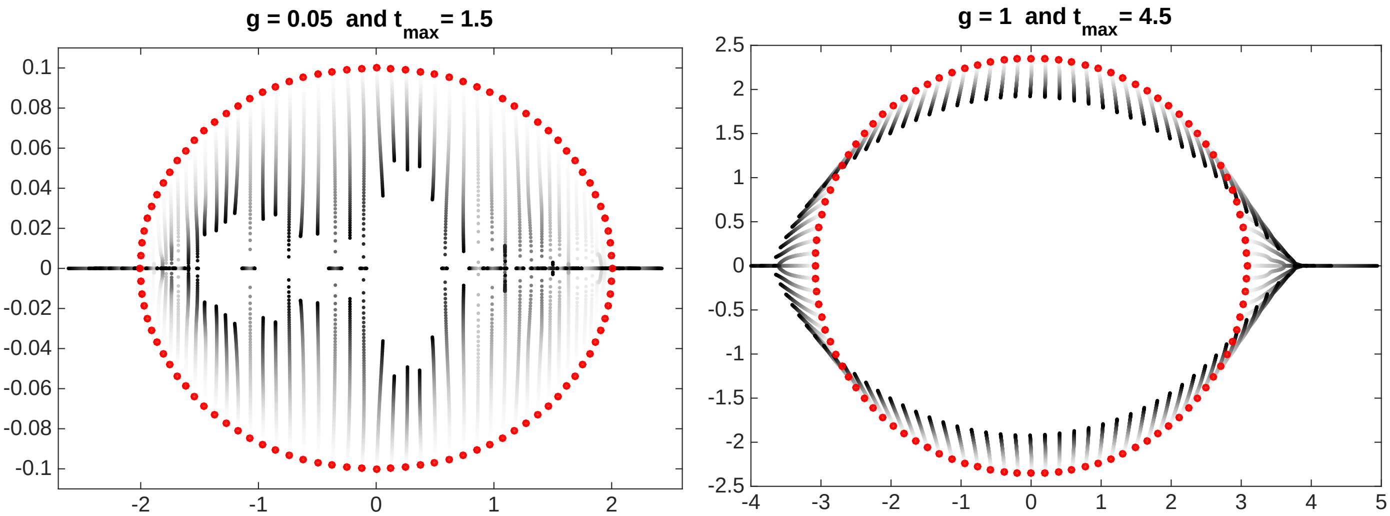


Figure 6: Dynamics of eigenvalues for small (left) g and larger (right) g. Also note the difference in range of the imaginary axis and the repulsion along the real axis in agreement with the discussion of the previous section.

So the net force on this eigenvalue is purely imaginary. In this limit (i.e., $g \ll 1$), the spectrum is approximately an ellipse with semi-minor axis $i\,2g$ and semi-major axis 2 (Figs. 5 and 6). The spectrum is quite flat for the eigenvalues with small real parts. Therefore the force on these eigenvalues is purely imaginary explaining their almost undeflected rush towards the real line (Fig. 6).

Although H is a circulant matrix and hence normal, $H + \delta t\, P$ is not, but the degree of non-normality is mild for small g as the following shows. The following combined with Corollary 1, imply that the expected force on any eigenvalue in this model is only due to its c.c. . We quantify the degree of non-normality by looking at $\left[H\left(t\right), H^{T}\left(t\right)\right] = t\left\{[H,P] - \left[H^{T},P\right]\right\}$, which is calculated to be

$$\left[H\left(t\right), H^{T}\left(t\right)\right] = 2t\,\sinh g \begin{bmatrix} 0 & p_1 - p_2 & 0 & \cdots & 0 & p_n - p_1 \\ p_1 - p_2 & 0 & p_2 - p_3 & & & 0 \\ 0 & p_2 - p_3 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & 0 & p_{n-2} - p_{n-1} & 0 \\ 0 & & & p_{n-2} - p_{n-1} & 0 & p_{n-1} - p_n \\ p_n - p_1 & 0 & \cdots & 0 & p_{n-1} - p_n & 0 \end{bmatrix}.$$

Let us take each p_i to be randomly distributed with mean μ, then the commutator matrix in an expectation sense (with respect to entries) is the zero matrix. Moreover, the entries are scaled by the pre-factor; for example, if we take p_i to be normally distributed with mean μ and variance σ^2 (i.e., $\mathcal{N}\left(\mu, \sigma^2\right)$, then $X_i \equiv \left(p_{i+1} - p_i\right) \sim \mathcal{N}\left(0, 2\sigma^2\right)$ and each entry has mean zero and variance $4t^2 \sinh^2 g$. Therefore fluctuations are small so long as g is small where the variance is approximately $4t^2 g^2$. In this limit one expects the dynamics, in an expectation sense, to be dominated by attraction of c.c. pairs (See Fig. 1).

