Santa Fe Institute Working Paper 16-07-015
arxiv.org:1607.06526 [math-ph]

Beyond the Spectral Theorem: Spectrally Decomposing Arbitrary Functions of Nondiagonalizable Operators

Paul M. Riechers* and James P. Crutchfield†
Complexity Sciences Center
Department of Physics
University of California at Davis
One Shields Avenue, Davis, CA 95616
(Dated: September 24, 2017)

Nonlinearities in finite dimensions can be linearized by projecting them into infinite dimensions. Unfortunately, the familiar linear operator techniques that one would then use often fail since the operators cannot be diagonalized. This curse is well known. It also occurs for finite-dimensional linear operators. We circumvent it via two tracks. First, using the well-known holomorphic functional calculus, we develop new practical results about spectral projection operators and the relationship between left and right generalized eigenvectors. Second, we generalize the holomorphic calculus to a *meromorphic functional calculus* that can decompose arbitrary functions of nondiagonalizable linear operators in terms of their eigenvalues and projection operators. This simultaneously simplifies and generalizes functional calculus so that it is readily applicable to analyzing complex physical systems. Together, these results extend the spectral theorem of normal operators to a much wider class, including circumstances in which poles and zeros of the function coincide with the operator spectrum. By allowing the direct manipulation of individual eigenspaces of nonnormal and nondiagonalizable operators, the new theory avoids spurious divergences. As such, it yields novel insights and closed-form expressions across several areas of physics in which nondiagonalizable dynamics arise, including memoryful stochastic processes, open nonunitary quantum systems, and far-from-equilibrium thermodynamics.

The technical contributions include the first full treatment of arbitrary powers of an operator, highlighting the special role of the zero eigenvalue. Furthermore, we show that the Drazin inverse, previously only defined axiomatically, can be derived as the negative-one power of singular operators within the meromorphic functional calculus and we give a new general method to construct it. We provide new formulae for constructing spectral projection operators and delineate the relations among projection operators, eigenvectors, and left and right generalized eigenvectors.

By way of illustrating its application, we explore several, rather distinct examples. First, we analyze stochastic transition operators in discrete and continuous time. Second, we show that nondiagonalizability can be a robust feature of a stochastic process, induced even by simple counting. As a result, we directly derive distributions of the time-dependent Poisson process and point out that nondiagonalizability is intrinsic to it and the broad class of hidden semi-Markov processes. Third, we show that the Drazin inverse arises naturally in stochastic thermodynamics and that applying the meromorphic functional calculus provides closed-form solutions for the dynamics of key thermodynamic observables. Fourth, we show that many memoryful processes have power spectra indistinguishable from white noise, despite being highly organized. Nevertheless, whenever the power spectrum is nontrivial, it is a direct signature of the spectrum and projection operators of the process' hidden linear dynamic, with nondiagonalizable subspaces yielding qualitatively distinct line profiles. Finally, we draw connections to the Ruelle–Frobenius–Perron and Koopman operators for chaotic dynamical systems and propose how to extract eigenvalues from a time-series.

PACS numbers: 02.50.-r 05.45.Tp 02.50.Ey 02.50.Ga

... the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience. A. Einstein [1, p. 165]

I. INTRODUCTION

Decomposing a complicated system into its constituent parts—reductionism—is one of science's most powerful strategies for analysis and understanding. Large-scale systems with linearly coupled components give one paradigm of this success. Each can be decomposed into an equivalent system of independent elements using a similarity transformation calculated from the linear algebra of the system's eigenvalues and eigenvectors. The physics of linear wave phenomena, whether of classical

* pmriechers@ucdavis.edu
† chaos@ucdavis.edu

light or quantum mechanical amplitudes, sets the standard of complete reduction rather high. The dynamics is captured by an "operator" whose allowed or exhibited "modes" are the elementary behaviors out of which composite behaviors are constructed by simply weighting each mode's contribution and adding them up.

However, one should not reduce a composite system more than is necessary nor, as is increasingly appreciated these days, more than one, in fact, can. Indeed, we live in a complex, nonlinear world whose constituents are strongly interacting. Often their key structures and memoryful behaviors emerge only over space and time. These are the complex systems. Yet, perhaps surprisingly, many complex systems with nonlinear dynamics correspond to linear operators in abstract high-dimensional spaces [2–4]. And so, there is a sense in which even these complex systems can be reduced to the study of independent nonlocal collective modes.

Reductionism, however, faces its own challenges even within its paradigmatic setting of linear systems: linear operators may have interdependent modes with irreducibly entwined behaviors. These irreducible components correspond to so-called nondiagonalizable subspaces. No similarity transformation can reduce them.

In this view, reductionism can only ever be a guide. The actual goal is to achieve a happy medium, as Einstein reminds us, of decomposing a system only to that level at which the parts are irreducible. To proceed, though, begs the original question, What happens when reductionism fails? To answer this requires revisiting one of its more successful implementations, spectral decomposition of completely reducible operators.

A. Spectral Decomposition

Spectral decomposition—splitting a linear operator into independent modes of simple behavior—has greatly accelerated progress in the physical sciences. The impact stems from the fact that spectral decomposition is not only a powerful mathematical tool for expressing the organization of large-scale systems, but also yields predictive theories with directly observable physical consequences [5]. Quantum mechanics and statistical mechanics identify the energy eigenvalues of Hamiltonians as the basic objects in thermodynamics: transitions among the energy eigenstates yield heat and work. The eigenvalue spectrum reveals itself most directly in other kinds of spectra, such as the frequency spectra of light emitted by the gases that permeate the galactic filaments of our universe [6]. Quantized transitions, an initially mystifying feature of atomic-scale systems, correspond to distinct eigenvectors and discrete spacing between eigenvalues. The corresponding theory of spectral decomposition established the quantitative foundation of quantum mechanics.

The applications and discoveries enabled by spectral decomposition and the corresponding spectral theory fill a long list. In application, direct-bandgap semiconducting materials can be turned into light-emitting diodes (LEDs) or lasers by engineering the spatially-inhomogeneous distribution of energy eigenvalues and the occupation of their corresponding states [7]. Before their experimental discovery, anti-particles were anticipated as the nonoccupancy of negative-energy eigenstates of the Dirac Hamiltonian [8].

The spectral theory, though, extends far beyond physical science disciplines. In large measure, this arises since the evolution of any object corresponds to a linear dynamic in a sufficiently high-dimensional state space. Even nominally nonlinear dynamics over several variables, the canonical mechanism of deterministic chaos, appear as linear dynamics in appropriate infinite-dimensional shift-spaces [4]. A nondynamic version of rendering nonlinearities into linearities in a higher-dimensional feature space is exploited with much success today in machine learning by support vector machines, for example [9]. Spectral decomposition often allows a problem to be simplified by approximations that use only the dominant contributing modes. Indeed, human-face recognition can be efficiently accomplished using a small basis of "eigenfaces" [10].

Certainly, there are many applications that highlight the importance of decomposition and the spectral theory of operators. However, a brief reflection on the mathematical history will give better context to its precise results, associated assumptions, and, more to the point, the generalizations we develop here in hopes of advancing the analysis and understanding of complex systems.

Following on early developments of operator theory by Hilbert and co-workers [11], the *spectral theorem for normal operators* reached maturity under von Neumann by the early 1930s [12, 13]. It became the mathematical backbone of much progress in physics since then, from classical partial differential equations to quantum physics. Normal operators, by definition, commute with their Hermitian conjugate: $A^\dagger A = AA^\dagger$. Examples include symmetric and orthogonal matrices in classical mechanics and Hermitian, skew-Hermitian, and unitary operators in quantum mechanics.

The spectral theorem itself is often identified as a collection of related results about normal operators; see, e.g., Ref. [14]. In the case of finite-dimensional vector spaces [15], the spectral theorem asserts that normal operators are diagonalizable and can always be diagonalized by a unitary transformation; that left and right eigenvec-

tors (or eigenfunctions) are simply related by complex-conjugate transpose; that these eigenvectors form a complete basis; and that functions of a normal operator reduce to the action of the function on each eigenvalue. Most of these qualities survive with only moderate provisos in the infinite-dimensional case. In short, the spectral theorem makes physics governed by normal operators tractable.

The spectral theorem, though, appears powerless when faced with nonnormal and nondiagonalizable operators. What then are we to do when confronted by, say, complex interconnected systems with nonunitary time evolution, by open systems, by structures that emerge on space and time scales different from the equations of motion, or by other novel physics governed by nonnormal and not-necessarily-diagonalizable operators? Where is the comparably constructive framework for calculations beyond the standard spectral theorem? Fortunately, portions of the necessary generalization have been made within pure mathematics [16], some finding applications in engineering and control [17, 18]. However, what is available is incomplete. And, even that which is available is often not in a form adapted to perform calculations that lead to quantitative predictions.

B. Synopsis

Here, we build on previous work in functional analysis and operator theory to provide both a rigorous and constructive foundation for physically relevant calculations involving not-necessarily-diagonalizable operators. In effect, we extend the spectral theorem for normal operators to a broader setting, allowing generalized "modes" of nondiagonalizable systems to be identified and manipulated. The meromorphic functional calculus we develop extends Taylor series expansion and standard holomorphic functional calculus to analyze arbitrary functions of not-necessarily-diagonalizable operators. It readily handles singularities arising when poles (or zeros) of the function coincide with poles of the operator's resolvent—poles that appear precisely at the operator's eigenvalues. Pole–pole and pole–zero interactions substantially modify the complex-analytic residues within the functional calculus. A key result is that the negative-one power of a singular operator exists in the meromorphic functional calculus. It is the *Drazin inverse*, a powerful tool that is receiving increased attention in stochastic thermodynamics and elsewhere. Furthermore, we derive consequences from the more familiar holomorphic functional calculus that readily allow spectral decomposition of nondiagonalizable operators in terms of spectral projections and left and right generalized eigenvectors—decanting the abstract mathematical theory into a more tractable framework for analyzing complex physical systems.

Taken together, the functional calculus, Drazin inverse, and methods to manipulate particular eigenspaces, are key to a thorough-going analysis of many complex systems, many now accessible for the first time. Indeed, the framework has already been fruitfully employed in several specific applications, including closed-form expressions for signal processing and information measures of hidden Markov processes [19–23], compressing stochastic processes over a quantum channel [24, 25], and stochastic thermodynamics [26, 27]. However, the techniques are sufficiently general they will be much more widely useful. We envision new opportunities for similar detailed analyses, ranging from biophysics to quantum field theory, wherever restrictions to normal operators and diagonalizability have been roadblocks.

With this broad scope in mind, we develop the mathematical theory first without reference to specific applications and disciplinary terminology. We later give pedagogical (yet, we hope, interesting) examples, exploring several niche, but important applications to finite hidden Markov processes, basic stochastic process theory, nonequilibrium thermodynamics, signal processing, and nonlinear dynamical systems. At a minimum, the examples and their breadth serve to better acquaint readers with the basic methods required to employ the theory.

We introduce the meromorphic functional calculus in §III through §IV, after necessary preparation in §II. §V A further explores and gives a new formula for eigenprojectors, which we refer to here simply as *projection operators*. §V B makes explicit their general relationship with eigenvectors and generalized eigenvectors and clarifies the orthonormality relationship among left and right generalized eigenvectors. §V B 4 then discusses simplifications of the functional calculus for special cases, while §VI A takes up the spectral properties of transition operators. The examples are discussed at length in §VI before we close in §VII with suggestions on future applications and research directions.

II. SPECTRAL PRIMER

The following is relatively self-contained, assuming basic familiarity with linear algebra at the level of Refs. [15, 17]—including eigen-decomposition and knowledge of the Jordan canonical form, partial fraction expansion (see Ref. [28]), and series expansion—and basic knowledge of complex analysis—including the residue theorem and calculation of residues at the level of Ref. [29]. For those lacking a working facility with these concepts, a quick review of §VI's applications may motivate reviewing them.

In this section, we introduce our notation and, in doing so, remind the reader of certain basic concepts in linear algebra and complex analysis that will be used extensively in the following.

To begin, we restrict attention to operators with finite representations and only sometimes do we take the limit of dimension going to infinity. That is, we do not consider infinite-rank operators outright. While this runs counter to previous presentations in mathematical physics that consider only infinite-dimensional operators, the upshot is that they—as limiting operators—can be fully treated with a countable point spectrum. We present examples of this later on. Accordingly, we restrict our attention to operators with at most a countably infinite spectrum. Such operators share many features with finite-dimensional square matrices, and so we recall several elementary but essential facts from matrix theory used repeatedly in the main development.

If A is a finite-dimensional square matrix, then its *spectrum* is simply the set Λ_A of its eigenvalues:

$$\Lambda_A = \left\{ \lambda \in \mathbb{C} : \det(\lambda I - A) = 0 \right\} ,$$

where $\det(\cdot)$ is the determinant of its argument and I is the identity matrix. The *algebraic multiplicity* a_λ of eigenvalue λ is the power of the term $(z - \lambda)$ in the characteristic polynomial $\det(zI - A)$. In contrast, the *geometric multiplicity* g_λ is the dimension of the kernel of the transformation $A - \lambda I$ or, equivalently, the number of linearly independent eigenvectors associated with the eigenvalue. The algebraic and geometric multiplicities are all equal when the matrix is diagonalizable.

Since there can be multiple subspaces associated with a single eigenvalue, corresponding to different Jordan blocks in the Jordan canonical form, it is structurally important to distinguish the *index* of the eigenvalue associated with the largest of these subspaces [30].

Definition 1. *Eigenvalue λ's* index ν_λ *is the size of the largest Jordan block associated with λ.*

If $z \notin \Lambda_A$, then $\nu_z = 0$. Note that the index of the operator A itself is sometimes discussed [31]. In such contexts, the index of A is ν_0. Hence, ν_λ corresponds to the index of $A - \lambda I$.

The index of an eigenvalue gives information beyond what the algebraic and geometric multiplicities themselves yield. Nevertheless, for $\lambda \in \Lambda_A$, it is always true that $\nu_\lambda - 1 \leq a_\lambda - g_\lambda \leq a_\lambda - 1$. In the diagonalizable case, $a_\lambda = g_\lambda$ and $\nu_\lambda = 1$ for all $\lambda \in \Lambda_A$.

The following employs basic features of complex analysis extensively in conjunction with linear algebra. Let us therefore review several elementary notions in complex analysis. Recall that a *holomorphic function* is one that is complex differentiable throughout the domain under consideration. A *pole* of order n at $z_0 \in \mathbb{C}$ is a singularity that behaves as $h(z)/(z - z_0)^n$ as $z \to z_0$, where $h(z)$ is holomorphic within a neighborhood of z_0 and $h(z_0) \neq 0$. We say that $h(z)$ has a *zero* of order m at z_1 if $1/h(z)$ has a pole of order m at z_1. A *meromorphic function* is one that is holomorphic except possibly at a set of isolated poles within the domain under consideration.

Defined over the continuous complex variable $z \in \mathbb{C}$, A's *resolvent*:

$$R(z; A) \equiv (zI - A)^{-1} ,$$

captures all of A's spectral information through the poles of $R(z; A)$'s matrix elements. In fact, the resolvent contains more than just A's spectrum: we later show that the order of each pole gives the index ν of the corresponding eigenvalue.

The spectrum Λ_A can be expressed in terms of the resolvent. Explicitly, the *point spectrum* (i.e., the set of eigenvalues) is the set of complex values z at which $zI - A$ is not a one-to-one mapping, with the implication that the inverse of $zI - A$ does not exist:

$$\Lambda_A = \left\{ \lambda \in \mathbb{C} : R(\lambda; A) \neq \mathrm{inv}(\lambda I - A) \right\} ,$$

where $\mathrm{inv}(\cdot)$ is the inverse of its argument. Later, via our investigation of the Drazin inverse, it should become clear that the resolvent operator can be self-consistently defined at the spectrum, despite the lack of an inverse.

For infinite-rank operators, the spectrum becomes more complicated. In that case, the *right point spectrum* (the point spectrum of A) need not be the same as the *left point spectrum* (the point spectrum of A's dual $A^\top$). Moreover, the spectrum may grow to include non-eigenvalues z for which the range of $zI - A$ is not dense in the vector space it transforms or for which $zI - A$ has dense range but the inverse of $zI - A$ is not bounded. These two settings give rise to the so-called *residual spectrum* and *continuous spectrum*, respectively [32]. To mitigate confusion, it should be noted that the point spectrum can be continuous, yet never coincides with the continuous spectrum just described. Moreover, understanding only countable point spectra is necessary to follow the developments here.

Each of A's eigenvalues λ has an associated *projection operator* A_λ, which is the *residue* of the resolvent as $z \to \lambda$ [14]. Explicitly:

$$A_\lambda = \mathrm{Res}\big((zI - A)^{-1}, z \to \lambda \big) ,$$

where $\mathrm{Res}(\,\cdot\,, z \to \lambda)$ is the element-wise residue of its first argument as $z \to \lambda$. The projection operators are

orthonormal:

$$A_\lambda A_\zeta = \delta_{\lambda,\zeta} A_\lambda \ . \tag{1}$$

and sum to the identity:

$$I = \sum_{\lambda \in \Lambda_A} A_\lambda \ . \tag{2}$$

The following discusses in detail and then derives several new properties of projection operators.

III. FUNCTIONAL CALCULI

In the following, we develop an extended *functional calculus* that makes sense of arbitrary functions $f(\cdot)$ of a linear operator A. Within any functional calculus, one considers how A's eigenvalues map to the eigenvalues of $f(A)$; which we call a *spectral mapping*. For example, it is known that holomorphic functions of bounded linear operators enjoy an especially simple spectral mapping theorem [33]:

$$\Lambda_{f(A)} = f(\Lambda_A) \ .$$

To fully appreciate the meromorphic functional calculus, we first state and compare the main features and limitations of alternative functional calculi.

A. Taylor series

Inspired by the Taylor expansion of scalar functions:

$$f(a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\xi)}{n!} (a - \xi)^n \ ,$$

a calculus for functions of an operator A can be based on the series:

$$f(A) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\xi)}{n!} (A - \xi I)^n \ , \tag{3}$$

where $f^{(n)}(\xi)$ is the n^{th} derivative of $f(z)$ evaluated at $z = \xi$.

This is often used, for example, to express the exponential of A as:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \ .$$

This particular series-expansion is convergent for any A since e^z is *entire*, in the sense of complex analysis. Unfortunately, even if it exists there is a limited domain of convergence for most functions. For example, suppose $f(z)$ has poles and choose a Maclaurin series; i.e., $\xi = 0$ in Eq. (3). Then the series only converges when A's spectral radius is less than the radius of the innermost pole of $f(z)$. Addressing this and related issues leads directly to alternative functional calculi.

B. Holomorphic functional calculus

Holomorphic functions are well behaved, smooth functions that are complex differentiable. Given a function $f(\cdot)$ that is holomorphic within a disk enclosed by a counterclockwise contour C, its Cauchy integral formula is given by:

$$f(a) = \frac{1}{2\pi i} \oint_C f(z) \, (z - a)^{-1} \, dz \ , \tag{4}$$

Taking this as inspiration, the holomorphic functional calculus performs a contour integration of the resolvent to extend $f(\cdot)$ to operators:

$$f(A) = \frac{1}{2\pi i} \oint_{C_{\Lambda_A}} f(z) \, (zI - A)^{-1} \, dz \ , \tag{5}$$

where C_{Λ_A} is a closed counterclockwise contour that encompasses Λ_A. Assuming that $f(z)$ is holomorphic at $z = \lambda$ for all $\lambda \in \Lambda_A$, a nontrivial calculation [30] shows that Eq. (5) is equivalent to the holomorphic calculus defined by:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda - 1} \frac{f^{(m)}(\lambda)}{m!} (A - \lambda I)^m A_\lambda \ . \tag{6}$$

After some necessary development, we will later derive Eq. (6) as a special case of our meromorphic functional calculus, such that Eq. (6) is valid whenever $f(z)$ is holomorphic at $z = \lambda$ for all $\lambda \in \Lambda_A$.

The holomorphic functional calculus was first proposed in Ref. [30] and is now in wide use; e.g., see Ref. [17, p. 603]. It agrees with the Taylor-series approach whenever the infinite series converges, but gives a functional calculus when the series approach fails. For example, using the principal branch of the complex logarithm, the holomorphic functional calculus admits $\log(A)$ for any nonsingular matrix, with the satisfying result that $e^{\log(A)} = A$. Whereas, the Taylor series approach fails to converge for the logarithm of most matrices even if the expansion for, say, $\log(1 - z)$ is used.

The major shortcoming of the holomorphic functional calculus is that it assumes $f(z)$ is holomorphic at Λ_A. Clearly, if $f(z)$ has a pole at some $z \in \Lambda_A$, then Eq. (6) fails. An example of such a failure is the negative-one

power of a singular operator, which we take up later on.

Several efforts have been made to extend the holomorphic functional calculus. For example, Refs. [34] and [35] define a functional calculus that extends the standard holomorphic functional calculus to include a certain class of meromorphic functions that are nevertheless still required to be *holomorphic on the point spectrum* (i.e., on the eigenvalues) of the operator. However, we are not aware of any previous work that introduces and develops the consequences of a functional calculus for functions that are meromorphic on the point spectrum—which we take up in the next few sections.

C. Meromorphic functional calculus

Meromorphic functions are holomorphic except at a set of isolated poles of the function. The resolvent of a finite-dimensional operator is meromorphic, since it is holomorphic everywhere except for poles at the eigenvalues of the operator. We will now also allow our function $f(z)$ to be meromorphic with possible poles that coincide with the poles of the resolvent.

Inspired again by the Cauchy integral formula of Eq. (4), but removing the restriction to holomorphic functions, our meromorphic functional calculus instead employs a partitioned contour integration of the resolvent:

$$f(A) = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} f(z) R(z; A)\, dz \ ,$$

where C_λ is a small counterclockwise contour around the eigenvalue λ. This and a spectral decomposition of the resolvent (to be derived later) extends the holomorphic calculus to a much wider domain, defining:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda - 1} A_\lambda \big(A - \lambda I\big)^m \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z-\lambda)^{m+1}}\, dz \ . \tag{7}$$

The contour is integrated using knowledge of $f(z)$ since meromorphic $f(z)$ can introduce poles and zeros at Λ_A that interact with the resolvent's poles.

The meromorphic functional calculus agrees with the Taylor-series approach whenever the series converges and agrees with the holomorphic functional calculus whenever $f(z)$ is holomorphic at Λ_A. However, when both the previous functional calculi fail, the meromorphic calculus extends the domain of $f(A)$ to yield surprising, yet sensible answers. For example, we show that within it, the negative-one power of a singular operator is the Drazin inverse—an operator that effectively inverts everything that is invertible.

The major assumption of our meromorphic functional calculus is that the domain of operators must have a spectrum that is at most countably infinite—e.g., A can be any compact operator. A related limitation is that singularities of $f(z)$ that coincide with Λ_A must be isolated singularities. Nevertheless, we expect that these restrictions can be lifted with proper treatment, as discussed in fuller context later.

IV. MEROMORPHIC SPECTRAL DECOMPOSITION

The preceding gave an overview of the relationship between alternative functional calculi and their trade-offs, highlighting the advantages of the meromorphic functional calculus. This section leverages these advantages and employs a partial fraction expansion of the resolvent to give a general spectral decomposition of almost any function of any operator. Then, since it plays a key role in applications, we apply the functional calculus to investigate the negative-one power of singular operators—thus *deriving*, what is otherwise an operator defined axiomatically, the Drazin inverse from first principles.

A. Partial fraction expansion of the resolvent

The elements of A's resolvent are proper rational functions that contain all of A's spectral information. (Recall that a *proper rational function* $r(z)$ is a ratio of polynomials in z whose numerator has degree strictly less than the degree of the denominator.) In particular, the resolvent's poles coincide with A's eigenvalues since, for $z \notin \Lambda_A$:

$$\begin{aligned} R(z; A) &= (zI - A)^{-1} \\ &= \frac{\mathcal{C}^\top}{\det(zI - A)} \\ &= \frac{\mathcal{C}^\top}{\prod_{\lambda \in \Lambda_A} (z-\lambda)^{a_\lambda}} \ , \end{aligned} \tag{8}$$

where a_λ is the algebraic multiplicity of eigenvalue λ and $\mathcal{C}$ is the matrix of *cofactors* of $zI - A$. That is, $\mathcal{C}$'s transpose $\mathcal{C}^\top$ is the *adjugate* of $zI - A$:

$$\mathcal{C}^\top = \mathrm{adj}(zI - A) \ ,$$

whose elements will be polynomial functions of z of degree less than $\sum_{\lambda \in \Lambda_A} a_\lambda$.

Recall that the partial fraction expansion of a proper rational function $r(z)$ with poles in Λ allows a unique decomposition into a sum of constant numerators divided

by monomials in $z-\lambda$ up to degree a_λ, when a_λ is the order of the pole of $r(z)$ at $\lambda \in \Lambda$ [28]. Equation (8) thus makes it clear that the resolvent has the unique partial fraction expansion:

$$R(z;A) = \sum_{\lambda\in\Lambda_A} \sum_{m=0}^{a_\lambda-1} \frac{1}{(z-\lambda)^{m+1}} A_{\lambda,m} \ , \tag{9}$$

where $\{A_{\lambda,m}\}$ is the set of matrices with constant entries (*not* functions of z) uniquely determined elementwise by the partial fraction expansion. However, $R(z;A)$'s poles are *not* necessarily of the same order as the algebraic multiplicity of the corresponding eigenvalues since the entries of $\mathcal{C}$, and thus of $\mathcal{C}^\top$, may have zeros at A's eigenvalues. This has the potential to render $A_{\lambda,m}$ equal to the zero matrix $\mathbf{0}$.

The Cauchy integral formula indicates that the constant matrices $\{A_{\lambda,m}\}$ of Eq. (9) can be obtained as the residues:

$$A_{\lambda,m} = \frac{1}{2\pi i} \oint_{C_\lambda} (z-\lambda)^m R(z;A) dz \ , \tag{10}$$

where the residues are calculated elementwise. The projection operators A_λ associated with each eigenvalue λ were already referenced in §II, but can now be properly introduced as the $A_{\lambda,0}$ matrices:

$$\begin{aligned} A_\lambda &= A_{\lambda,0} && (11)\\ &= \frac{1}{2\pi i} \oint_{C_\lambda} R(z;A) dz \ . && (12) \end{aligned}$$

Since $R(z;A)$'s elements are rational functions, as we just showed, it is analytic except at a finite number of isolated singularities—at A's eigenvalues. In light of the residue theorem, this motivates the Cauchy-integral-like formula that serves as the starting point for the meromorphic functional calculus:

$$f(A) = \sum_{\lambda\in\Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} f(z) R(z;A) dz \ . \tag{13}$$

Let's now consider several immediate consequences.

B. Decomposing the identity

Even the simplest applications of Eq. (13) yield insight. Consider the identity as the operator function $f(A) = A^0 = I$ that corresponds to the scalar function $f(z) = z^0 = 1$. Then, Eq. (13) implies:

$$\begin{aligned} I &= \sum_{\lambda\in\Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} R(z;A) dz \\ &= \sum_{\lambda\in\Lambda_A} A_\lambda \ . \end{aligned}$$

This shows that the projection operators are, in fact, a decomposition of the identity, as anticipated in Eq. (2).

C. Dunford decomposition, decomposed

For $f(A) = A$, Eqs. (13) and (10) imply that:

$$\begin{aligned} A &= \sum_{\lambda\in\Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} z R(z;A) dz \\ &= \sum_{\lambda\in\Lambda_A} \left[\lambda \tfrac{1}{2\pi i} \oint_{C_\lambda} R(z;A) dz + \tfrac{1}{2\pi i} \oint_{C_\lambda} (z-\lambda) R(z;A) dz \right] \\ &= \sum_{\lambda\in\Lambda_A} \left(\lambda A_{\lambda,0} + A_{\lambda,1} \right) \ . \end{aligned} \tag{14}$$

We denote the important set of nilpotent matrices $A_{\lambda,1}$ that project onto the generalized eigenspaces by relabeling them:

$$\begin{aligned} N_\lambda &\equiv A_{\lambda,1} && (15)\\ &= \frac{1}{2\pi i} \oint_{C_\lambda} (z-\lambda) R(z;A) dz \ . && (16) \end{aligned}$$

Equation (14) is the unique *Dunford decomposition* [16]: $A = D + N$, where $D \equiv \sum_{\lambda\in\Lambda_A} \lambda A_\lambda$ is diagonalizable, $N \equiv \sum_{\lambda\in\Lambda_A} N_\lambda$ is nilpotent, and D and N commute: $[D,N] = \mathbf{0}$. This is also known as the *Jordan–Chevalley decomposition*.

The special case where A is diagonalizable implies that $N = \mathbf{0}$. And so, Eq. (14) simplifies to:

$$A = \sum_{\lambda\in\Lambda_A} \lambda A_\lambda \ .$$

D. The resolvent, resolved

As shown in Ref. [14] and can be derived from Eqs. (12) and (16):

$$\begin{aligned} A_\lambda A_\zeta &= \delta_{\lambda,\zeta} A_\lambda \text{ and} \\ A_\lambda N_\zeta &= \delta_{\lambda,\zeta} N_\lambda \ . \end{aligned}$$

Due to these, our spectral decomposition of the Dunford decomposition implies that:

$$\begin{aligned} N_\lambda &= A_\lambda\Big(A - \sum_{\zeta\in\Lambda_A} \zeta A_\zeta\Big) \\ &= A_\lambda\big(A - \lambda A_\lambda\big) \\ &= A_\lambda\big(A - \lambda I\big) \ . \end{aligned} \tag{17}$$

Moreover:

$$A_{\lambda,m} = A_\lambda\big(A - \lambda I\big)^m \ . \tag{18}$$

It turns out that for $m > 0$: $A_{\lambda,m} = N_\lambda^m$. (See also Ref. [14, p. 483].) This leads to a generalization of the projection operator orthonormality relations of Eq. (1). Most generally, the operators of $\{A_{\lambda,m}\}$ are mutually related by:

$$A_{\lambda,m}A_{\zeta,n} = \delta_{\lambda,\zeta}A_{\lambda,m+n} \ . \tag{19}$$

Finally, if we recall that the index ν_λ is the dimension of the largest associated subspace, we find that the index of λ characterizes the nilpotency of N_λ: $N_\lambda^m = \mathbf{0}$ for $m \geq \nu_\lambda$. That is:

$$A_{\lambda,m} = \mathbf{0} \qquad \text{for } m \geq \nu_\lambda \ . \tag{20}$$

Returning to Eq. (9), we see that all $A_{\lambda,m}$ with $m \geq \nu_\lambda$ are zero-matrices and so do not contribute to the sum. Thus, we can rewrite Eq. (9) as:

$$R(z;A) = \sum_{\lambda\in\Lambda_A}\sum_{m=0}^{\nu_\lambda-1} \frac{1}{(z-\lambda)^{m+1}} A_{\lambda,m} \tag{21}$$

or:

$$R(z;A) = \sum_{\lambda\in\Lambda_A}\sum_{m=0}^{\nu_\lambda-1} \frac{1}{(z-\lambda)^{m+1}} A_\lambda\big(A - \lambda I\big)^m \ , \tag{22}$$

for $z \notin \Lambda_A$.

The following sections sometimes use $A_{\lambda,m}$ in place of $A_\lambda\big(A - \lambda I\big)^m$. This is helpful both for conciseness and when applying Eq. (19). Nonetheless, the equality in Eq. (18) is a useful one to keep in mind.

E. Meromorphic functional calculus

In light of Eq. (13), Eq. (21) together with Eq. (18) allow us to express any function of an operator simply and solely in terms of its spectrum (i.e., its eigenvalues for the finite dimensional case), its projection operators, and itself:

$$f(A) = \sum_{\lambda\in\Lambda_A}\sum_{m=0}^{\nu_\lambda-1} A_{\lambda,m}\,\frac{1}{2\pi i}\oint_{C_\lambda} \frac{f(z)}{(z-\lambda)^{m+1}}\,dz \ . \tag{23}$$

In obtaining Eq. (23) we finally derived Eq. (7), as promised earlier in § III C. Effectively, by modulating the modes associated with the resolvent's singularities, the scalar function $f(\cdot)$ is mapped to the operator domain, where its action is expressed in each of A's independent subspaces.

F. Evaluating the residues

Interpretation aside, how does one use this result? Equation (23) says that the spectral decomposition of $f(A)$ reduces to the evaluation of several residues, where:

$$\mathrm{Res}\big(g(z),\ z\to\lambda\big) = \frac{1}{2\pi i}\oint_{C_\lambda} g(z)\,dz \ .$$

So, to make progress with Eq. (23), we must evaluate function-dependent residues of the form:

$$\mathrm{Res}\left(f(z)/(z-\lambda)^{m+1},\, z\to\lambda\right) \ .$$

If $f(z)$ were holomorphic at each λ, then the order of the pole would simply be the power of the denominator. We could then use Cauchy's differential formula for holomorphic functions:

$$f^{(n)}(a) = \frac{n!}{2\pi i}\oint_{C_a} \frac{f(z)}{(z-a)^{n+1}}dz \ , \tag{24}$$

for $f(z)$ holomorphic at a. And, the meromorphic calculus would reduce to the holomorphic calculus. Often, $f(z)$ will be holomorphic at least at *some* of A's eigenvalues. And so, Eq. (24) is still locally a useful simplification in those special cases.

In general, though, $f(z)$ introduces poles and zeros at $\lambda \in \Lambda_A$ that change their orders. This is exactly the impetus for the generalized functional calculus. The residue of a complex-valued function $g(z)$ around its isolated pole λ of order $n+1$ can be calculated from:

$$\mathrm{Res}\big(g(z),\ z\to\lambda\big) = \frac{1}{n!}\lim_{z\to\lambda}\frac{d^n}{dz^n}\left[(z-\lambda)^{n+1}g(z)\right] \ .$$

G. Decomposing A^L

Equation (23) says that we can explicitly derive the spectral decomposition of powers of the operator A. Of course, we already did this for the special cases of A^0 and

A^1. The goal, though, is to do this in general.