Remark. We intentionally did not use the Frobenius norm to quantify normality as it would not appreciate the entries being in an expectation sense zero despite $\mathbb{E}\left(p\right) \neq 0$. Take as a measure of non-normality the ratio $\left\|\left[H\left(t\right), H^{T}\left(t\right)\right]\right\|_F / \left\|H\right\|_F^2$. But $\left\|H\right\|_F^2 = 2n\,\cosh 2g$ and $\left\|\left[H\left(t\right), H^{T}\left(t\right)\right]\right\|_F^2 = 2t\,\sinh g\left\{2\sum_i\left(p_{i+1} - p_i\right)^2\right\}$, where $p_{n+1} = p_1$. and $\sum_i X_i^2 \sim 2\sigma^2 \chi_{n-1}^2$, where χ_{n-1}^2 denotes a chi-square distribution with $n-1$ degrees of freedom and we have

$$\frac{\left\|\left[H\left(t\right), H^{T}\left(t\right)\right]\right\|_F}{\left\|H\right\|_F^2} \sim \frac{\sigma\sqrt{2t\,\sinh g}}{n\,\cosh 2g}\chi_{n-1} \quad .$$

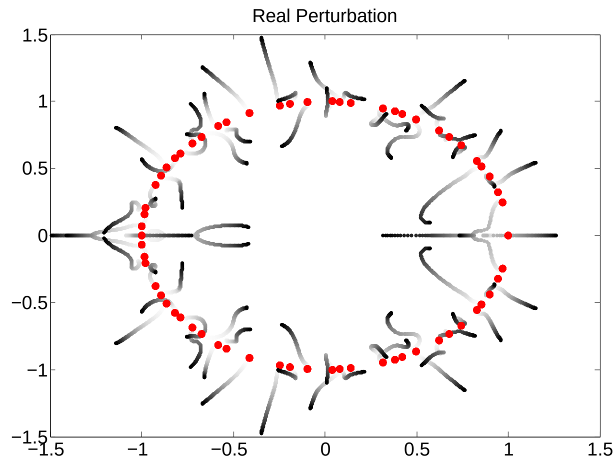


Figure 7: Trajectories of the eigenvalues of $M + tP$, where M is an orthogonal matrix and $t_{max} = 2$, when they become black. P is a real random gaussian matrix with norm 1.

The mean of the $\chi-$distribution is $\sqrt{2}\Gamma\left[(n+1)/2\right]/\Gamma\left(n/2\right)$, which for large n is approximately $\sqrt{2n}$ and we have

$$\mathbb{E}\left[\frac{\left\|\left[H\left(t\right),H^{T}\left(t\right)\right]\right\|_{F}}{\|H\|_{F}^{2}}\right]=\frac{2\sigma\sqrt{t}\,\sinh g}{\sqrt{n}\,\cosh 2g}\quad .$$

First lets consider $g \gg 1$, where H becomes proportional to a permutation matrix. In this limit the foregoing expectation is zero. Second let take $g \ll 1$, in which case the expectation is approximately $2\sigma\sqrt{\frac{tg}{n}}$.

G. Further illustrations

As the first illustration, suppose we perturb a 64×64 real orthogonal matrix, M, with $t\, P$, where P is a real random matrix with Gaussian entries. In Figure 7 we show the motion of the eigenvalues for $t \in [0, 2]$.

An application of this work is a better understanding of the origin of real eigenvalues in the Hatano-Nelson model as discussed in Sec. I (Fig. 1). In Fig. 6 one can see the formation of wings mentioned Sec. I. Despite the base case being the same, the motion of the eigenvalues of the Hatano-Nelson model in Figs. 1, a and Fig. 6 is much more uniform than in Fig. 1,b. In the latter plot in the course of the evolution there were 50 intervals with a different random P acting in each. Although continuous, this makes the motion jittery.

In Fig. 8 we take M and P to be two independent 32×32 real random Gaussian matrices with unit $2-$norm. Here we also see cases where c.c. pairs become real eigenvalues first and then as a result of yet another encounter leave the real line by forming a c.c. pair with the newly encounter eigenvalue.

As our last illustrations, we take M to be a random orthogonal matrix and take the perturbation matrix, P, to be a random ± 1 matrix (Fig. 9 left) or another random orthogonal matrix (Fig. 9 right). In the latter, M and P both have eigenvalues on the unit circle (see Fig. 8).

IV. SPECIAL CASE OF STOCHASTIC DYNAMICS OF THE EIGENVALUES

The eigenvalues of a continuous stochastic process are continuous in t.

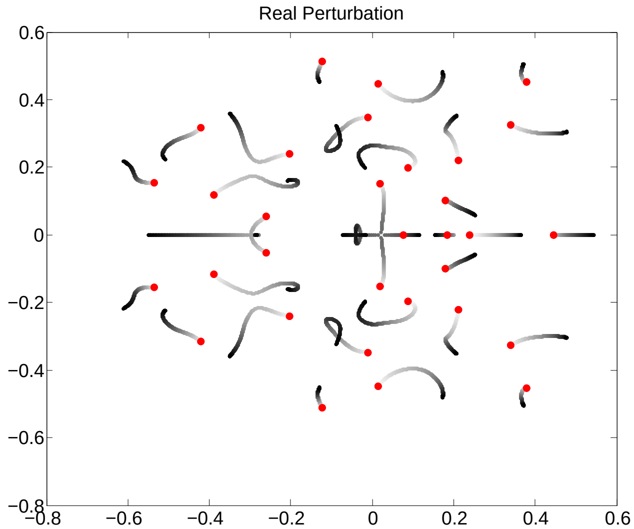


Figure 8: Eigenvalues of $M+tP$, where M is a real Gaussian random matrix of size 32 with unit 2−norm. P is a real random Gaussian matrix with unit 2−norm. We took $t_{max}=0.5$.

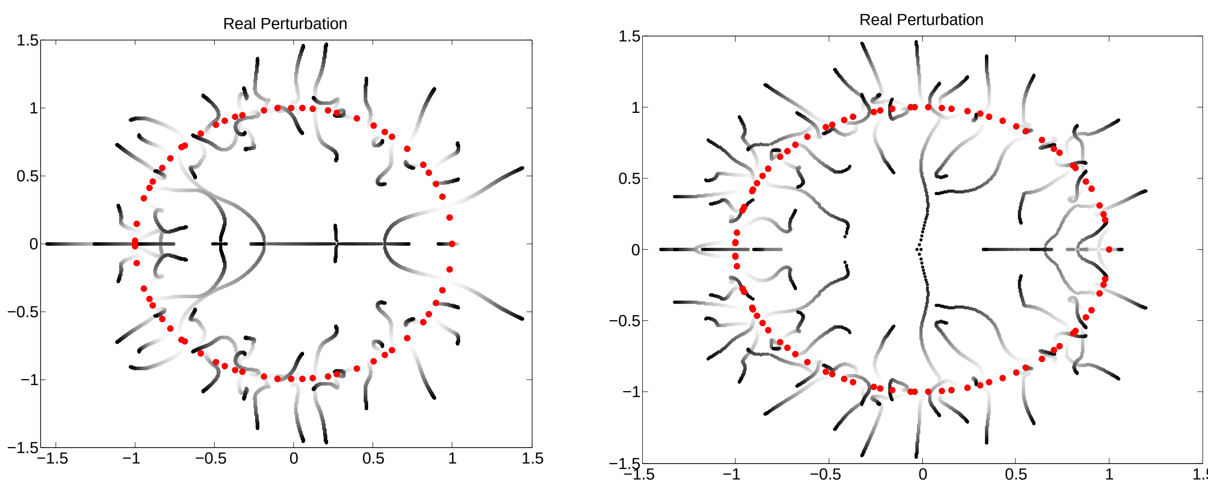


Figure 9: Eigenvalues of $M+tP$, where M is an orthogonal matrix. Left: P is a random ± 1 matrix with norm 2 and final time is $t_{max}=0.74$. Right: P is a real random orthogonal matrix with final time is $t_{max}=1$.

A. Discrete stochastic process

Suppose $M(t)$ is a discrete stochastically varying matrix with $M \equiv M(0)$ being a fixed real $n\times n$ matrix. We discretize time $0=t_0<t_1<\cdots$, and define the evolution of $M(t)$ for any $t_i \le \delta t \le t_{i+1}$ by a piece-wise linear stochastic process

$$M(t_i+\delta t) = M(t_i) + \delta t\, P(t_i) \quad , \tag{31}$$

where $P(t_i)$ is the *impulse* which is a random matrix with independent entries, denoted by p_{jk}, whose expectation values are zero i.e., $\mathbb{E}\left(p_{jk}\right)=0\;\forall j,k$ and $\delta t\in[0,t_{i+1}-t_i)$.

Had we used $\sqrt{\delta t}$ in Eq. 31, and defined the process such that $M(t_i+\delta t)-M(t_i)=\sqrt{\delta t}\,P_i \sim N(0,\delta t)_{\mathbb{R}^n\times\mathbb{R}^n}$, then $M(t)$ would define a *discrete Wiener process,* which is very special type of a stochastic process. The square root of δt is to satisfy the requirement that the variance grows linearly with time. The natural geometry would then be a random walk on the space of $n\times n$ real matrices [5, Chapter 3].

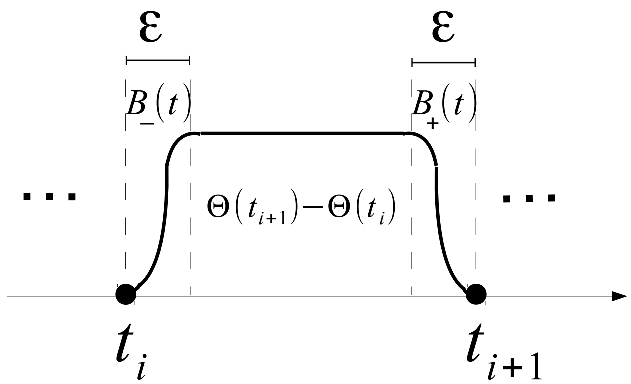


Figure 10: The smooth window function $W_\epsilon\,(t;\ t_i, t_{i+1})$ for $t_i \le t \le t_{i+1}$.