For $f(A) = A^L \to f(z) = z^L$, $z = 0$ can be either a zero or a pole of $f(z)$, depending on the value of L. In either case, an eigenvalue of $\lambda = 0$ will distinguish itself in the residue calculation of A^L via its unique ability to change the order of the pole (or zero) at $z = 0$. For example, at this special value of λ and for integer $L > 0$, $\lambda = 0$ induces poles that *cancel* with the zeros of $f(z) = z^L$, since z^L has a zero at $z = 0$ of order L. For integer $L < 0$, an eigenvalue of $\lambda = 0$ *increases* the order of the $z = 0$ pole of $f(z) = z^L$. For all other eigenvalues, the residues will be as expected. Hence, from Eq. (23) and inserting $f(z) = z^L$, for any $L \in \mathbb{C}$:

$$A^L = \left[\sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda - 1} A_\lambda \big(A - \lambda I\big)^m \overbrace{\left(\frac{1}{2\pi i} \oint_{C_\lambda} \frac{z^L}{(z-\lambda)^{m+1}}\, dz \right)}^{= \frac{1}{m!} \lim_{z \to \lambda} \frac{d^m}{dz^m} z^L = \frac{\lambda^{L-m}}{m!} \prod_{n=1}^{m} (L-n+1)} \right] + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} A_0 A^m \underbrace{\left(\frac{1}{2\pi i} \oint_{C_0} z^{L-m-1}\, dz \right)}_{= \delta_{L,m}}$$

$$= \left[\sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda \big(A - \lambda I\big)^m \right] + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m \ , \tag{25}$$

where $\binom{L}{m}$ is the generalized binomial coefficient:

$$\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1) \quad \text{with } \binom{L}{0} = 1 \ , \tag{26}$$

and $[0 \in \Lambda_A]$ is the Iverson bracket which takes on value 1 if zero is an eigenvalue of A and 0 if not. $A_{\lambda,m}$ was replaced by $A_\lambda (A - \lambda I)^m$ to suggest the more explicit calculations involved with evaluating any A^L. Equation (25) applies to any linear operator with only isolated singularities in its resolvent.

The eigen-decomposition of polynomials implied by Eq. (25) makes the contribution of the zero eigenvalue more explicit than previous treatments and enables closed-form expressions, e.g., for correlation functions, where the zero eigenvalue makes a qualitatively distinct contribution. Consequentially, this formulation can lead to the recognition of coexistent finite and infinite range physical phenomena of different mechanistic origin.

If L is a nonnegative integer such that $L \geq \nu_\lambda - 1$ for all $\lambda \in \Lambda_A$, then:

$$A^L = \sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_{\lambda,m} \ , \tag{27}$$

where $\binom{L}{m}$ is now reduced to the traditional binomial coefficient $L!/(m!(L-m)!)$.

H. Drazin inverse

If L is any negative integer, then $\binom{-|L|}{m}$ can be written as a traditional binomial coefficient $(-1)^m \binom{|L|+m-1}{m}$, yielding:

$$A^{-|L|} = \sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda - 1} (-1)^m \tbinom{|L|+m-1}{m} \lambda^{-|L|-m} A_{\lambda,m} \ , \tag{28}$$

for $-|L| \in \{-1, -2, -3, \dots\}$.

Thus, negative powers of an operator can be consistently defined even for noninvertible operators. In light of Eqs. (25) and (28), it appears that the zero eigenvalue does not even contribute to the function. It is well known, in contrast, that it wreaks havoc on the naive, oft-quoted definition of a matrix's negative power:

$$A^{-1} \stackrel{?}{=} \frac{\mathrm{adj}(A)}{\det(A)} = \frac{\mathrm{adj}(A)}{\prod_{\lambda \in \Lambda_A} \lambda^{a_\lambda}} \ ,$$

since this would imply dividing by zero. If we can accept large positive powers of singular matrices—for which the zero eigenvalue does not contribute—it seems fair to also accept negative powers that likewise involve no contribution from the zero eigenvalue.

Editorializing aside, we note that extending the definition of A^{-1} to the domain including singular operators via Eqs. (25) and (28) implies that:

$$\begin{aligned} A^{|L|} A^{-|\ell|} &= A^{-|\ell|} A^{|L|} \\ &= A^{|L|-|\ell|} \qquad \text{for } |L| \geq |\ell| + \nu_0 \ , \end{aligned}$$

which is a very sensible and desirable condition. Moreover, we find that $AA^{-1} = I - A_0$.

Specifically, the negative-one power of any square matrix is in general *not* the same as the matrix inverse since $\mathrm{inv}(A)$ need not exist. However, it is consistently defined via Eq. (28) to be:

$$A^{-1} = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} (-1)^m \lambda^{-1-m} A_{\lambda,m} \ . \tag{29}$$

This is the *Drazin inverse* $A^{\mathcal{D}}$ of A. Note that it is *not* the same as the Moore–Penrose pseudo-inverse [36, 37].

Although the Drazin inverse is usually defined axiomatically to satisfy certain criteria [38], it is naturally *derived* as the negative one power of a singular operator in the meromorphic functional calculus. We can check that it indeed satisfies the axiomatic criteria for the Drazin inverse, enumerated according to historical precedent:

$$\begin{aligned} &(1^{\nu_0}) && A^{\nu_0} A^{\mathcal{D}} A = A^{\nu_0} \\ &(2) && A^{\mathcal{D}} A A^{\mathcal{D}} = A^{\mathcal{D}} \\ &(5) && [A, A^{\mathcal{D}}] = 0 \ , \end{aligned}$$

which gives rise to the Drazin inverse's moniker as the $\{1^{\nu_0}, 2, 5\}$-inverse [38]. The analytical form of Eq. (29) has been teased out previously by other means; see, e.g., Ref. [38] and for other settings see Refs. [39, 40]. Nevertheless, due to its utility in application, it is noteworthy and appealing that the Drazin inverse falls out organically in the meromorphic functional calculus, as the negative-one power, in contrast to its otherwise rather esoteric axiomatic origin.

While A^{-1} always exists, the resolvent is nonanalytic at $z = 0$ for a singular matrix. Effectively, the meromorphic functional calculus removes the nonanalyticity of the resolvent in evaluating A^{-1}. As a result, as we can see from Eq. (29), the Drazin inverse inverts what is invertible; the remainder is zeroed out.

Of course, whenever A *is* invertible, A^{-1} is equal to $\mathrm{inv}(A)$. However, we should not confuse this coincidence with equivalence. Moreover, despite historic notation there is no reason that the negative-one power should in general be equivalent to the inverse. Especially, if an operator is not invertible! To avoid confusing A^{-1} with $\mathrm{inv}(A)$, we use the notation $A^{\mathcal{D}}$ for the Drazin inverse of A. Still, $A^{\mathcal{D}} = \mathrm{inv}(A)$, whenever $0 \notin \Lambda_A$.

Amusingly, this extension of previous calculi lets us resolve an elementary but fundamental question: What is 0^{-1}? It is certainly not infinity. Indeed, it is just as close to negative infinity! Rather: $0^{-1} = 0 \neq \mathrm{inv}(0)$.

Although Eq. (29) is a constructive way to build the Drazin inverse, it imposes more work than is actually necessary. Using the meromorphic functional calculus, we can derive a new, simple construction of the Drazin inverse that requires only the original operator and the eigenvalue-0 projector.

First, assume that λ is an isolated singularity of $R(z; A)$ with finite separation at least ϵ distance from the nearest neighboring singularity. And, consider the operator-valued function f^ϵ_λ defined via the RHS of:

$$\begin{aligned} A_\lambda &= f^\epsilon_\lambda(A) \\ &= \tfrac{1}{2\pi i} \oint_{\lambda + \epsilon e^{i\phi}} (\zeta I - A)^{-1} \, d\zeta \ , \end{aligned}$$

with $\lambda + \epsilon e^{i\phi}$ defining an ϵ-radius circular contour around λ. Then we see that:

$$\begin{aligned} f^\epsilon_\lambda(z) &= \tfrac{1}{2\pi i} \oint_{\lambda + \epsilon e^{i\phi}} (\zeta - z)^{-1} \, d\zeta \\ &= \big[z \in \mathbb{C} : |z - \lambda| < \epsilon \big] \ , \end{aligned} \tag{30}$$

where $[z \in \mathbb{C} : |z - \lambda| < \epsilon]$ is the Iverson bracket that takes on value 1 if z is within ϵ-distance of λ and 0 if not.

Second, we use this to find that, for any $c \in \mathbb{C} \setminus \{0\}$:

$$\begin{aligned} (A + cA_0)^{-1} &= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda - 1} A_{\lambda,m} \tfrac{1}{2\pi i} \oint_{C_\lambda} \frac{\big(z + c f^\epsilon_0(z)\big)^{-1}}{(z - \lambda)^{m+1}} \, dz \\ &= A^{\mathcal{D}} + \sum_{m=0}^{\nu_0 - 1} A_0 A^m \tfrac{1}{2\pi i} \oint_{C_0} \frac{(z + c)^{-1}}{z^{m+1}} \\ &= A^{\mathcal{D}} + \sum_{m=0}^{\nu_0 - 1} A_0 A^m (-1)^m / c^{m+1} \ , \end{aligned} \tag{31}$$

where we asserted that the contour C_0 exists within the finite ϵ-ball about the origin.

Third, we note that $A + cA_0$ is invertible for all $c \neq 0$; this can be proven by multiplying each side of Eq. (31) by $A + cA_0$. Hence, $(A + cA_0)^{-1} = \mathrm{inv}(A + cA_0)$ for all $c \neq 0$.

Finally, multiplying each side of Eq. (31) by $I - A_0$, and recalling that $A_{0,0} A_{0,m} = A_{0,m}$, we find a useful expression for calculating the Drazin inverse of any linear operator A, given only A and A_0. Specifically:

$$A^{\mathcal{D}} = (I - A_0)(A + cA_0)^{-1} \ . \tag{32}$$

which is valid for any $c \in \mathbb{C} \setminus \{0\}$. Eq. (32) generalizes the result found specifically for $c = -1$ in Ref. [41].

For the special case of $c = -1$, it is worthwhile to also consider the alternative construction of the Drazin

inverse implied by Eq. (31):

$$A^{\mathcal{D}} = (A - A_0)^{-1} + A_0 \Big(\sum_{m=0}^{\nu_0 - 1} A^m \Big) \ . \tag{33}$$

By a spectral mapping ($\lambda \to 1 - \lambda$, for $\lambda \in \Lambda_T$), the Perron–Frobenius theorem and Eq. (31) yield an important consequence for any stochastic matrix T. The Perron–Frobenius theorem guarantees that T's eigenvalues along the unit circle are associated with a diagonalizable subspace. In particular, $\nu_1 = 1$. Spectral mapping of this result means that T's eigenvalue 1 maps to the eigenvalue 0 of $I - T$ and $T_1 = (I - T)_0$. Moreover:

$$[(I - T) + T_1]^{-1} = (I - T)^{\mathcal{D}} + T_1 \ ,$$

since $\nu_0 = 1$. This corollary of Eq. (31) (with $c = 1$) corresponds to a number of important and well known results in the theory of Markov processes. Indeed, $Z \equiv (I - T + T_1)^{-1}$ is called the *fundamental matrix* in that setting [42].

I. Consequences and generalizations

For an infinite-rank operator A with a continuous spectrum, the meromorphic functional calculus has the natural generalization:

$$f(A) = \frac{1}{2\pi i} \oint_{C_{\Lambda_A}} f(z)(zI - A)^{-1} \, dz \ , \tag{34}$$

where the contour C_{Λ_A} encloses the (possibly continuous) spectrum of A without including any unbounded contributions from $f(z)$ outside of C_{Λ_A}. The function $f(z)$ is expected to be meromorphic within C_{Λ_A}. This again deviates from the holomorphic approach, since the holomorphic functional calculus requires that $f(z)$ is analytic in a neighborhood around the spectrum; see § VII of Ref. [43]. Moreover, Eq. (34) allows an extension of the functional calculus of Refs. [34, 35, 44], since the function can be meromorphic at the point spectrum in addition being meromorphic on the residual and continuous spectra.

In either the finite- or infinite-rank case, whenever $f(z)$ *is* analytic in a neighborhood around the spectrum, the meromorphic functional calculus agrees with the holomorphic. Whenever $f(z)$ is *not* analytic in a neighborhood around the spectrum, the function is undefined in the holomorphic approach. In contrast, the meromorphic approach extends the function to the operator-valued domain and does so with novel consequences.

In particular, when $f(z)$ is *not analytic in a neighborhood around the spectrum*—say $f(z)$ is nonanalytic within A's spectrum at $\Xi_f \subset \Lambda_A$—then we expect to lose both homomorphism and spectral mapping properties:

- Loss of homomorphism: $f_1(A) f_2(A) \neq (f_1 \cdot f_2)(A)$;
- Loss of naive spectral mapping: $f(\Lambda_A \setminus \Xi_f) \subset \Lambda_{f(A)}$.

A simple example of both losses arises with the Drazin inverse, above. There, $f_1(z) = z^{-1}$. Taking this and $f_2(z) = z$ combined with singular operator A leads to the loss of homomorphism: $A^{\mathcal{D}} A \neq I$. As for the second property, the spectral mapping can be altered for the candidate spectra at Ξ_f via pole–pole or pole–zero interactions in the complex contour integral. For $f(A) = A^{-1}$, how does A's eigenvalue of 0 get mapped into the new spectrum of $A^{\mathcal{D}}$? A naive application of the spectral mapping theorem might seem to yield an undefined quantity. But, using the meromorphic functional calculus self-consistently maps the eigenvalue as $0^{-1} = 0$. It remains to be explored whether the full spectral mapping is preserved for any function $f(A)$ under the meromorphic interpretation of $f(\lambda)$.

It should now be apparent that extending functions via the meromorphic functional calculus allows one to express novel mathematical properties, some likely capable of describing new physical phenomena. At the same time, extra care is necessary. The situation is reminiscent of the loss of commutativity in non-Abelian operator algebra: not all of the old rules apply, but the gain in nuance allows for mathematical description of important phenomena.

We chose to focus primarily on the finite-rank case here since it is sufficient to demonstrate the utility of the general projection-operator formalism. Indeed, there are ample nontrivial applications in the finite-rank setting that deserve attention. To appreciate these, we now turn to address the construction and properties of general eigenprojectors.

V. CONSTRUCTING DECOMPOSITIONS

At this point, we see that projection operators are fundamental to functions of an operator. This prompts the practical question of how to actually calculate them. The next several sections address this by deriving expressions with both theoretical and applied use. We first develop the projection operators associated with index-one eigenvalues. We then explicate the relationship between eigenvectors, generalized eigenvectors, and projection operators for normal, diagonalizable, and general matrices. Finally, we show how the general results specialize in several common cases of interest. After these, we turn to examples and applications.

A. Projection operators of index-one eigenvalues

To obtain the projection operators associated with each index-one eigenvalue $\lambda \in \{\zeta \in \Lambda_A : \nu_\zeta = 1\}$, we apply the functional calculus to an appropriately chosen function of A, finding:

$$\begin{aligned}\prod_{\substack{\zeta\in\Lambda_A\\ \zeta\neq\lambda}} (A-\zeta I)^{\nu_\zeta} &= \sum_{\xi\in\Lambda_A} \sum_{m=0}^{\nu_\xi-1} \frac{A_{\xi,m}}{2\pi i} \oint_{C_\xi} \frac{\prod_{\substack{\zeta\in\Lambda_A\\ \zeta\neq\lambda}} (z-\zeta)^{\nu_\zeta}}{(z-\xi)^{m+1}}\, dz \\ &= A_\lambda \frac{1}{2\pi i} \oint_{C_\lambda} \frac{\prod_{\substack{\zeta\in\Lambda_A\\ \zeta\neq\lambda}} (z-\zeta)^{\nu_\zeta}}{z-\lambda}\, dz \\ &= A_\lambda \prod_{\substack{\zeta\in\Lambda_A\\ \zeta\neq\lambda}} (\lambda-\zeta)^{\nu_\zeta} \ .\end{aligned}$$

Therefore, if $\nu_\lambda = 1$:

$$A_\lambda = \prod_{\substack{\zeta\in\Lambda_A\\ \zeta\neq\lambda}} \left(\frac{A-\zeta I}{\lambda-\zeta}\right)^{\nu_\zeta} \ . \tag{35}$$

As convenience dictates in our computations, we let $\nu_\zeta \to a_\zeta - g_\zeta + 1$ or even $\nu_\zeta \to a_\zeta$ in Eq. (35), since multiplying A_λ by $(A-\zeta I)/(\lambda-\zeta)$ has no effect for $\zeta \in \Lambda_A \setminus \{\lambda\}$ if $\nu_\lambda = 1$.

Equation (35) generalizes a well known result that applies when the index of *all* eigenvalues is one. That is, when the operator is diagonalizable, we have:

$$A_\lambda = \prod_{\substack{\zeta\in\Lambda_A\\ \zeta\neq\lambda}} \frac{A-\zeta I}{\lambda-\zeta} \ .$$

To the best of our knowledge, Eq. (35) is original.