B. Smoothly varying discrete stochastic process

In practice nothing develops infinitely fast and no impulse acts instantly. It is more satisfactory to have a controlled smooth, albeit potentially rapidly changing, formulation of the stochastic impulse. To this end, in what follows we define a smooth version of the stochastic impulse, denoted by $P_\epsilon\,(t)$, that in the limit of $\epsilon \to 0$ becomes Eq. 31. Let

$$P_\epsilon\,(t) = \sum_{i\geq 0} P\,(t_i)\ \ W_\epsilon\,(t;\ t_i, t_{i+1})\,, \tag{32}$$

where each $P\,(t_i)$ is as before and we define the *window* function $W_\epsilon\,(t;\ t_i, t_{i+1})$ to be (see Fig. 10)

$$W_\epsilon\,(t;\ t_i, t_{i+1}) = \begin{cases} \Theta\,(t_{i+1}) - \Theta\,(t_i) & t_i + \epsilon < t < t_{i+1} - \epsilon \\ B_-\,(t\ ;\ t_i, t_{i+1}) & t \le t_i + \epsilon \\ B_+\,(t\ ;\ t_i, t_{i+1}) & t \ge t_{i+1} - \epsilon \end{cases}$$

with $\epsilon < (t_{i+1} - t_i)\,/2$, and Θ being the Heaviside function. B^+ and B^- are the right and left sections of the modified bump function [24] respectively shown in Fig. 10, such that they reach zero at t_i and t_{i+1} and are scaled to have 1 as their maxima (Fig. 10). Mathematically, they are

$$B_-(t\ ;\ t_i, t_{i+1}) = \begin{cases} e^{1-\frac{1}{1-[(t-t_i-\epsilon)/\epsilon]^2}} & t_i \le t \le t_i + \epsilon \\ 0 & \text{otherwise}\,, \end{cases}$$

and

$$B_+(t\ ;\ t_i, t_{i+1}) = \begin{cases} e^{1-\frac{1}{1-[(t-t_{i+1}+\epsilon)/\epsilon]^2}} & t_{i+1} - \epsilon \le t \le t_{i+1} \\ 0 & \text{otherwise}\,. \end{cases}$$

We think of $B_\pm(t\ ;\ t_i, t_{i+1})$ as equations for the *boundary layers* near every t_i. Moreover, the desired independence of time intervals in the discrete stochastic process is guaranteed by the independence of $P\,(t_i)$'s and their confinement to $t_i \le t \le t_{i+1}$ by $W_\epsilon\,(t;\ t_i, t_{i+1})$.

It is easy to check that $P_\epsilon\,(t)$ is differentiable everywhere [4]. More specifically, for $\epsilon > 0$, it is smooth (C^∞) everywhere but on a set of measure zero (i.e., all t_i), where it is only C^1 . We shall need the differentiable property below. With these definitions the smooth version of Eq. 31 reads

---

[4] It is possible to construct C^∞ versions of such window functions such as the Planck-taper window function[24, pp. 127-134]; however, W has a simple form with the properties needed here.

$$\dot{M}_\epsilon(t) = P_\epsilon(t) \quad , \tag{33}$$

with the base case $M \equiv M_\epsilon(0)$ being a fixed real $n \times n$ matrix. $M(t)$, which is not differentiable at any t_i, is recovered by $\lim_{\epsilon \to 0} M_\epsilon(t)$.

Since $P_\epsilon(t)$ is random and differentiable, $\lambda(t)$'s are distinct with probability one. That is $\lambda_i(t)$ has an open neighborhood around it that does not contain any other eigenvalue and can be taken to be a smooth function of t [2]. We shall investigate the eigenvalues of $M(t)$.

The smoothing procedure is not necessary if one is only interested in the spectral properties inside a single interval such as done in Sec. III G. In this case, attraction holds for $t \in (t_i, t_{i+1})$; i.e., outside the boundary layers.

From our derivations leading to Eq. 33, it now becomes obvious that, for small ϵ, inside the boundary layers, i.e., times $(t_i - \epsilon, t_i + \epsilon)$, the dominant force on any eigenvalue is the inertial force $\mathbf{u_i}^*(t)\ \dot{P}_\epsilon(t)\ \mathbf{v_i}(t)$, because

$$\mathbf{u_i}^*(t)\ \dot{P}_\epsilon(t)\ \mathbf{v_i}(t)\ \sim\ 1/\epsilon^2 \qquad t_j < t < t_j + \epsilon \tag{34}$$

$$\mathbf{u_i}^*(t)\ \dot{P}_\epsilon(t)\ \mathbf{v_i}(t)\ \sim\ -1/\epsilon^2 \qquad t_j - \epsilon < t < t_j \quad . \tag{35}$$

For all other times this term is zero and the interaction of the eigenvalues, given by the second term (Eq. 13), governs the force. We will further discuss this and the continuum limit in the next section.

Below to simplify notation, we let $\lambda_i(t) = \lambda_i$, $\mathbf{v_i}(t) = \mathbf{v_i}$, and $\mathbf{u_i}(t) = \mathbf{u_i}$, whereby, Eq. 10 reads

$$\begin{aligned} \ddot{\lambda}_i(t) &= \left(\mathbf{u}_i^*\ \dot{P}_\epsilon(t)\ \mathbf{v_i}\right) + 2\sum_{j \neq i} \frac{c_{ij}\ c_{ji}}{\lambda_i - \lambda_j} && (36) \\ &\doteq (\text{Stochastic Force}) + \sum_{j \neq i} \{\ \text{Force of } \lambda_j \text{ on } \lambda_i\ \}\,, && (37) \end{aligned}$$

where $c_{ij} = \mathbf{u_i^*}\ P_\epsilon(t)\ \mathbf{v_j}$.

Corollary 2. *(**attraction**) Let $M \equiv M_\epsilon(0)$ be a real matrix that evolves according to $\dot{M}_\epsilon(t) = P_\epsilon(t)$, where $P_\epsilon(t)$ is given by Eq. 32. Then for all t, any c.c. pair of eigenvalues of $M_\epsilon(t)$ attract (as in Definition 1). Moreover the expected stochastic force is zero.*

Proof. The attraction immediately follows from the previous proof of attraction for small perturbations. Let us denote $P_\epsilon(t)$ and its components by P_ϵ and $p_{m\ell}$ respectively. The first variation of the eigenvalues in Eq. 7, using index notation reads $\dot{\lambda}_i = \mathbf{u}_i^{*,m}\ p_{m\ell}\ \mathbf{v}_i'^{\ell}$.

We comment that for small $t_i \leq \delta t \leq t_{i+1}$, $\mathbf{u_i}^*$ and $\mathbf{v_i}$ are taken to be eigenvectors of $M(t_i)$ which are independent of $P_\epsilon(t_i + \delta t)$. Hence the right hand sides, in the proof below, are accurate up to $\mathcal{O}(\delta t)$. The first variation is

$$\mathbb{E}\left[\dot{\lambda}_i\right] = \mathbf{u_i}^{*,m}\ \mathbb{E}\left[p_{m\ell}\right]\ \mathbf{v_i}'^{\ell} = 0 \quad , \tag{38}$$

since $\mathbb{E}[p_{m\ell}] = 0$ by assumption. From Eq. 10 we have

$$\mathbb{E}\left[\ddot{\lambda}_i\right] \ = \ \mathbf{u_i}^{*,m}\ \mathbb{E}\left[\left(\dot{P}_\epsilon\right)_{m\ell}\right]\ \mathbf{v_i}'^{\ell} + 2\mathbb{E}\sum_{j \neq i} \frac{\left(\mathbf{u_j^*}\ P_\epsilon\ \mathbf{v_i}\right)\left(\mathbf{u_i^*}\ P_\epsilon\ \mathbf{v_j}\right)}{\lambda_i - \lambda_j} \quad .$$

where $\mathbb{E}\left[\left(\dot{P}_\epsilon\right)_{m\ell}\right] = \sum_i \mathbb{E}\left(P_{m\ell}(t_i)\right) \dot{W}_\epsilon(t;\ t_i, t_{i+1}) = 0$. We have

$$\mathbb{E}\left[\ddot{\lambda}_i\right] \ = \ -i\frac{\mathbb{E}\left[p^2\right]\ \|\mathbf{u_i}^*\|_2^2}{\mathrm{Im}(\lambda_i)} + 2\mathbb{E}\left[p^2\right] \sum_{j \neq \{i, \bar{i}\}} \frac{\left(\mathbf{v_i^T v_j}\right)\left(\mathbf{u_i^*}\overline{\mathbf{u_j}}\right)}{\lambda_i - \lambda_j} \tag{39}$$

□

Remark 7. C.c. eigenvalues and eigenvectors that ultimately become real or those that are initially real and eventually become a c.c. pair must first become equal. Since the motion of the eigenvalues is continuous and the matrix is real, the transition from a c.c. pair to two real eigenvalues or vice versa requires that they first become equal. This is corollary is obvious but perhaps interesting in that the degeneracy of eigenvalues is forced under a generic evolution.

Now suppose we want to define a *continuous stochastic process* where in the equations of motion we first take $\epsilon \to 0$ and then $\delta t \to 0$. The first limit will produce two Dirac delta functions at each t_i whereby $\lim_{t\to t_i^<} \dot{P}(t) = -P(t_{i-1})\,\delta(t_i)$ and $\lim_{t\to t_i^>} \dot{P}(t) = P(t_i)\,\delta(t_i)$; therefore the function is not differentiable at t_i. Lastly, $\delta t \to 0$ will ensure that the stochastic process is nowhere differentiable as one expects from continuous Brownian motion ideas. So what does this mean for eigenvalue attraction? The infinitesimally close delta function impulses dominate the time evolution (Eqs. 34 and 35) with a zero mean force on any eigenvalue.

Remark 8. For a real stochastic process as before the requirement of $M_\epsilon(t)$ being normal for all times demands that $\left[M_\epsilon(t_i) - M_\epsilon^T(t_i), P_\epsilon(t_i)\right] = 0$ which generally is not met for generic $P_\epsilon(t_i)$.

V. FURTHER DISCUSSIONS AND OPEN PROBLEMS

Strongly attracting c.c. pairs ultimately coalesce on the real line and scatter like billiard balls and move about on the real line. Thereafter they can act like "normal" eigenvalues and repel. In particular, the inevitability of collision between an eigenvalue and its c.c., prevents any eigenvalue to cross the real line (change the sign of its imaginary part), without a second encounter.