Since eigenvalues can have index larger than one, not all projection operators of a nondiagonalizable operator can be found directly from Eq. (35). Even so, it serves three useful purposes. First, it gives a practical reduction of the eigen-analysis by finding all projection operators of index-one eigenvalues. Second, if there is only one eigenvalue that has index larger than one—what we call the *almost diagonalizable case*—then Eq. (35), together with the fact that the projection operators must sum to the identity, *does* give a full solution to the set of projection operators. Third, Eq. (35) is a powerful theoretical tool that we can use directly to spectrally decompose functions, for example, of a stochastic matrix whose eigenvalues on the unit circle are guaranteed to be index-one by the Perron–Frobenius theorem.

Although index-one expressions have some utility, we need a more general procedure to obtain all projection operators of any linear operator. Recall that, with full generality, projection operators can also be calculated directly via residues, as in Eq. (12).

An alternative procedure—one that extends a method familiar at least in quantum mechanics—is to obtain the projection operators via eigenvectors. However, quantum mechanics always concerns itself with a subset of diagonalizable operators. What is the necessary generalization? For one, left and right eigenvectors are no longer simply conjugate transposes of each other. More severely, a full set of spanning eigenvectors is no longer guaranteed and we must resort to *generalized* eigenvectors. Since the relationships among eigenvectors, generalized eigenvectors, and projection operators are critical to the practical calculation of many physical observables of complex systems, we collect these results in the next section.

B. Eigenvectors, generalized eigenvectors, and projection operators

Two common questions regarding projection operators are: Why not just use eigenvectors? And, why not use the Jordan canonical form? First, the eigenvectors of a defective matrix do not form a complete basis with which to expand an arbitrary vector. One needs generalized eigenvectors for this. Second, some functions of an operator require removing, or otherwise altering, the contribution from select eigenspaces. This is most adroitly handled with the projection operator formalism where different eigenspaces (correlates of Jordan blocks) can effectively be treated separately. Moreover, even for simple cases where eigenvectors suffice, the projection operator formalism simply can be more calculationally or mathematically convenient.

That said, it is useful to understand the relationship between projection operators and generalized eigenvectors. For example, it is often useful to create projection operators from generalized eigenvectors. This section clarifies their connection using the language of matrices. In the most general case, we show that the projection operator formalism is usefully concise.

1. *Normal matrices*

Unitary, Hermitian, skew-Hermitian, orthogonal, symmetric, and skew-symmetric matrices are all special cases of normal matrices. As noted, normal matrices are those that commute with their Hermitian adjoint (complex-conjugate transpose): $AA^\dagger = A^\dagger A$. Moreover, a matrix is normal if and only if it can be diagonalized by a unitary transformation: $A = U\Lambda U^\dagger$, where the columns of the unitary matrix U are the orthonormal right eigenvectors

of A corresponding to the eigenvalues ordered along the diagonal matrix Λ. For an M-by-M matrix A, the eigenvalues in Λ_A are ordered and enumerated according to the possibly degenerate M-tuple $(\Lambda_A) = (\lambda_1, \ldots, \lambda_M)$. Since an eigenvalue $\lambda \in \Lambda_A$ has algebraic multiplicity $a_\lambda \geq 1$, λ appears a_λ times in the ordered tuple.

Assuming A is normal, each projection operator A_λ can be constructed as the sum of all ket–bra pairs of right-eigenvectors corresponding to λ composed with their conjugate transpose. We later introduce bras and kets more generally via generalized eigenvectors of the operator A and its dual $A^\top$. However, since the complex-conjugate transposition rule between dual spaces is only applicable to a ket basis derived from a normal operator, we put off using the bra-ket notation for now so as not to confuse the more familiar "normal" case with the general case.

To explicitly demonstrate this relationship between projection operators, eigenvectors, and their Hermitian adjoints in the case of normality, observe that:

$$\begin{aligned}
A &= U\Lambda U^\dagger \\
&= \begin{bmatrix}\vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_M\end{bmatrix} \begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M\end{bmatrix} \begin{bmatrix}\vec{u}_1^\dagger \\ \vec{u}_2^\dagger \\ \vdots \\ \vec{u}_M^\dagger\end{bmatrix} \\
&= \begin{bmatrix}\lambda_1\vec{u}_1 & \lambda_2\vec{u}_2 & \cdots & \lambda_M\vec{u}_M\end{bmatrix} \begin{bmatrix}\vec{u}_1^\dagger \\ \vec{u}_2^\dagger \\ \vdots \\ \vec{u}_M^\dagger\end{bmatrix} \\
&= \sum_{j=1}^{M} \lambda_j \vec{u}_j \vec{u}_j^\dagger \\
&= \sum_{\lambda\in\Lambda_A} \lambda A_\lambda \ .
\end{aligned}$$

Evidently, for normal matrices A:

$$A_\lambda = \sum_{j=1}^{M} \delta_{\lambda,\lambda_j} \vec{u}_j \vec{u}_j^\dagger \ .$$

And, since $\vec{u}_i^\dagger \vec{u}_j = \delta_{i,j}$, we have an orthogonal set $\{A_\lambda\}_{\lambda\in\Lambda_A}$ with the property that:

$$\begin{aligned}
A_\zeta A_\lambda &= \sum_{i=1}^{M}\sum_{j=1}^{M} \delta_{\zeta,\lambda_i}\delta_{\lambda,\lambda_j} \vec{u}_i\vec{u}_i^\dagger \vec{u}_j\vec{u}_j^\dagger \\
&= \sum_{i=1}^{M}\sum_{j=1}^{M} \delta_{\zeta,\lambda_i}\delta_{\lambda,\lambda_j} \vec{u}_i\delta_{i,j}\vec{u}_j^\dagger \\
&= \sum_{i=1}^{M} \delta_{\zeta,\lambda_i}\delta_{\lambda,\lambda_i} \vec{u}_i\vec{u}_i^\dagger \\
&= \delta_{\zeta,\lambda} A_\lambda \ .
\end{aligned}$$

Moreover:

$$\begin{aligned}
\sum_{\lambda\in\Lambda_A} A_\lambda &= \sum_{j=1}^{M} \vec{u}_j\vec{u}_j^\dagger \\
&= UU^\dagger \\
&= I \ ,
\end{aligned}$$

and so on. All of the expected properties of projection operators can be established again in this restricted setting.

The rows of $U^{-1} = U^\dagger$ are A's left eigenvectors. In this case, they are simply the conjugate transpose of the right eigenvectors. Note that conjugate transposition is the familiar transformation rule between ket and bra spaces in quantum mechanics (see, e.g., Ref. [45])—a consequence of the restriction to normal operators, as we will show. Importantly, a more general formulation of quantum mechanics would *not* have this same restricted correspondence between the dual ket and bra spaces.

To elaborate on this point, recall that vector spaces admit dual spaces and dual bases. However, there is no sense of a dual correspondence of a single ket or bra without reference to a full basis [15]. Implicitly in quantum mechanics, the basis is taken to be the basis of eigenstates of any Hermitian operator, nominally since observables are self-adjoint.

To allude to an alternative, we note that $\vec{u}_j^\dagger\vec{u}_j$ is not only the Hermitian form of inner product $\langle\vec{u}_j, \vec{u}_j\rangle$ (where $\langle\cdot,\cdot\rangle$ denotes the inner product) of the right eigenvector $\vec{u}_j$ with itself, but importantly also the simple dot-product of the left eigenvector $\vec{u}_j^\dagger$ and the right eigenvector $\vec{u}_j$, where $\vec{u}_j^\dagger$ acts as a linear functional on $\vec{u}_j$. Contrary to the substantial effort devoted to the inner-product-centric theory of Hilbert spaces, this latter interpretation of $\vec{u}_j^\dagger\vec{u}_j$—in terms of linear functionals and a left-eigenvector basis for linear functionals—is what generalizes to a consistent and constructive framework for the spectral theory beyond normal operators, as we will see shortly.

2. Diagonalizable matrices

By definition, diagonalizable matrices can be diagonalized, but not necessarily via a unitary transformation. All diagonalizable matrices can nevertheless be diagonalized via the transformation: $A = P\Lambda P^{-1}$, where the columns of the square matrix P are the not-necessarily-orthogonal right eigenvectors of A corresponding to the eigenvalues ordered along the diagonal matrix Λ and where the rows of P^{-1} are A's left eigenvectors. Importantly, the left eigenvectors need not be the Hermitian adjoint of the right eigenvectors. As a particular example, this more general setting is required for almost any transition dynamic of a Markov chain. In other words, the transition dynamic of any interesting complex network with irreversible processes serves as an example of a nonnormal operator.

Given the M-tuple of possibly-degenerate eigenvalues $(\Lambda_A) = (\lambda_1, \lambda_2, \ldots, \lambda_M)$, there is a corresponding M-tuple of linearly-independent right-eigenvectors $(|\lambda_1\rangle, |\lambda_2\rangle, \ldots, |\lambda_M\rangle)$ and a corresponding M-tuple of linearly-independent left-eigenvectors $(\langle\lambda_1|, \langle\lambda_2|, \ldots, \langle\lambda_M|)$ such that:

$$A\,|\lambda_j\rangle = \lambda_j\,|\lambda_j\rangle$$

and:

$$\langle\lambda_j|\,A = \lambda_j\,\langle\lambda_j|$$

with the orthonormality condition that:

$$\langle\lambda_i|\lambda_j\rangle = \delta_{i,j}\ .$$

To avoid misinterpretation, we stress that the bras and kets that appear above are the left and right eigenvectors, respectively, and typically do *not* correspond to complex-conjugate transposition.

With these definitions in place, the projection operators for a diagonalizable matrix can be written:

$$A_\lambda = \sum_{j=1}^{M} \delta_{\lambda,\lambda_j}\,|\lambda_j\rangle\,\langle\lambda_j|\ .$$

Then:

$$\begin{aligned}
A &= \sum_{\lambda\in\Lambda_A} \lambda A_\lambda \\
&= \sum_{j=1}^{M} \lambda_j\,|\lambda_j\rangle\,\langle\lambda_j| \\
&= \begin{bmatrix}\lambda_1\,|\lambda_1\rangle & \lambda_2\,|\lambda_2\rangle & \cdots & \lambda_M\,|\lambda_M\rangle\end{bmatrix} \begin{bmatrix}\langle\lambda_1| \\ \langle\lambda_2| \\ \vdots \\ \langle\lambda_M|\end{bmatrix} \\
&= \begin{bmatrix}|\lambda_1\rangle & |\lambda_2\rangle & \cdots & |\lambda_M\rangle\end{bmatrix} \begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M\end{bmatrix} \begin{bmatrix}\langle\lambda_1| \\ \langle\lambda_2| \\ \vdots \\ \langle\lambda_M|\end{bmatrix} \\
&= P\Lambda P^{-1}\ .
\end{aligned}$$

So, we see that the projection operators introduced earlier in a coordinate-free manner have a concrete representation in terms of left and right eigenvectors when the operator is diagonalizable.

3. Any matrix

Not all matrices can be diagonalized, but all square matrices can be put into *Jordan canonical form* via the transformation: $A = YJY^{-1}$ [17]. Here, the columns of the square matrix Y are the linearly independent right eigenvectors and generalized right eigenvectors corresponding to the Jordan blocks ordered along the diagonal of the block-diagonal matrix J. And, the rows of Y^{-1} are the corresponding left eigenvectors and generalized left eigenvectors, but reverse-ordered within each block, as we will show.

Let there be n Jordan blocks forming the n-tuple $(J_1, J_2, \ldots, J_n)$, with $1 \le n \le M$. The k^{th} Jordan block J_k has dimension m_k-by-m_k:

$$J_k = \underbrace{\left.\begin{bmatrix}\lambda_k & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda_k & 1 & & & 0 & 0 \\ & 0 & \lambda_k & & & & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & & & & \lambda_k & 1 & 0 \\ 0 & 0 & & & 0 & \lambda_k & 1 \\ 0 & 0 & 0 & \cdots & & 0 & \lambda_k\end{bmatrix}\right\}}_{m_k \text{ columns}} \quad m_k \text{ rows}$$

such that:

$$\sum_{k=1}^{n} m_k = M \ .$$

Note that eigenvalue $\lambda \in \Lambda_A$ corresponds to g_λ different Jordan blocks, where g_λ is the geometric multiplicity of the eigenvalue λ. Indeed:

$$n = \sum_{\lambda \in \Lambda_A} g_\lambda \ .$$

Moreover, the index ν_λ of the eigenvalue λ is defined as the size of the largest Jordan block corresponding to λ. So, we write this in the current notation as:

$$\nu_\lambda = \max\{\delta_{\lambda,\lambda_k} m_k\}_{k=1}^{n} \ .$$

If the index of any eigenvalue is greater than one, then the conventional eigenvectors do not span the M-dimensional vector space. However, the set of M generalized eigenvectors does form a basis for the vector space [46].

Given the n-tuple of possibly-degenerate eigenvalues $(\Lambda_A) = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, there is a corresponding n-tuple of m_k-tuples of linearly-independent generalized right-eigenvectors:

$$\left((|\lambda_1^{(m)}\rangle)_{m=1}^{m_1}, (|\lambda_2^{(m)}\rangle)_{m=1}^{m_2}, \ldots, (|\lambda_n^{(m)}\rangle)_{m=1}^{m_n} \right) \ ,$$

where:

$$\left(|\lambda_k^{(m)}\rangle \right)_{m=1}^{m_k} \equiv \left(|\lambda_k^{(1)}\rangle, |\lambda_k^{(2)}\rangle, \ldots, |\lambda_k^{(m_k)}\rangle \right)$$

and a corresponding n-tuple of m_k-tuples of linearly-independent generalized left-eigenvectors:

$$\left((\langle\lambda_1^{(m)}|)_{m=1}^{m_1}, (\langle\lambda_2^{(m)}|)_{m=1}^{m_2}, \ldots, (\langle\lambda_n^{(m)}|)_{m=1}^{m_n} \right) \ ,$$

where:

$$\left(\langle\lambda_k^{(m)}| \right)_{m=1}^{m_k} \equiv \left(\langle\lambda_k^{(1)}|, \langle\lambda_k^{(2)}|, \ldots, \langle\lambda_k^{(m_k)}| \right)$$

such that:

$$(A - \lambda_k I)\, |\lambda_k^{(m+1)}\rangle = |\lambda_k^{(m)}\rangle \tag{36}$$

and:

$$\langle\lambda_k^{(m+1)}|\, (A - \lambda_k I) = \langle\lambda_k^{(m)}| \ , \tag{37}$$

for $0 \leq m \leq m_k - 1$, where $|\lambda_j^{(0)}\rangle = \vec{0}$ and $\langle\lambda_j^{(0)}| = \vec{0}$. Specifically, $|\lambda_k^{(1)}\rangle$ and $\langle\lambda_k^{(1)}|$ are conventional right and left eigenvectors, respectively.

Most directly, the generalized right and left eigenvectors can be found as the nontrivial solutions to:

$$(A - \lambda_k I)^m\, |\lambda_k^{(m)}\rangle = \vec{0}$$

and:

$$\langle\lambda_k^{(m)}|\, (A - \lambda_k I)^m = \vec{0} \ ,$$

respectively.

It should be clear from Eq. (36) and Eq. (37) that:

$$\begin{aligned} \langle\lambda_k^{(m)}|(A - \lambda_k I)^\ell|\lambda_k^{(n)}\rangle &= \langle\lambda_k^{(m-\ell)}|\lambda_k^{(n)}\rangle \\ &= \langle\lambda_k^{(m)}|\lambda_k^{(n-\ell)}\rangle \ , \end{aligned}$$

for $m, n, \in \{0, 1, \ldots, m_k\}$ and $\ell \geq 0$. At the same time, it is then easy to show that:

$$\langle\lambda_k^{(m)}|\lambda_k^{(n)}\rangle = \langle\lambda_k^{(m+n)}|\lambda_k^{(0)}\rangle = 0, \quad \text{if } m + n \leq m_k \ ,$$

where $m, n \in \{0, 1, \ldots, m_k\}$. Imposing appropriate normalization, we find that:

$$\langle\lambda_j^{(m)}|\lambda_k^{(n)}\rangle = \delta_{j,k}\delta_{m+n,m_k+1} \ . \tag{38}$$

Hence, we see that the left eigenvectors and generalized eigenvectors are a dual basis to the right eigenvectors and generalized eigenvectors. Interestingly though, within each Jordan subspace, *the most generalized left eigenvectors are dual to the least generalized right eigenvectors*, and vice versa.

(To be clear, in this terminology "least generalized" eigenvectors are the standard eigenvectors. For example, the $\langle\lambda_k^{(1)}|$ satisfying the standard eigenvector relation $\langle\lambda_k^{(1)}|\, A = \lambda_k\, \langle\lambda_k^{(1)}|$ is the least generalized left eigenvector of subspace k. By way of comparison, the "most generalized"' right eigenvector of subspace k is $|\lambda_k^{(m_k)}\rangle$ satisfying the most generalized eigenvector relation $(A - \lambda_k I)\, |\lambda_k^{(m_k)}\rangle = |\lambda_k^{(m_k-1)}\rangle$ for subspace k. The orthonormality relation shows that the two are dual correspondents: $\langle\lambda_k^{(1)}|\lambda_k^{(m_k)}\rangle = 1$, while all other eigen-bra–eigen-ket closures utilizing these objects are null.)