The motion of the eigenvalues is constrained by the reality of the matrix; the eigenvalue distribution remains symmetric about the real axis. As can be seen in the Figures, an interesting scenario is when two c.c. eigenvalues attract and coalesce on the real line, after which they move in opposite directions on the real line till one of them collides with another (real) eigenvalue. Subsequently, the newly encountered eigenvalue and the one of the original c.c. eigenvalues can momentarily become equal, then move off the real line as a new c.c. pair (see Fig. 8and the plot on the left of Fig. 9 for examples). At times they simply repel each other and remain real.

We emphasize that the proof of c.c. attraction is one of the many forces and at any given instance the net force on any eigenvalue (Eq. 10 and Eq. 36) is the result of the sum of forces of the remaining $n-1$ eigenvalues. In particular, a random collision can take place in the complex plane between eigenvalues that are not c.c.s and deviate the path of an eigenvalue that initially moved towards the real line.

It should be clear that the attraction proved in this work does *not* imply that the long-time behavior is an aggregation of all the eigenvalues on the real line. For a fixed δt, and over long times, the process loses memory of the initial condition (i.e., $M(t=0)$) and ultimately behaves like a random walk on the space of $\mathbb{R}^{n\times n}$ matrices. In fact, Edelman, Kostlan, and Shub [25] showed that for an $n \times n$ real random matrix whose entries are drawn from a standard normal distribution, the expected number of real eigenvalues is $\sqrt{\frac{2n}{\pi}} + O(1)$. Later Tao and Vu [26] proved that matrices whose entries are jointly independent, exponentially decaying, and whose moments match the real gaussian ensemble to fourth order have $\sqrt{\frac{2n}{\pi}} + o\left(\sqrt{n}\right)$ real eigenvalues.

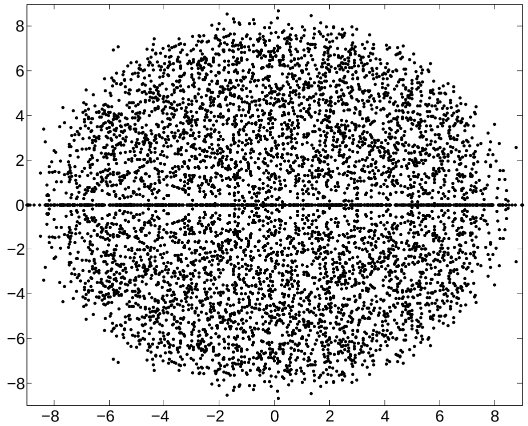
It would be interesting to calculate relaxation times for real deterministic matrices that evolve stochastically and see how long it takes for the matrix to start acting "typical" whereafter the results just mentioned determine the expected behavior.

For any simple eigenvalue λ_i, the condition number is [1, p. 474],

$$\kappa_i = \frac{\|\mathbf{u_i}^*\|\,\|\mathbf{v_i}\|}{|\mathbf{u_i}^*\mathbf{v_i}|} = \|\mathbf{u_i}^*\| \equiv \frac{1}{\left|\cos\theta_0^\ell\right|}$$

by the orthogonality condition (Eq. 6) and the normality (unit length) of $\mathbf{v}_i$. By the Cauchy-Schwartz inequality $\kappa_i \geq 1$. An eigenvalue for which $\kappa_i = 1$ is called a normal eigenvalue and is stable under perturbation, whereas an ill-conditioned eigenvalue has $\kappa_i \gg 1$. The right and left eigenvectors associated to an ill-conditioned eigenvalue can become almost orthogonal implying $\kappa_i \gg 1$ or equivalently $\|\mathbf{u_i}^*\| \gg 1$.

The eigenvalues of normal matrices (e.g., Hermitian, unitary) are very stable under small perturbations. This is not generally the case for non-normal (e.g., Toeplitz) matrices, where small perturbations can change the spectrum significantly [1, 27].



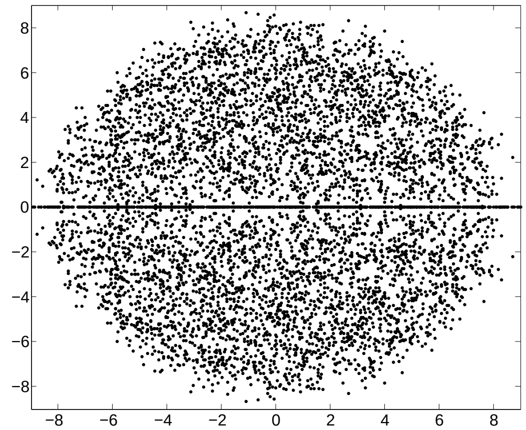


Figure 11: Eigenvalues of 100 instances of 64×64 real matrices. Left: Entries drawn from a standard normal distribution. Right: Entries are random ± 1. Note the aggregation of the eigenvalues on the real axis and their lower nearby density.

Corollary 3. *The expected force of attraction of the c.c. eigenvalues is directly proportional to the square of the* $2-$*norm of the left eigenvector and the variance of the entries of the perturbation matrix.*

This is an immediate consequence of Eq. 39. The numerator in Eq. 39 can change the strength of interaction, most notably because of $\|\mathbf{u_i}^*\|_2^2$, which for non-normal matrices can become quite large. Therefore, c.c. eigenvalues that are distant; i.e., Im (λ_i) is not necessarily small, can attract strongly if they are ill-conditioned.

A. Conjecture: Does the low density of eigenvalues near the real line result from repulsion?

It was previously argued that the relative low density of eigenvalues of real random matrices, seen near the real axis, results from a repulsion of eigenvalues from the real line [28], [29, Section 6.1](see Fig. 11). Preceding [29], Edelman derived the distribution of the eigenvalues for standard normal random matrices and, interestingly, argued that one might think of the real axis as attracting the nearby eigenvalues [30, Section 2 Following Theorem 6.2] (preprint appeared in 1993).

One can conceive of a potentially more complete explanation, where the interaction and dynamics of the eigenvalues take the center stage and not a mysterious interaction with the real axis. To do so, one might need to relate every instance of a real random matrix to the limit of a stochastic process with a base case contained in the deformations of the particular matrix (see below).

In the considerations above, the c.c. eigenvalues of a real matrix move more rapidly towards the real line. Moreover, colliding eigenvalues on the real line, have a large acceleration when they shoot off the real line. This is easily seen from Eq. 13; real eigenvalues have high accelerations as $\lambda_i - \lambda_j$ is small and their subsequent motion is either on the real axis or is into the complex plane.

In either case when there is an imaginary component to the acceleration, its magnitude is quite high. The former corresponds to high accelerations towards the real line which result in the c.c. eigenvalues becoming real. The latter is a strong repulsion away from the real line, shooting the eigenvalues into the complex plane away from the real axis. Therefore, at any given time, on average, one expects a smaller number of eigenvalues to be in the vicinity of the real axis. For large times, when the *equilibrium* is reached, $M(t)$ will have lost the memory of the initial conditions. At all subsequent times, under the stochastic evolution, some eigenvalues become real (because of attraction) and some move off the real line (because of collisions) and on average about $\sqrt{n}$ of the eigenvalues will be found on the real line.

Large forces between nearby eigenvalues is in no way special to the ones with small imaginary components. However, the reality of the matrix causes an anisotropy– the acceleration of the eigenvalues in the imaginary direction becomes larger.

Conjecture 1. *The low density of eigenvalues of real random matrices near the real axis is the result of the large imaginary components of the acceleration into (attraction of c.c. pair) and away from (colliding real eigenvalues) the real axis.*

The relative lack of stability of eigenvalues may explain their aggregation on the real line, as well as, their low nearby density. In order to settle this conjecture, a first step might be to construct any $n \times n$ real random matrix as a limit of a dynamical process like we did above. In particular, deformations of a given random matrix are also random, so in a way, one can conceive of the stochastic process to be the deformations of a matrix in the neighborhood of the random matrix. Then one can relate the expectation of finding an eigenvalue to the expectation of the time it spends anywhere on the complex plane, which, for real random matrices, would be lower in the vicinity of the real axis.

It is our belief that the repulsion of eigenvalues away from the real line is only part of the story in accounting for their relative low density near the real axis.

B. Further opportunities for future work

A list of other open problems includes:

1. What is the probability of collision of eigenvalues on the the real line?[5] The answer to this question would be a helpful step in proving the conjecture above.
2. Estimation of c_{ij} would help quantify the direction and strength of interaction between pairs of eigenvalues.
3. One could give an estimate of the imaginary part of Eq. 30. By doing so one can solve the differential equation to calculate the time it takes for any eigenvalue, λ_k, to reach the real axis, which is the time that λ_k and $\overline{\lambda_k}$ collide and momentarily become degenerate. Indeed, let $u \equiv \operatorname{Im}(\lambda_k)$. The imaginary part of the differential equation (Eq. 30) is of the form $\ddot{u} = f(u)$, which can be solved by first multiplying both sides by $\dot{u}$. That is $\dot{u}\ddot{u} = \frac{1}{2}\frac{d}{dt}\left(\dot{u}\right)^2 = \dot{u}f(u)$ and one has $d\left(\dot{u}\right)^2 = 2f(u)\ du$, hence $\dot{u}(t) = \sqrt{\left(\dot{u}(0)\right)^2 + 2\int f(u)\ du}$, which can be integrated once more to solve for t when $u(t) = 0$.
4. Study of eigenvector localization, especially as they become more real.
5. Toeplitz matrices provide excellent examples of matrices that can be asymmetric and arise in various applications [1]. It would be interesting to better understand the role of the symbol (e.g. its singularity) in connection with this work.
6. Do new features appear in the operator limit?
7. Application of this work in other areas such as open quantum systems [31, 32], PT-symmetric material [12] and biophysics [13].

Acknowledgements

I thank Leo P. Kadanoff, Steven G. Johnson, Tony Iarrobino, Gil Strang, Chris King for comments and the James Franck Institute at University of Chicago and the Perimeter Institute Canada, for having hosted me over the summer of 2013. I acknowledge the National Science Foundation's support through grant DMS. 1312831.