With these details worked out, we find that the projection operators for a nondiagonalizable matrix can be written as:

$$A_\lambda = \sum_{k=1}^{n} \sum_{m=1}^{m_k} \delta_{\lambda,\lambda_k}\, |\lambda_k^{(m)}\rangle\, \langle\lambda_k^{(m_k+1-m)}| \ . \tag{39}$$

And, we see that a projection operator includes all of its left and right eigenvectors and all of its left and right generalized eigenvectors. This implies that the identity operator must also have a decomposition in terms of both

eigenvectors and generalized eigenvectors:

$$\begin{aligned} I &= \sum_{\lambda\in\Lambda_A} A_\lambda \\ &= \sum_{k=1}^{n}\sum_{m=1}^{m_k} |\lambda_k^{(m)}\rangle\, \langle\lambda_k^{(m_k+1-m)}| \ . \end{aligned}$$

Let $\left[|\lambda_k^{(m)}\rangle\right]_{m=1}^{m_k}$ denote the column vector:

$$\left[|\lambda_k^{(m)}\rangle\right]_{m=1}^{m_k} = \begin{bmatrix} |\lambda_k^{(1)}\rangle \\ \vdots \\ |\lambda_k^{(m_k)}\rangle \end{bmatrix} \ ,$$

and let $\left[\langle\lambda_k^{(m_k+1-m)}|\right]_{m=1}^{m_k}$ denote the column vector:

$$\left[\langle\lambda_k^{(m_k+1-m)}|\right]_{m=1}^{m_k} = \begin{bmatrix} \langle\lambda_k^{(m_k)}| \\ \vdots \\ \langle\lambda_k^{(1)}| \end{bmatrix} \ .$$

Then, using the above results, and the fact that Eq. (37) implies that $\langle\lambda_k^{(m+1)}|\, A = \lambda_k\, \langle\lambda_k^{(m+1)}| + \langle\lambda_k^{(m)}|$, we derive the explicit generalized-eigenvector decomposition of the nondiagonalizable operator A:

$$\begin{aligned} A &= \big(\sum_{\lambda\in\Lambda_A} A_\lambda\big) A \\ &= \sum_{k=1}^{n}\sum_{m=1}^{m_k} |\lambda_k^{(m)}\rangle\, \langle\lambda_k^{(m_k+1-m)}|\, A \\ &= \sum_{k=1}^{n}\sum_{m=1}^{m_k} |\lambda_k^{(m)}\rangle \left(\lambda_k\, \langle\lambda_k^{(m_k+1-m)}| + \langle\lambda_k^{(m_k-m)}|\right) \\ &= \begin{bmatrix} \left[|\lambda_1^{(m)}\rangle\right]_{m=1}^{m_1} \\ \left[|\lambda_2^{(m)}\rangle\right]_{m=1}^{m_2} \\ \vdots \\ \left[|\lambda_n^{(m)}\rangle\right]_{m=1}^{m_n} \end{bmatrix}^\top \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_n \end{bmatrix} \begin{bmatrix} \left[\langle\lambda_1^{(m_1+1-m)}|\right]_{m=1}^{m_1} \\ \left[\langle\lambda_2^{(m_2+1-m)}|\right]_{m=1}^{m_2} \\ \vdots \\ \left[\langle\lambda_n^{(m_n+1-m)}|\right]_{m=1}^{m_n} \end{bmatrix} \\ &= YJY^{-1} \ , \end{aligned}$$

where, defining Y as:

$$Y = \begin{bmatrix} \left[|\lambda_1^{(m)}\rangle\right]_{m=1}^{m_1} \\ \left[|\lambda_2^{(m)}\rangle\right]_{m=1}^{m_2} \\ \vdots \\ \left[|\lambda_n^{(m)}\rangle\right]_{m=1}^{m_n} \end{bmatrix}^\top \ ,$$

we are forced by Eq. (38) to recognize that:

$$Y^{-1} = \begin{bmatrix} \left[\langle\lambda_1^{(m_1+1-m)}|\right]_{m=1}^{m_1} \\ \left[\langle\lambda_2^{(m_2+1-m)}|\right]_{m=1}^{m_2} \\ \vdots \\ \left[\langle\lambda_n^{(m_n+1-m)}|\right]_{m=1}^{m_n} \end{bmatrix}$$

since then $Y^{-1}Y = I$, and we recall that the inverse is guaranteed to be unique.

The above demonstrates an explicit construction for the Jordan canonical form. One advantage we learn from this explicit decomposition is that the complete set of left eigenvectors and left generalized eigenvectors (encapsulated in Y^{-1}) can be obtained from the inverse of the matrix of the complete set of right eigenvectors and generalized right eigenvectors (encoded in Y) and vice versa. One unexpected lesson, though, is that the generalized left eigenvectors appear in reverse order within each Jordan block.

Using Eqs. (39) and (18) with Eq. (37), we see that the nilpotent operators $A_{\lambda,m}$ with $m > 0$ further link the various generalized eigenvectors within each subspace k. Said more suggestively, generalized modes of a nondiagonalizable subspace are necessarily cooperative.

It is worth noting that the left eigenvectors and generalized left eigenvectors form a basis for all linear functionals of the vector space spanned by the right eigenvectors and generalized right eigenvectors. Moreover, the left eigenvectors and generalized left eigenvectors are exactly the dual basis to the right eigenvectors and generalized right eigenvectors by their orthonormality properties. However, neither the left nor right eigen-basis is a priori more fundamental to the operator. Sympathetically, the right eigenvectors and generalized eigenvectors form a (dual) basis for all linear functionals of the vector space spanned by the left eigenvectors and generalized eigenvectors.

4. Simplified calculi for special cases

In special cases, the meromorphic functional calculus reduces the general expressions above to markedly simpler forms. And, this can greatly expedite practical calculations and provide physical intuition. Here, we show which reductions can be used under which assumptions.

For functions of operators with a countable spectrum, recall that the general form of the meromorphic functional calculus is:

$$f(A) = \sum_{\lambda\in\Lambda_A} \sum_{m=0}^{\nu_\lambda-1} A_{\lambda,m}\, \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z-\lambda)^{m+1}}\, dz \ . \quad (40)$$

Equations (18) and (39) gave the method to calculate $A_{\lambda,m}$ in terms of eigenvectors and generalized eigenvectors.

When the operator is *diagonalizable* (not necessarily normal), this reduces to:

$$f(A) = \sum_{\lambda \in \Lambda_A} A_\lambda \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z-\lambda)} dz \ , \tag{41}$$

where A_λ can now be constructed from conventional right and left eigenvectors, although $\langle \lambda_j |$ is *not* necessarily the conjugate transpose of $|\lambda_j\rangle$.

When the function is *analytic* on the spectrum of the (not necessarily diagonalizable) operator, then our functional calculus reduces to the holomorphic functional calculus:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda - 1} \frac{f^{(m)}(\lambda)}{m!} A_{\lambda,m} \ . \tag{42}$$

When the function is *analytic* on the spectrum of a *diagonalizable* (not necessarily normal) operator this reduces yet again to:

$$f(A) = \sum_{\lambda \in \Lambda_A} f(\lambda) A_\lambda \ . \tag{43}$$

When the function is *analytic* on the spectrum of a *diagonalizable* (not necessarily normal) operator with *no degeneracy* this reduces even further to:

$$f(A) = \sum_{\lambda \in \Lambda_A} f(\lambda) \frac{|\lambda\rangle \langle \lambda|}{\langle \lambda | \lambda \rangle} \ . \tag{44}$$

Finally, recall that an operator is *normal* when it commutes with its conjugate transpose. If the function is *analytic* on the spectrum of a *normal* operator, then we recover the simple form enabled by the spectral theorem of normal operators familiar in physics. That is, Eq. (43) is applicable, but now we have the extra simplification that $\langle \lambda_j |$ is simply the conjugate transpose of $|\lambda_j\rangle$: $\langle \lambda_j | = |\lambda_j\rangle^\dagger$.

VI. EXAMPLES AND APPLICATIONS

To illustrate the use and power of the meromorphic functional calculus, we now adapt it to analyze a suite of applications from quite distinct domains. First, we point to a set of example calculations for finite-dimensional operators of stochastic processes. Second, we show that the familiar Poisson process is intrinsically nondiagonalizable, and hint that nondiagonalizability may be common more generally in semi-Markov processes. Third, we illustrate how commonly the Drazin inverse arises in nonequilibrium thermodynamics, giving a roadmap to developing closed-from expressions for a number of key observables. Fourth, we turn to signal analysis and comment on power spectra of processes generated by nondiagonalizable operators. Finally, we round out the applications with a general discussion of Ruelle–Frobenius–Perron and Koopman operators for nonlinear dynamical systems.

A. Spectra of stochastic transition operators

The preceding employed the notation that A represents a general linear operator. In the following examples, we reserve the symbol T for the operator of a stochastic transition dynamic. If the state-space is finite and has a stationary distribution, then T has a representation that is a nonnegative row-stochastic—all rows sum to unity—transition matrix.

The transition matrix's nonnegativity guarantees that for each $\lambda \in \Lambda_T$ its complex conjugate $\overline{\lambda}$ is also in Λ_T. Moreover, the projection operator associated with the complex conjugate of λ is the complex conjugate of T_λ: $T_{\overline{\lambda}} = \overline{T_\lambda}$.

If the dynamic induced by T has a stationary distribution over the state space, then the spectral radius of T is unity and all of T's eigenvalues lie on or within the unit circle in the complex plane. The maximal eigenvalues have unity magnitude and $1 \in \Lambda_T$. Moreover, an extension of the Perron–Frobenius theorem guarantees that eigenvalues on the unit circle have algebraic multiplicity equal to their geometric multiplicity. And, so, $\nu_\zeta = 1$ for all $\zeta \in \{\lambda \in \Lambda_T : |\lambda| = 1\}$.

T's index-one eigenvalue of $\lambda = 1$ is associated with stationarity of the associated Markov process. T's other eigenvalues on the unit circle are roots of unity and correspond to deterministic periodicities within the process.

All of these results carry over from discrete to continuous time. In continuous time, where $e^{tG} = T_{t_0 \to t_0 + t}$, T's stationary eigenvalue of unity maps to G's stationary eigenvalue of zero. If the dynamic has a stationary distribution over the state space, then the rate matrix G is row-sum zero rather than row-stochastic. T's eigenvalues, on or within the unit circle, map to G's eigenvalues with nonpositive real part in the left-hand side of the complex plane.

To reduce ambiguity in the presence of multiple operators, functions of operators, and spectral mapping, we occasionally denote eigenvectors with subscripted operators on the eigenvalues within the bra or ket. For example, $|0_G\rangle = |1_T\rangle \neq |0_\mathcal{G}\rangle = |1_\mathcal{T}\rangle \neq |0_T\rangle$ disambiguates the identification of $|0\rangle$ when we have operators G, T, $\mathcal{G}$, and

$\mathcal{T}$ with $T = e^{\tau G}$, $\mathcal{T} = e^{\tau \mathcal{G}}$, and $0 \in \Lambda_G, \Lambda_{\mathcal{G}}, \Lambda_T$.

B. Randomness and memory in correlated processes

The generalized spectral theory developed here has recently been applied to give the first closed-form expressions for many measures of complexity for stochastic processes that can be generated by probabilistic finite automata [19–23]. Rather than belabor the Kolmogorov–Chaitin notion of complexity which is inherently uncomputable [47], the new analytic framework here infuses *computational mechanics* [48] with a means to compute very practical answers about an observed system's organization and to address the challenges of prediction.

For example, we can now answer the obvious questions regarding prediction: How random is a process? How much information is shared between the past and the future? How *far* into the past must we look to predict what is predictable about the future? How *much* about the observed history must be remembered to predict what is predictable about the future? And so on. The Supplementary Materials of Ref. [19] exploit the generalized spectral theory to answer these (and more) questions for the symbolic dynamics of a chaotic map, the spacetime domain for an elementary cellular automata, and the chaotic crystallographic structure of a close-packed polytypic material as determined from experimental X-ray diffractograms.

In the context of the current exposition, the most notable feature of the analyses across these many domains is that our questions, which entail tracking an observer's state of knowledge about a process, necessarily *induce* a nondiagonalizable metadynamic that becomes the central object of analysis in each case. (This metadynamic is the so-called *mixed-state presentation* of Refs. [49, 50].)

This theme, and the inherent nondiagonalizability of prediction, is explored in greater depth elsewhere [22, 23]. We also found that another nondiagonalizable dynamic is induced even in the context of quantum communication when determining how much memory reduction can be achieved if we generate a classical stochastic process using quantum mechanics [24].

We mention the above nondiagonalizable metadynamics primarily as a pointer to concrete worked-out examples where the generalized spectral theory has been employed to analyze finitary hidden Markov processes via explicitly calculated, generalized eigenvectors and projection operators. We now return to a more self-contained discussion, where we show that nondiagonalizability can be induced by the simple act of counting. Moreover, the theory developed is then applied to deliver quick and powerful results.

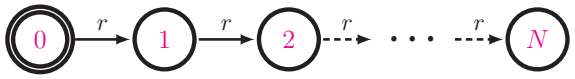


FIG. 1: Explicit Markov-chain representation of the continuous-time truncated Poisson dynamic, giving interstate transition rates r among the first $N+1$ counter-states. (State self-transition rates $-r$ are not depicted.) Taking the limit of $N \to \infty$ recovers the full Poisson counting distribution. It can either be time-homogeneous (transition-rate parameter r is time-independent) or time-inhomogeneous (parameter r is time-dependent).

C. Poisson point processes

The functional calculus leads naturally to a novel perspective on the familiar Poisson counting process—a familiar stochastic process class used widely across physics and other quantitative sciences to describe "completely random" event durations that occur over a continuous domain [51–54]. The calculus shows that the basic Poisson distribution arises as the signature of a simple nondiagonalizable dynamic. More to the point, we derive the Poisson distribution directly, without requiring the limit of the discrete-time binomial distribution, as conventionally done [29].

Consider all possible counts, up to some arbitrarily large integer N. The dynamics among these first $N+1$ counter states constitute what can be called the *truncated Poisson dynamic*. We recover the full Poisson distribution as $N \to \infty$. A Markov chain for the truncated Poisson dynamic is shown in Fig. 1. The corresponding rate matrix G, for any arbitrarily large truncation N of the possible count, is:

$$G = \begin{bmatrix} -r & r & & & \\ & -r & r & & \\ & & \ddots & \ddots & \\ & & & -r & r \\ & & & & -r \end{bmatrix} ,$$

where G_{ij} is the rate of transitioning to state (count) j given that the system is in state (count) i. Elements not on either the main diagonal or first superdiagonal are zero. This can be rewritten succinctly as:

$$G = -rI + rD_1 ,$$

where I is the identity operator in N-dimensions and D_1 is the upshift-by-1 matrix in N-dimensions, with zeros everywhere, except 1s along the first superdiagonal. Let us also define the upshift-by-m matrix D_m with zeros everywhere except 1s along the m^{th} superdiagonal, such that $D_m = D_1^m$ and $D_m^n = D_{m\cdot n}$, with $D_0 = I$. Operationally, if $\langle \delta_\ell|$ is the probability distribution over counter states that is peaked solely at state ℓ, then $\langle \delta_\ell| D_m = \langle \delta_{\ell+m}|$.

For any arbitrarily large N, G's eigenvalues are given by $\det(G - \lambda I) = (-r - \lambda)^{N+1} = 0$, from which we see that its spectrum is the singleton: $\Lambda_G = \{-r\}$. Moreover, since it has algebraic multiplicity $a_{-r} = N + 1$ and geometric multiplicity $g_{-r} = 1$, the index of the $-r$ eigenvalue is $\nu_{-r} = N + 1$. Since $-r$ is the only eigenvalue, and all projection operators must sum to the identity, we must have the eigenprojection: $G_{-r} = I$. The lesson is that the Poisson point process is highly nondiagonalizable.

1. Homogeneous Poisson processes

When the transition rate r between counter states is constant in time, the net counter state-to-state transition operator from initial time 0 to later time t is given simply by:

$$T(t) = e^{tG} \ .$$

The functional calculus allows us to directly evaluate e^{tG} for the Poisson nondiagonalizable transition-rate operator G; we find:

$$\begin{aligned} T(t) &= e^{tG} \\ &= \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_\lambda - 1} G_\lambda (G - \lambda I)^m \Big(\tfrac{1}{2\pi i} \oint_{C_\lambda} \frac{e^{tz}}{(z-\lambda)^m} \, dz \Big) \\ &= \lim_{N \to \infty} \sum_{m=0}^{N} I (G + rI)^m \frac{1}{m!} \underbrace{\lim_{z \to -r} \frac{d^m}{dz^m} e^{tz}}_{t^m e^{-rt}} \\ &= \sum_{m=0}^{\infty} (rD_1)^m \frac{t^m e^{-rt}}{m!} \\ &= \sum_{m=0}^{\infty} D_m \frac{(rt)^m e^{-rt}}{m!} \ . \end{aligned}$$

Consider the orthonormality relation $\langle \delta_i | \delta_j \rangle = \delta_{i,j}$ between counter states, where $|\delta_j\rangle$ is represented by 0s everywhere except for a 1 at counter-state j. It effectively measures the occupation probability of counter-state j. Employing the result for $T(t)$, we find the simple consequence that:

$$\begin{aligned} \langle \delta_0 | \, T(t) \, | \delta_n \rangle &= \frac{(rt)^n e^{-rt}}{n!} \\ &= \langle \delta_m | \, T(t) \, | \delta_{m+n} \rangle \ . \end{aligned}$$

That is, the probability that the counter is incremented by n in a time interval t is independent of the initial count and given by: $(rt)^n e^{-rt}/n!$.

Let us emphasize that these steps derived the Poisson distribution directly, rather than as the typical limit of the binomial distribution. Our derivation depended critically on spectral manipulations of a highly nondiagonalizable operator. Moreover, our result for the transition dynamic $T(t)$ allows a direct analysis of how *distributions* over counts evolve in time, as would be necessary, say, in a Bayesian setting with unknown prior count. This type of calculus can immediately be applied to the analysis of more sophisticated processes, for which we can generally expect nondiagonalizability to play an important functional role.

2. Inhomogeneous Poisson processes

Let us now generalize to time-inhomogeneous Poisson processes, where the transition rate r between count events is instantaneously uniform, but varies in time as $r(t)$. Conveniently, the associated rate matrices at different times commute with each other. Specifically, with $G_a = -aI + aD_1$ and $G_b = -bI + bD_1$, we see that:

$$[G_a, \, G_b] = 0 \ .$$

Therefore, the net counter state-to-state transition operator from time t_0 to time t_f is given by:

$$\begin{aligned} T_{t_0, t_f} &= e^{\int_{t_0}^{t_f} G(t) \, dt} \\ &= e^{\left(\int_{t_0}^{t_f} r(t) \, dt \right) (-I + D_1)} \\ &= e^{\langle r \rangle (\Delta t)(-I + D_1)} \\ &= e^{(\Delta t) G_{\langle r \rangle}} \ , \end{aligned} \tag{45}$$

where $\Delta t = t_f - t_0$ is the time elapsed and:

$$\langle r \rangle = \tfrac{1}{\Delta t} \int_{t_0}^{t_f} r(t) \, dt$$

is the average rate during that time. Given Eq. (45), the functional calculus proceeds just as in the time-homogeneous case to give the analogous net transition

dynamic:

$$T_{t_0,t_f} = \sum_{m=0}^{\infty} D_m \frac{\left(\langle r\rangle \Delta t\right)^m e^{-\langle r\rangle \Delta t}}{m!} \ .$$

The probability that the count is incremented by n during the time interval Δt follows directly:

$$\langle \delta_m | \, T_{t_0,t_f} \, | \delta_{m+n} \rangle = \frac{\left(\langle r\rangle \Delta t\right)^n e^{-\langle r\rangle \Delta t}}{n!} \ .$$

With relative ease, our calculus allowed us to derive an important result for stochastic process theory that is nontrivial to derive by other means. Perhaps surprisingly, we see that the probability distribution over final counts induced by any rate trajectory $r(t)$ is the same as if the transition rate were held fixed at mean $\langle r\rangle$ throughout the duration. Moreover, we can directly analyze the net evolution of distributions over counts using the derived transition operator T_{t_0,t_f}.

Note that the nondiagonalizability of the Poisson dynamic is robust in a physical sense. That is, even varying the rate parameter in time in an erratic way, the inherent structure of counting imposes a fundamental nondiagonalizable nature. That nondiagonalizability can be robust in a physical sense is significant, since one might otherwise be tempted to argue that nondiagonalizability is extremely fragile due to numerical perturbations within any matrix representation of the operator. This is simply not the case since such perturbations are physically forbidden. Rather, this simple example challenges us with the fact that some processes, even those familiar and widely used, are intrinsically nondiagonalizable. On the positive side, it appears that spectral methods can now be applied to analyze them. And, this will be particularly important in more complex, memoryful processes [55–58], including the hidden semi-Markov processes [51, 59] that are, roughly speaking, the cross-product of hidden finite-state Markov chains and renewal processes.

D. Stochastic thermodynamics

The previous simple examples started to demonstrate the spectral methods of the functional calculus. Next, we show a novel application of the meromorphic functional calculus to environmentally driven mesoscopic dynamical systems, selected to give a new set of results within nonequilibrium thermodynamics. In particular, we analyze functions of singular transition-rate operators. Notably, we show that the Drazin inverse arises naturally in the general solution of Green–Kubo relations. We mention that it also arises when analyzing moments of the excess heat produced in the driven transitions atop either equilibrium steady states or nonequilibrium steady states [26].

1. Dynamics in independent eigenspaces

An important feature of the functional calculus is its ability to address particular eigenspaces independently when necessary. This feature is often taken for granted in the case of normal operators; say, in physical dynamical systems when analyzing stationary distributions or dominant decay modes. Consider a singular operator $\mathcal{L}$ that is not necessarily normal and not necessarily diagonalizable and evaluate the simple yet ubiquitous integral $\int_0^\tau e^{t\mathcal{L}}\, dt$. Via the meromorphic functional calculus we find:

$$\begin{aligned}
\int_0^\tau e^{t\mathcal{L}}\, dt &= \sum_{\lambda\in\Lambda_\mathcal{L}} \sum_{m=0}^{\nu_\lambda - 1} \mathcal{L}_{\lambda,m} \tfrac{1}{2\pi i} \oint_{C_\lambda} \frac{\int_0^\tau e^{tz}\, dt}{(z-\lambda)^{m+1}}\, dz \\
&= \Big(\sum_{m=0}^{\nu_0 - 1} \mathcal{L}_{0,m} \tfrac{1}{2\pi i} \oint_{C_0} \frac{z^{-1}(e^{\tau z} - 1)}{z^{m+1}}\, dz \Big) \\
&\quad + \sum_{\lambda\in\Lambda_\mathcal{L}\setminus 0} \sum_{m=0}^{\nu_\lambda - 1} \mathcal{L}_{\lambda,m} \tfrac{1}{2\pi i} \oint_{C_\lambda} \frac{z^{-1}(e^{\tau z} - 1)}{(z-\lambda)^{m+1}}\, dz \\
&= \Big(\sum_{m=0}^{\nu_0 - 1} \tfrac{\tau^{m+1}}{(m+1)!} \mathcal{L}_{0,m} \Big) + \mathcal{L}^\mathcal{D} \left(e^{\tau\mathcal{L}} - I \right) \ , \quad (46)
\end{aligned}$$

where $\mathcal{L}^\mathcal{D}$ is the Drazin inverse of $\mathcal{L}$, discussed earlier.

The pole–pole interaction (z^{-1} with z^{-m-1}) at $z = 0$ distinguished the 0-eigenspace in the calculations and required the meromorphic functional calculus for direct analysis. The given solution to this integral will be useful in the following.

Next, we consider the case where $\mathcal{L}$ is the transition-rate operator among the states of a structured stochastic dynamical system. This leads to several novel consequence within stochastic thermodynamics.

2. Green–Kubo relations

Let us reconsider the above integral in the case when the singular operator $\mathcal{L}$—let us call it G—is a transition-rate operator that exhibits a single stationary distribution. By the spectral mapping $\ln \Lambda_{e^G}$ of the eigenvalue $1 \in \Lambda_{e^G}$ addressed in the Perron–Frobenius theorem, G's zero eigenmode is diagonalizable. And, by assuming a single attracting stationary distribution, the zero eigenvalue has algebraic multiplicity $a_0 = 1$. Equation (46)

then simplifies to:

$$\int_0^\tau e^{tG}\,dt = \tau\,|0_G\rangle\langle 0_G| + G^{\mathcal{D}}\left(e^{\tau G} - I\right) \ . \tag{47}$$

Since G is a transition-rate operator, the above integral corresponds to integrated time evolution. The Drazin inverse $G^{\mathcal{D}}$ concentrates on the transient contribution beyond the persistent stationary background. In Eq. (47), the subscript within the left and right eigenvectors explicitly links the eigenvectors to the operator G, to reduce ambiguity. Specifically, the projector $|0_G\rangle\langle 0_G|$ maps any distribution to the stationary distribution.

Green–Kubo-type relations [60, 61] connect the out-of-steady-state transport coefficients to the time integral of steady-state autocorrelation functions. They are thus very useful for understanding out-of-steady-state dissipation due to steady-state fluctuations. (Steady state here refers to either equilibrium or nonequilibrium steady state.) Specifically, the Green–Kubo relation for a transport coefficient, κ say, is typically of the form:

$$\kappa = \int_0^\infty \left(\langle A(0)A(t)\rangle_{\text{s.s.}} - \langle A\rangle_{\text{s.s.}}^2\right) dt \ ,$$

where $A(0)$ and $A(t)$ are some observable of the stationary stochastic dynamical system at time 0 and time t, respectively, and the subscript $\langle\cdot\rangle_{\text{s.s.}}$ emphasizes that the expectation value is to be taken according to the steady-state distribution.

Using:

$$\begin{aligned}\langle A(0)A(t)\rangle_{\text{s.s.}} &= \text{tr}\big(|0_G\rangle\langle 0_G|\,A\,e^{tG}A\big)\\ &= \langle 0_G|\,A\,e^{tG}A\,|0_G\rangle \ ,\end{aligned}$$

the transport coefficient κ can be written more explicitly in terms of the relevant transition-rate operator G for the stochastic dynamics:

$$\begin{aligned}\kappa &= \lim_{\tau\to\infty}\int_0^\tau \langle 0_G|\,A\,e^{tG}A\,|0_G\rangle\ dt - \tau\,\langle 0_G|\,A\,|0_G\rangle^2\\ &= \lim_{\tau\to\infty}\langle 0_G|\,A\Big(\int_0^\tau e^{tG}\,dt\Big)A\,|0_G\rangle - \tau\,\langle 0_G|\,A\,|0_G\rangle^2\\ &= \lim_{\tau\to\infty}\langle 0_G|\,A\,G^{\mathcal{D}}\big(e^{\tau G} - I\big)A\,|0_G\rangle\\ &= -\,\langle A\,G^{\mathcal{D}}A\rangle_{\text{s.s.}} \ .\end{aligned} \tag{48}$$

Thus, we learn that relations of Green–Kubo form are direct signatures of the Drazin inverse of the transition-rate operator for the stochastic dynamic.

The result of Eq. (48) holds quite generally. For example, if the steady state has some number of periodic flows, the result of Eq. (48) remains valid. Alternatively, in the case of nonperiodic chaotic flows—where G will be the logarithm of the Ruelle–Frobenius–Perron operator, as described later in § VI F 1—$|0_G\rangle\langle 0_G|$ still induces the average over the steady-state trajectories.

In the special case where the transition-rate operator is diagonalizable, $-\langle A\,G^{\mathcal{D}}A\rangle_{\text{s.s.}}$ is simply the integrated contribution from a weighted sum of decaying exponentials. Transport coefficients then have a solution of the simple form:

$$\kappa = -\sum_{\lambda\in\Lambda_G\setminus 0}\frac{1}{\lambda}\,\langle 0_G|\,A\,G_\lambda A\,|0_G\rangle \ . \tag{49}$$

Note that the minus sign keeps κ positive since $\text{Re}(\lambda) < 0$ for $\lambda \in \Lambda_G \setminus \{0\}$. Also, recall that G's eigenvalues with nonzero imaginary part occur in complex-conjugate pairs and $G_{\overline{\lambda}} = \overline{G_\lambda}$. Moreover, if $G_{i,j}$ is the classical transition-rate from state i to state j (to disambiguate from the transposed possibility), then $\langle 0_G|$ is the stationary distribution. (The latter is sometimes denoted $\langle\boldsymbol{\pi}|$ in the Markov process literature.) And, $|0_G\rangle$ is a column vector of all ones (sometimes denoted $|\mathbf{1}\rangle$) which acts to integrate contributions throughout the state space.

A relationship of the form of Eq. (48), between the Drazin inverse of a classical transition-rate operator and a particular Green–Kubo relation was recently found in Ref. [62] for the friction tensor for smoothly-driven transitions atop nonequilibrium steady states. Subsequently, a truncation of the eigen-expansion of the form of Eq. (49) was recently used in a similar context to bound a universal tradeoff between power, precision, and speed [63]. Equation (48) shows that a fundamental relationship between a physical property and a Drazin inverse is to be expected more generally whenever the property can be related to integrated correlation.

Notably, if a Green–Kubo-like relation integrates a cross-correlation, say between $A(t)$ and $B(t)$ rather than an autocorrelation, then we have only the slight modification:

$$\int_0^\infty \big(\langle A(0)B(t)\rangle_{\text{s.s.}} - \langle A\rangle_{\text{s.s.}}\langle B\rangle_{\text{s.s.}}\big)dt = -\langle A\,G^{\mathcal{D}}B\rangle_{\text{s.s.}} . \tag{50}$$

The foregoing analysis bears on both classical and quantum dynamics. G may be a so-called linear *superoperator* in the quantum regime [64]; for example, the *Lindblad superoperator* [65, 66] that evolves density operators. A Liouville-space representation [67] of the superoperator, though, exposes the superficiality of the distinction between superoperator and operator. At an abstract level, time evolution can be discussed uniformly across subfields and reinterpretations of Eq. (50) will be found in each associated physical theory.

Reference [26] presents additional constructive results

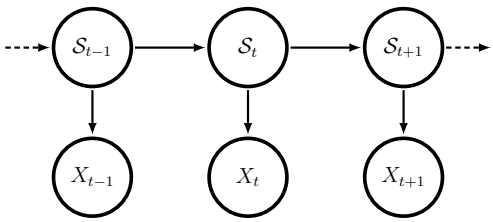


FIG. 2: Bayes network for a state-emitting hidden Markov model graphically depicts the structure of conditional independence among random variables for the latent state $\{\mathcal{S}_n\}_{n\in\mathbb{Z}}$ at each time n and the random variables $\{X_n\}_{n\in\mathbb{Z}}$ for the observation at each time n.

that emphasize the ubiquity of integrated correlation and Drazin inverses in the transitions between steady states [68], relevant to the fluctuations within any physical dynamic. Overall, these results support the broader notion that dissipation depends on the structure of correlation.

Frequency-dependent generalizations of integrated correlation have a corresponding general solution. To be slightly less abstract, we give novel representative formulae for a particular application: the general solution to power spectra of a process generated by any countable-state hidden Markov chain.

E. Power spectra

A signal's power spectrum quantifies how its power is distributed across frequency [69]. For a discrete-domain process it is:

$$P(\omega) = \lim_{N\to\infty} \tfrac{1}{N} \left\langle \left| \sum_{n=1}^{N} X_n e^{-i\omega n} \right|^2 \right\rangle \ , \qquad (51)$$

where ω is the angular frequency and X_n is the random variable for the observation at time n. For a wide-sense stationary stochastic process, the power spectrum is also determined from the signal's autocorrelation function $\gamma(\tau)$:

$$P(\omega) = \lim_{N\to\infty} \tfrac{1}{N} \sum_{\tau=-N}^{N} \big(N - |\tau|\big)\gamma(\tau) e^{-i\omega\tau} \ , \qquad (52)$$

where the autocorrelation function for a wide-sense stationary stochastic process is defined:

$$\gamma(\tau) = \left\langle \overline{X_n} X_{n+\tau} \right\rangle_n \ .$$

The windowing function $N - |\tau|$ appearing in Eq. (52) is a direct consequence of Eq. (51). It is not imposed externally, as is common practice in signal analysis. This is important to subsequent derivations.

The question we address is how to calculate the correlation function and power spectrum given a model of the signal's generator. To this end, we briefly introduce hidden Markov models as signal generators and then use the meromorphic calculus to calculate their autocorrelation and power spectra in closed-form. This leads to several lessons. First, we see that the power spectrum is a direct fingerprint of the resolvent of the generator's time-evolution operator, analyzed along the unit circle. Second, spectrally decomposing the not-necessarily-diagonalizable time evolution operator, we derive the most general qualitative behavior of the autocorrelation function and power spectra. Third, contributions from eigenvalues on the unit circle must be extracted and dealt with separately. Contributions from eigenvalues on the unit circle correspond to Dirac delta functions—the analog of Bragg peaks in diffraction. Whereas, eigen-contributions from inside the unit circle correspond to diffuse peaks, which become sharper for eigenvalues closer to the unit circle. Finally, nondiagonalizable eigenmodes yield qualitatively different line profiles than their diagonalizable counterparts. In short, when applied to signal analysis our generalized spectral decomposition has directly measurable consequences. This has been key to analyzing low-dimensional disordered materials, for example, when adapted to X-ray diffraction spectra [20, 21, 70].