* Electronic address: q.eigenman@gmail.com

[1] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra* (Princeton University Press, 2005).
[2] T. Tao, *Poincaré's legacies: pages from year two of a mathematical blog* (American Mathematical Society, 2009).
[3] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon Press, 1981).
[4] L. Erdös, B. Schlein, and H.-T. Yau, Int Math Res Notices **3**, 436 (2010).
[5] T. Tao, *Topics in Random Matrix Theory (Graduate Studies in Mathematics)* (American Mathematical Society, March 21, 2012).
[6] F. Dyson, J. Mathematical Phys. **3**, 1191 (1962).

[5] This question was posed to us by Freeman Dyson.

[7] N. Hatano and D. R. Nelson, Physical Review B **56**, 8651 (1997).
[8] E. Brézin and A. Zee, Nuclear Physics B **509**, 599 (1998).
[9] P. Brouwer, P. Silvestrov, and C. Beenakker, Physical Review B **56**, 55 (1997).
[10] J. Feinberg and A. Zee, (1997), `arXiv:9706218` .
[11] H. Widom, Operator Theory: Advances and Applications **71**, 1 (1994).
[12] Z. Lin, H. Ramezani, T. Eichelkraut, T. Kottos, H. Cao, and D. N. Christodoulides, Phys. Rev. Lett. **106**, 213901 (2011).
[13] D. R. Nelson, Annu. Rev. Biophys. **41**, 371 (2012).
[14] I. Y. Goldsheid and B. A. Khoruzhenko, Physical Review Letters **80**, 2897 (1998).
[15] J. Bloch, F. Bruckmann, N. Meyer, and S. Schierenberg, Journal of High Energy Physics **8**, 1 (2012).
[16] P. W. Anderson, Physical Review **109**, 1492 (1958).
[17] G. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, 1st ed. (Academic Press, 1990).
[18] P. A. M. Dirac, *The principles of quantum mechanics*, Vol. 4 (Clarendon Press Oxford, 1958).
[19] J. H. Wilkinson, *The algebraic eigenvalue problem*, Vol. 87 (Clarendon Press Oxford, 1965).
[20] T. Tao and V. Vu, Acta mathematica **206**, 127 (2011).
[21] L. N. Trefethen and D. B. III, *Numerical Linear Algebra* (Siam, 1997).
[22] P. Bourgade and H.-T. Yau, arXiv preprint arXiv:1312.1301 (2013).
[23] P. Stoica and R. L. Moses, *Introduction to spectral analysis*, Vol. 1 (Prentice hall Upper Saddle River, 1997).
[24] L. W. Tu, *An Introduction to Manifold* (New York, Springer, 2008).
[25] A. Edelman, E. Kostlan, and M. Shub, J. Amer. Math. Soc. **7**, 247 (1994).
[26] T. Tao and V. Vu, (2012), `arXiv:1206.1893[math.PR]` .
[27] A. Boettcher, M. Embree, and V. I. Sokolov, Math. Comp. **72**, 1329 (2003).
[28] A. Mays, Asia Pacific Mathematics Newsletter **4**, 8 (2014).
[29] A. Edelman and E. Kostlan, Bulletin (new series) of the American Mathematical Society **32** (1995).
[30] A. Edelman, Journal of Multivariate Analysis **60**, 203 (1997).
[31] F. Ticozzi, R. Lucchese, P. Cappellaro, and L. Viola, Automatic Control, IEEE Transactions on **57**, 1931 (2012).
[32] F. Ticozzi and L. Viola, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **370**, 5259 (2012).

VI. APPENDIX: MATLAB CODE

Algorithm 1 Code for plots shown in the paper. The base case can be changed by (un)commenting.

Contents


```
% Ramis Movassagh, Spring 2015
% This code shows the motion of the eigenvalues as described in the paper
clear; n=64;  % Size of the matrices
```

Defining the base case M(0) :

```
%[H dc]=qr(randn(n)); % random orthogonal
% H=randn(n);  % random Gaussian
```

Hatano-Nelson matrix as M(0) :

```
g=0.2;
row=([0 exp(g) zeros(1,n-2)]); col=([0 exp(-g) zeros(1,n-2)]);
H=toeplitz(col,row); H(1,n)=exp(-g); H(n,1)=exp(+g);
```

Additive Randomness

```
D=diag(randn(1,n));  % random diagonal Gaussian
% D=(randn(n));  % random dense Gaussian
D=D/norm(D); % Make norm one
```

Time evolution and motion of eigenvalues in the complex plane

```
p=3; % sets the darkening rate of trajectories
tMax=2; % Maximum time of evolution
dt=0.01; % Time steps
figure(1); plot(eig(H),'r.','MarkerSize',25); hold on; % Unperturbed
for t=0:dt:tMax
    figure(1)
    plot(eig(H+t*D),'k.','MarkerSize',10,'color',...
        [(1-(t/tMax)^p) (1-(t/tMax)^p) (1-(t/tMax)^p)]); hold on;
    title(['The current t is: ',num2str(t)]);pause(0.1);
end
figure(1); plot(eig(H),'r.','MarkerSize',25); hold off;
```