Let the 4-tuple $\mathcal{M} = \big(\boldsymbol{\mathcal{S}}, \mathcal{A}, \mathcal{P}, T\big)$ be a discrete-time *state-emitting hidden Markov model* (HMM) that generates the stationary stochastic process $\ldots X_{-2}X_{-1}X_0X_1X_2\ldots$ according to the following. $\boldsymbol{\mathcal{S}}$ is the (finite) set of latent states of the hidden Markov chain and $\mathcal{A} \subseteq \mathbb{C}$ is the observable alphabet. $\mathcal{S}_t$ is the random variable for the hidden state at time t that takes on values $s \in \boldsymbol{\mathcal{S}}$. X_t is the random variable for the observation at time t that takes on values $x \in \mathcal{A}$. Given the latent state at time t, the possible observations are distributed according to the conditional probability density functions: $\mathcal{P} = \big\{\mathrm{p}(X_t = x|\mathcal{S}_t = s)\big\}_{s\in\boldsymbol{\mathcal{S}}}$. For each $s \in \boldsymbol{\mathcal{S}}$, $\mathrm{p}(X_t = x|\mathcal{S}_t = s)$ may be abbreviated as $\mathrm{p}(x|s)$ since the probability density function in each state is assumed not to change over t. Finally, the latent-state-to-state stochastic transition matrix T has elements $T_{i,j} = \Pr(\mathcal{S}_{t+1} = s_j|\mathcal{S}_t = s_i)$, which give the probability of transitioning from latent state s_i to s_j given that the system is in state s_i, where $s_i, s_j \in \boldsymbol{\mathcal{S}}$. It is important for the subsequent derivation that we use $\Pr(\cdot)$ to denote a probability in contrast to $\mathrm{p}(\cdot)$ which denotes a probability *density*. The Bayes network

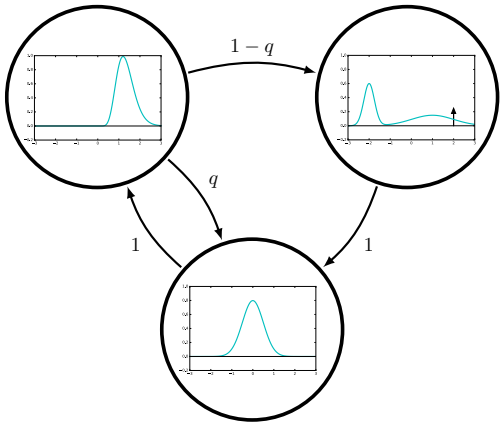


FIG. 3: Simple 3-state state-emitting HMM that generates a stochastic process according to the state-to-state transition dynamic T and the probability density functions (pdfs) $\{\mathrm{p}(x|s)\}_{s\in\boldsymbol{\mathcal{S}}}$ associated with each state. Theorem 1 asserts that its power spectrum will be the same (with only constant offset) as the power spectrum generated from the alternative process where the pdfs in each state are solely concentrated at the Platonic average value $\langle x\rangle_{\mathrm{p}_s(x)}$ of the former pdf associated with the state.

diagram of Fig. 2 depicts the structure of conditional independence among the random variables.

1. *Continuous-value, discrete-state and -time processes*

Figure 3 gives a particular HMM with continuous observable alphabet $\mathcal{A} = \mathbb{R}$ distributed according to the probability density function shown within each latent state. Processes generated as the observation of a function of a Markov chain can be of either finite or infinite Markov order. (They are, in fact, *typically* infinite Markov order in the space of processes [71].)

Directly calculating, one finds that the autocorrelation function, for $\tau > 0$, for any such HMM is:

$$\begin{aligned}
\gamma(\tau) &= \left\langle \overline{X_n} X_{n+\tau} \right\rangle_n \\
&= \int_{x\in\mathcal{A}} \int_{x'\in\mathcal{A}} \overline{x} x' \mathrm{p}(X_0 = x, X_\tau = x')\, dx\, dx' \\
&= \sum_{s\in\boldsymbol{\mathcal{S}}} \sum_{s'\in\boldsymbol{\mathcal{S}}} \int_{x\in\mathcal{A}} \int_{x'\in\mathcal{A}} \overline{x} x' \mathrm{p}(X_0 = x, X_\tau = x', \mathcal{S}_0 = s, \mathcal{S}_\tau = s')\, dx\, dx' \\
&= \sum_{s\in\boldsymbol{\mathcal{S}}} \sum_{s'\in\boldsymbol{\mathcal{S}}} \int_{x\in\mathcal{A}} \int_{x'\in\mathcal{A}} \overline{x} x' \Pr(\mathcal{S}_0 = s, \mathcal{S}_\tau = s')\, \mathrm{p}(X_0 = x|\mathcal{S}_0 = s)\, \mathrm{p}(X_\tau = x'|\mathcal{S}_\tau = s')\, dx\, dx' \\
&= \sum_{s\in\boldsymbol{\mathcal{S}}} \sum_{s'\in\boldsymbol{\mathcal{S}}} \langle \boldsymbol{\pi}|\delta_s\rangle \langle \delta_s| T^\tau |\delta_{s'}\rangle \langle \delta_{s'}|\mathbf{1}\rangle \left(\int_{x\in\mathcal{A}} \overline{x}\, \mathrm{p}(x|s)\, dx\right) \left(\int_{x'\in\mathcal{A}} x'\, \mathrm{p}(x'|s')\, dx'\right) \\
&= \langle \boldsymbol{\pi}| \left(\sum_{s\in\boldsymbol{\mathcal{S}}} \langle \overline{x}\rangle_{\mathrm{p}(x|s)} |\delta_s\rangle \langle \delta_s|\right) T^\tau \left(\sum_{s'\in\boldsymbol{\mathcal{S}}} \langle x\rangle_{\mathrm{p}(x|s')} |\delta_{s'}\rangle \langle \delta_{s'}|\right) |\mathbf{1}\rangle \ ,
\end{aligned}$$

where:

$$\mathrm{p}(X_0 = x, X_\tau = x', \mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \Pr(\mathcal{S}_0 = s, \mathcal{S}_\tau = s')\mathrm{p}(X_0 = x, X_\tau = x'|\mathcal{S}_0 = s, \mathcal{S}_\tau = s')$$

holds by definition of conditional probability. The decomposition of:

$$\mathrm{p}(X_0 = x, X_\tau = x'|\mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \mathrm{p}(X_0 = x|\mathcal{S}_0 = s)\mathrm{p}(X_\tau = x'|\mathcal{S}_\tau = s')$$

for $\tau \neq 0$ follows from the conditional independence in the relevant Bayesian network shown in Fig. 2. Moreover, the equality:

$$\Pr(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \langle \boldsymbol{\pi}|\delta_s\rangle \langle \delta_s| T^\tau |\delta_{s'}\rangle \langle \delta_{s'}|\mathbf{1}\rangle$$

can be derived by marginalizing over all possible intervening state sequences. Note that $|\delta_s\rangle$ is the column vector of all 0s except for a 1 at the index corresponding to state s and $\langle\delta_s|$ is simply its transpose. Recall that $\langle\boldsymbol{\pi}| = \langle 1_T|$ is the stationary distribution induced by T over latent states and $|\mathbf{1}\rangle = |1_T\rangle$ is a column vector of all ones. Note also that $\langle\boldsymbol{\pi}|\delta_s\rangle = \Pr(s)$ and $\langle\delta_{s'}|\mathbf{1}\rangle = 1$.

Since the autocorrelation function is symmetric in τ and:

$$\begin{aligned}\gamma(0) &= \left\langle |x|^2 \right\rangle_{\mathrm{p}(x)} \\ &= \langle\boldsymbol{\pi}| \sum_{s\in\boldsymbol{\mathcal{S}}} \left\langle |x|^2 \right\rangle_{\mathrm{p}(x|s)} |\delta_s\rangle \ ,\end{aligned}$$

we find the full autocorrelation function is given by:

$$\gamma(\tau) = \begin{cases} \left\langle |x|^2 \right\rangle & \text{if } \tau = 0 \\ \langle\boldsymbol{\pi}|\,\overline{\Omega}\,T^{|\tau|-1}\Omega\,|\mathbf{1}\rangle & \text{if } |\tau| \geq 1 \end{cases} \ ,$$

where Ω is the $|\boldsymbol{\mathcal{S}}|$-by-$|\boldsymbol{\mathcal{S}}|$ matrix defined by:

$$\Omega = \sum_{s\in\boldsymbol{\mathcal{S}}} \langle x\rangle_{\mathrm{p}(x|s)} |\delta_s\rangle \langle\delta_s| \, T \ . \tag{53}$$

The power spectrum is then calculated via Eq. (52) using the meromorphic calculus. In particular, the power spectrum decomposes naturally into a discrete part and a continuous part. Full details will be given elsewhere, but the derivation is similar to that given in Ref. [20] for the special case of diffraction patterns from HMMs. We note that it is important to treat individual eigenspaces separately, as our generalized calculus naturally accommodates. The end result, for the continuous part of the power spectrum, is:

$$P_{\mathrm{c}}(\omega) = \left\langle |x|^2 \right\rangle + 2\,\mathrm{Re}\,\langle\boldsymbol{\pi}|\,\overline{\Omega}\left(e^{i\omega}I - T\right)^{-1}\Omega\,|\mathbf{1}\rangle \ . \tag{54}$$

All of the ω-dependence is in the resolvent. Using the spectral expansion of the resolvent given by Eq. (21) allows us to better understand the qualitative possibilities for the shape of the power spectrum:

$$P_{\mathrm{c}}(\omega) = \left\langle |x|^2 \right\rangle + \sum_{\lambda\in\Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} 2\,\mathrm{Re}\frac{\langle\boldsymbol{\pi}|\,\overline{\Omega}\,T_{\lambda,m}\Omega\,|\mathbf{1}\rangle}{(e^{i\omega}-\lambda)^{m+1}} \ . \tag{55}$$

Note that $\langle\boldsymbol{\pi}|\,\overline{\Omega}\,T_{\lambda,m}\Omega\,|\mathbf{1}\rangle$ is a complex-valued scalar and all of the frequency dependence now handily resides in the denominator.

The discrete portion (delta functions) of the power spectrum is:

$$\begin{aligned}P_{\mathrm{d}}(\omega) = \sum_{k=-\infty}^{\infty} \sum_{\substack{\lambda\in\Lambda_T \\ |\lambda|=1}} & 2\pi\,\delta(\omega - \omega_\lambda + 2\pi k) \\ & \times \mathrm{Re}\big(\lambda^{-1}\,\langle\boldsymbol{\pi}|\,\overline{\Omega}\,T_\lambda\Omega\,|\mathbf{1}\rangle\big) \ ,\end{aligned} \tag{56}$$

where ω_λ is related to λ by $\lambda = e^{i\omega_\lambda}$. An extension of the Perron–Frobenius theorem guarantees that the eigenvalues of T on the unit circle have index $\nu_\lambda = 1$.

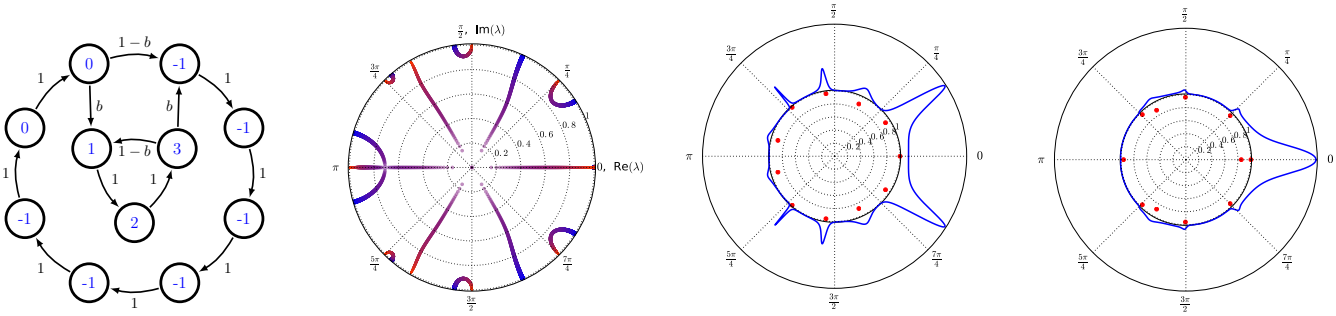
When plotted as a function of the angular frequency ω around the unit circle, the power spectrum suggestively appears to emanate from the eigenvalues $\lambda \in \Lambda_T$ of the hidden linear dynamic. See Fig. 4 for the analysis of an example parametrized process and the last two panels for this display mode for the power spectra.

Eigenvalues of T *on* the unit circle yield Dirac delta functions in the power spectrum. Eigenvalues of T *within* the unit circle yield more diffuse line profiles, increasingly diffuse as the magnitude of the eigenvalues retreats toward the origin. Moreover, the integrated magnitude of each contribution is determined by projecting pairwise observation operators onto the eigenspace emanating the contribution. Finally, we note that nondiagonalizable eigenmodes yield qualitatively different line profiles.

Remarkably, the power spectrum generated by such a process is the same as the that generated by a potentially much simpler one—a process that is a function of the same underlying Markov chain but instead emits the state-dependent *expectation value* of the observable within each state:

Theorem 1. *Let $\mathcal{P} = \{p_s(x)\}_{s\in\boldsymbol{\mathcal{S}}}$ be any set of probability distribution functions over the domain $\mathcal{A} \subseteq \mathbb{C}$. Let $\mathcal{B} = \{\langle x\rangle_{p_s(x)}\}_{s\in\boldsymbol{\mathcal{S}}}$ and let $\mathcal{Q} = \{\delta(x - \langle x\rangle_{p_s(x)})\}_{s\in\boldsymbol{\mathcal{S}}}$. Then, the power spectrum generated by any hidden Markov model $\mathcal{M} = (\boldsymbol{\mathcal{S}}, \mathcal{A}, \mathcal{P}, T)$ differs at most by a constant offset from the power spectrum generated by the hidden Markov model $\mathcal{M}' = (\boldsymbol{\mathcal{S}}, \mathcal{B}, \mathcal{Q}, T)$ that has the same latent Markov chain but in any state $s \in \boldsymbol{\mathcal{S}}$ emits, with probability one, the average value $\langle x\rangle_{p_s(x)}$ of the state-conditioned probability density function $p_s(x) \in \mathcal{P}$ of $\mathcal{M}$.*

Proof. *From Eqs.* (54) *and* (56)*, we see that $P_c(\omega) + P_d(\omega) - \left\langle |x|^2 \right\rangle$ depends only on T and $\{\langle x\rangle_{p(x|s)}\}_{s\in\boldsymbol{\mathcal{S}}}$. This shows that all HMMs that share the same T and $\{\langle x\rangle_{p(x|s)}\}_{s\in\boldsymbol{\mathcal{S}}}$ have the same power spectrum $P(\omega) = P_c(\omega) + P_d(\omega)$ besides a constant offset determined by*



(a) A b-parametrized HMM with mean values of each state's pdf $\langle x \rangle_{\mathrm{p}(x|s)}$ indicated as the number inside each state.

(b) Eigenvalue evolution for all $\lambda \in \Lambda_T$ sweeping transition parameter b from 1 to 0.

(c) Power spectrum and eigenvalues at $b = 3/4$.

(d) Power spectrum and eigenvalues at $b = 1/4$.

FIG. 4: Parametrized HMM generator of a stochastic process, its eigenvalue evolution, and two coronal spectrograms showing power spectra emanating from eigen-spectra.

differences in $\langle |x|^2 \rangle$.

One immediate consequence is that *any hidden Markov chain with any arbitrary set of zero-mean distributions attached to each state*, i.e.:

$$\mathcal{P} \in \left\{ \{\mathrm{p}(x|s)\}_{s \in \boldsymbol{\mathcal{S}}} : \langle x \rangle_{\mathrm{p}(x|s)} = 0 \text{ for all } s \in \boldsymbol{\mathcal{S}} \right\} ,$$

generates a flat power spectrum with the appearance of white noise. On the one hand, this strongly suggests to data analysts to look beyond power spectra when attempting to extract a process' full architecture. On the other, whenever a process's power spectrum *is* structured, it is a direct fingerprint of the resolvent of the hidden linear dynamic. In short, the power spectrum is a filtered image of the resolvent along the unit circle.

The power spectrum of a particular stochastic process is shown in Fig. 4 and using *coronal spectrograms*, introduced in Ref. [20], it illustrates how the observed spectrum can be thought of as emanating from the spectrum of the hidden linear dynamic, as all power spectra must. Figure 4a shows the state-emitting HMM with state-to-state transition probabilities parametrized by b; the mean values $\langle x \rangle_{\mathrm{p}(x|s)}$ of each state's pdf $\mathrm{p}(x|s)$ are indicated as the blue number inside each state. The process generated depends on the actual pdfs and the transition parameter b although, and this is our point, the power spectrum is ignorant to the details of the pdfs.

The evolution of the eigenvalues Λ_T of the transition dynamic among latent states is shown from thick blue to thin red markers in Fig. 4b, as we sweep the transition parameter b from 1 to 0. A subset of the eigenvalues pass continuously but very quickly through the origin of the complex plane as b passes through 1/2. The continuity of this is not immediately apparent numerically, but can be revealed with a finer increment of b near $b \approx 1/2$. Notice the persistent eigenvalue of $\lambda_T = 1$, which is guaranteed by the Perron–Frobenius theorem.

In Fig. 4c and again, at another parameter setting, in Fig. 4d, we show the continuous part of the *power* spectrum $P_{\mathrm{c}}(\omega)$ (plotted around the unit circle in solid blue) and the *eigen*-spectrum (plotted as red dots on and within the unit circle) of the state-to-state transition matrix for the 11-state hidden Markov chain (leftmost panel) that generates it. There is also a δ-function contribution to the power spectrum at $\omega = 0$ (corresponding to $\lambda_T = 1$). This is not shown. These coronal spectrograms illustrate how the power spectrum emanates from the HMM's eigen-spectrum, with sharper peaks when the eigenvalues are closer to the unit circle. This observation is fully explained by Eq. (55). The integrated magnitude of each peak depends on $\langle \boldsymbol{\pi}| \, \overline{\Omega} \, |\lambda\rangle \, \langle \lambda| \, \Omega \, |\mathbf{1}\rangle$.

Interestingly, the apparent continuous spectrum component is the shadow of the discrete spectrum of nonunitary dynamics. This suggests that resonances in various physics domains concerned with a continuous spectrum can be modeled as simple consequences of nonunitary dynamics. Indeed, hints of this appear in the literature [72–74].

2. *Continuous-time processes*

We close this exploration of conventional signal analysis methods using the meromorphic calculus by commenting on continuous-time processes. Analogous formulae can be derived with similar methods for continuous-time hidden Markov jump processes and continuous-time deterministic (possibly chaotic) dynamics in terms of the generator G of time evolution. For example, the continuous part $P_{\mathrm{c}}(\omega)$ of the power spectrum from a continuous-

time deterministic dynamic has the form:

$$P_{\mathrm{c}}(\omega) = 2\,\mathrm{Re}\,\langle \boldsymbol{\pi}|\,\overline{\Omega}\left(i\omega I - G\right)^{-1}\Omega\,|\mathbf{1}\rangle \ .$$

Appealing to the resolvent's spectral expansion again allows us to better understand the possible shapes of their power spectrum:

$$P_{\mathrm{c}}(\omega) = \sum_{\lambda\in\Lambda_G}\sum_{m=0}^{\nu_\lambda-1} 2\,\mathrm{Re}\frac{\langle \boldsymbol{\pi}|\,\overline{\Omega}\,G_{\lambda,m}\Omega\,|\mathbf{1}\rangle}{(i\omega-\lambda)^{m+1}} \ . \tag{57}$$

Since all of the frequency-dependence has been isolated in the denominator and $\langle \boldsymbol{\pi}|\,\overline{\Omega}\,G_{\lambda,m}\Omega\,|\mathbf{1}\rangle$ is a frequency-independent complex-valued constant, peaks in $P_{\mathrm{c}}(\omega)$ can only arise via contributions of the form $\mathrm{Re}\frac{c}{(i\omega-\lambda)^n}$ for $c\in\mathbb{C}$, $\omega\in\mathbb{R}$, $\lambda\in\Lambda_G$, and $n\in\mathbb{Z}_+$. This provides a rich starting point for application and further theoretical investigation. For example, Eq. (57) helps explain the shapes of power spectra of nonlinear dynamical systems, as have appeared, e.g., in Ref. [75]. Furthermore, it suggests an approach to the inverse problem of inferring the spectrum of the hidden linear dynamic via power spectra. In the next section, however, we develop a more general proposal for inferring eigenvalues from a time series. Further developments will appear elsewhere.

F. Operators for chaotic dynamics

Since trajectories in state-space can be generated independently of each other, any nonlinear dynamic corresponds to a linear operation on an infinite-dimensional vector-space of complex-valued distributions (in the sense of generalized functions) over the original state-space. For example, the well-known Lorenz ordinary differential equations [76] are nonlinear over its three given state-space variables—x, y, and z. Nevertheless, the dynamic is linear in the infinite-dimensional vector space $D(\mathbb{R}^3)$ of distributions over $\mathbb{R}^3$. Although $D(\mathbb{R}^3)$ is an unwieldy state-space, the dynamics there might be well approximated by a finite truncation of its modes.

1. *Ruelle–Frobenius–Perron and Koopman operators*

The preceding operator formalism applies, in principle at least. The question, of course, is, Is it practical and does it lead to constructive consequences? Let's see. The right eigenvector is either $|0_G\rangle$ or $|1_T\rangle$ with $T = e^{\tau G}$ as the Ruelle–Frobenius–Perron transition operator [77, 78]. Equivalently, it is also $\boldsymbol{\pi}$, the stationary distribution, with support on attracting subsets of $\mathbb{R}^3$ in the case of the Lorenz dynamic. The corresponding left-eigenvector $\mathbf{1}$, either $\langle 0_G|$ or $\langle 1_T|$, is uniform over the space. Other modes of the operator's action, according to the eigenvalues and left and right eigenvectors and generalized eigenvectors, capture the decay of arbitrary distributions on $\mathbb{R}^3$.

The meromorphic spectral methods developed above give a view of the Koopman operator and Koopman modes of nominally nonlinear dynamical systems [4] that is complementary to the Ruelle–Frobenius–Perron operator. The Koopman operator K is the adjoint—in the sense of vector spaces, not inner product spaces—of the Ruelle–Frobenius–Perron operator T: effectively, the transpose $K = T^\top$. Moreover, it has the same spectrum with only right and left swapping of the eigenvectors and generalized eigenvectors.

The Ruelle–Frobenius–Perron operator T is usually associated with the evolution of probability density, while the Koopman operator K is usually associated with the evolution of linear functionals of probability density. The duality of perspectives is associative in nature: $\langle f|\left(T^n\,|\rho_0\rangle\right)$ corresponds to the Ruelle–Frobenius–Perron perspective with T acting on the density ρ and $\left(\langle f|\,T^n\right)|\rho_0\rangle$ corresponds to the Koopman operator $T^\top = K$ acting on the observation function f. Allowing an observation vector $\vec{f} = [f_1, f_2, \ldots f_m]$ of linear functionals, and inspecting the most general form of K^n given by Eq. (25) together with the generalized eigenvector decomposition of the projection operators of Eq. (39), yields the most general form of the dynamics in terms of Koopman modes. Each Koopman mode is a length-m vector-valued functional of a Ruelle–Frobenius–Perron right eigenvector or generalized eigenvector.

Both approaches suffer when their operators are defective. Given the meromorphic calculus' ability to work around a wide class of such defects, adapting it the Ruelle–Frobenius–Perron and Koopman operators suggests that it may lift their decades-long restriction to only analyzing highly idealized (e.g., hyperbolic) chaotic systems.

2. *Eigenvalues from a time series*

Let's explore an additional benefit of this view of the Ruelle–Frobenius–Perron and Koopman operators, by proposing a novel method to extract the eigenvalues of a nominally nonlinear dynamic. Let $O_N(f,z)$ be (z^{-1} times) the z-transform [79, pp. 257–262] of a length-N sequence of τ-spaced type-f observations of a dynamical

system:

$$\begin{aligned} O_N(f,z) &\equiv z^{-1}\sum_{n=0}^{N} z^{-n}\,\langle f|T^n|\rho_0\rangle \\ &\to_{N\to\infty} \langle f|(zI-T)^{-1}|\rho_0\rangle \\ &= \sum_{\lambda\in\Lambda_T}\sum_{m=0}^{\nu_\lambda-1} \frac{\langle f|\,T_{\lambda,m}\,|\rho_0\rangle}{(re^{i\omega}-\lambda)^{m+1}} \ , \end{aligned}$$

as $N\to\infty$ for $|z|=r>1$. Note that $\langle f|T^n|\rho_0\rangle$ is simply the f-observation of the system at time $n\tau$, when the system started in state ρ_0. We see that this z-transform of observations automatically induces the resolvent of the hidden linear dynamic. If the process is continuous-time, then $T=e^{\tau G}$ implies $\lambda_T=e^{\tau\lambda_G}$, so that the eigenvalues should shift along the unit circle if τ changes; but the eigenvalues should be invariant to τ in the appropriate τ-dependent conformal mapping of the inside of the unit circle of the complex plane to the left half complex plane. Specifically, for any experimentally accessible choice of inter-measurement temporal spacing τ, the fundamental set of continuous-time eigenvalues Λ_G can be obtained from $\lambda_G=\frac{1}{\tau}\ln\lambda_T$, where each $\lambda_T\in\Lambda_T$ is extrapolated from $c/(re^{i\omega}-\lambda_T)^n$ curves fit to $O_N(f,re^{i\omega})$ for $c\in\mathbb{C}$, large N, and fixed r.

The square magnitude of $O_N(f,z)$ is related to the power spectrum generated by f-type observations of the system. Indeed, the power spectrum generated by any type of observation of a nominally nonlinear system is a direct fingerprint of the eigenspectrum and resolvent of the hidden linear dynamic. This suggests many opportunities for inferring eigenvalues and projection operators from frequency-domain transformations of a time series.

VII. CONCLUSION

The original, abstract spectral theory of normal operators rose to central importance when, in the early development of quantum mechanics, the eigenvalues of Hermitian operators were detected experimentally in the optical spectra of energetic transitions of excited electrons. We extended this powerful theory by introducing the meromorphic functional calculus, and unraveling the consequences of both the holomorphic and meromorphic functional calculi in terms of spectral projection operators and their associated left and right generalized eigenvectors. The result is a tractable spectral theory of *nonnormal* operators. Our straightforward examples suggest that the spectral properties of these general operators should also be experimentally accessible in the behavior of complex—open, strongly interacting—systems. We see a direct parallel with the success of the original spectral theory of normal operators as it made accessible the phenomena of the quantum mechanics of closed systems. This turns on nondiagonalizability and appreciating how ubiquitous it is.

Nondiagonalizability has consequences for settings as simple as counting, as shown in § VI C. Moreover, there we found that nondiagonalizability can be robust. The Drazin inverse, the negative-one power in the meromorphic functional calculus, is quite common in the nonequilibrium thermodynamics of open systems, as we showed in § VI D. Finally, we showed that the spectral character of nonnormal and nondiagonalizable operators manifests itself physically, as illustrated by Figs. 4c and 4d of § VI E. Our new formulae for spectral projection operators and the orthonormality relation among left and right generalized eigenvectors will thus likely find use in the analytic treatment of complex physical systems.

From the perspective of functional calculus, nonunitary time evolution, open systems, and non-Hermitian generators are closely related concepts since they all rely on the manipulation of nonnormal operators. Moreover, each domain is gaining traction. Nonnormal operators have recently drawn attention, from the nonequilibrium thermodynamics of nanoscale systems [80] to large-scale cosmological evolution [81]. In another arena entirely, complex directed networks [82] correspond to nonnormal and not-necessarily-diagonalizable weighted digraphs. There are even hints that nondiagonalizable network structures can be optimal for implementing certain dynamical functionalities [83]. The opportunity here should be contrasted with the well established field of spectral graph theory [84] that typically considers consequences of the spectral theorem for normal operators applied to the symmetric (and thus normal) adjacency matrices and Laplacian matrices. It seems that the meromorphic calculus and its generalized spectral theory will enable a *spectral weighted digraph theory* beyond the purview of current spectral graph theory.

Even if the underlying dynamic is diagonalizable, particular questions or particular choices of observable often *induce* a nondiagonalizable hidden linear dynamic. The examples already showed this arising from the simple imposition of counting or assuming a Poissonian dynamic. In more sophisticated examples, we recently found nondiagonalizable dynamic structures in quantum memory reduction [24], classical complexity measures [19], and prediction [22, 23].

Our goal has been to develop tractable, exact analytical techniques for nondiagonalizable systems. We did not discuss numerical implementation of algorithms that naturally accompany its practical application. Nevertheless, the theory does suggest new algorithms—for the Drazin inverse, projection operators, power spectra, and more.

Guided by the meromorphic calculus, such algorithms can be made robust despite the common knowledge that numerics with nondiagonalizable matrices is sensitive in certain ways.

The extended spectral theory we have drawn out of the holomorphic and meromorphic functional calculi complement efforts to address nondiagonalizability, e.g., via pseudospectra [85, 86]. It also extends and simplifies previously known results, especially as developed by Dunford [16]. Just as the spectral theorem for normal operators enabled much theoretical progress in physics, we hope that our generalized and tractable analytic framework yields rigorous understanding for much broader classes of complex system. Importantly, the analytic framework should enable a new *theory* of complex systems beyond the limited purview of numerical investigations.

While the infinite-dimensional theory is in principle readily obtained from the present framework, special care must be taken to guarantee a similar level of tractability and generality. Nevertheless, even the finite-dimensional theory enables a new level of tractability for analyzing not-necessarily-diagonalizable systems, including nonnormal dynamics. Future work will take full advantage of the operator theory, with more emphasis on infinite-dimensional systems and continuous spectra. Another direction forward is to develop creation and annihilation operators within nondiagonalizable dynamics. In the study of complex stochastic information processing, for example, this would allow analytic study of infinite-memory processes generated by, say, stochastic pushdown and counter automata [58, 87–89]. In a physical context, such operators may aid in the study of open quantum field theories. One might finally speculate that the Drazin inverse will help tame the divergences that arise there.

ACKNOWLEDGMENTS

JPC thanks the Santa Fe Institute for its hospitality. The authors thank John Mahoney, Sarah Marzen, Gregory Wimsatt, and Alec Boyd for helpful discussions. We especially thank Gregory Wimsatt for his assistance with § V B 3. This material is based upon work supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under contracts W911NF-12-1-0234, W911NF-13-1-0390, and W911NF-13-1-0340.

[1] A. Einstein. On the method of theoretical physics. *Philosophy of Science*, 1(2):163–169, April 1934. The Herbert Spencer Lecture, delivered at Oxford (10 June 1953). 1

[2] B. O. Koopman. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931. 2

[3] P. Gaspard, G. Nicolis, A. Provata, and S. Tasaki. Spectral signature of the pitchfork bifurcation: Liouville equation approach. *Phys. Rev. E*, 51:74–94, Jan 1995.

[4] M. Budii, R. Mohr, and I. Mezic. Applied Koopmanism. *Chaos*, 22(4), 2012. 2, 26

[5] N. Trefethen. Favorite eigenvalue problems. *SIAM News*, 44(10), Dec 2011. 2

[6] A. Sandage and G. A. Tammann. Steps toward the Hubble constant. VII-distances to NGC 2403, M101, and the Virgo cluster using 21 centimeter line widths compared with optical methods: The global value of H sub 0. *Astrophys. J.*, 210:7–24, 1976. 2

[7] A.G. Milnes. Semiconductor heterojunction topics: Introduction and overview. *Solid-State Electronics*, 29(2):99 – 121, 1986. 2

[8] P. A. M. Dirac. Theory of electrons and positrons. In *Nobel Lecture, Physics 1922–1941*. Elsevier Publishing Company, Amsterdam, 1965. 2

[9] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. 2

[10] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Opt. Soc. Am. A*, 4(3):519–524, Mar 1987. 2

[11] R. Courant and D. Hilbert. *Methods of mathematical physics: first English edition*, volume 1. Interscience Publishers, 1953. 2

[12] J. von Neumann. Zur algebra der funktionaloperationen und theorie der normalen operatoren. *Math. Annalen*, 102:370–427, 1930. 2

[13] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, New Jersey, 1955. 2

[14] S. Hassani. *Mathematical Physics*. Springer, New York, 1999. 2, 4, 7, 8

[15] R. R. Halmos. *Finite-Dimensional Vector Spaces*. D. Van Nostrand Company, 1958. 2, 3, 13

[16] N. Dunford. Spectral operators. *Pacific J. Math.*, 4(3):321–354, 1954. 3, 7, 28

[17] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM Press, Philadephia, Pennsylvannia, 2000. 3, 5, 14

[18] P. J. Antsaklis and A. N. Michel. *A Linear Systems Primer*. Springer Science & Business Media, New York, New York, 2007. 3

[19] J. P. Crutchfield, C. J. Ellison, and P. M. Riechers. Exact complexity: Spectral decomposition of intrinsic computation. *Phys. Lett. A*, 380(9-10):998–1002, 2015. 3, 18,

27

[20] P. M. Riechers, D. P. Varn, and J. P. Crutchfield. Diffraction patterns of layered close-packed structures from hidden Markov models. arxiv.org:1410.5028. 22, 24, 25

[21] P. M. Riechers, D. P. Varn, and J. P. Crutchfield. Pairwise correlations in layered close-packed structures. *Acta Cryst. A*, 71:423–443, 2015. 22

[22] P. M. Riechers and J. P. Crutchfield. Spectral simplicity of apparent complexity, Part I: The nondiagonalizable metadynamics of prediction. *arXiv:1705.08042*. 18, 27

[23] P. M. Riechers and J. P. Crutchfield. Spectral simplicity of apparent complexity, Part II: Exact complexities and complexity spectra. *arXiv:1706.00883*. 3, 18, 27

[24] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Minimized state complexity of quantum-encoded cryptic processes. *Phys. Rev. A*, 93:052317, May 2016. 3, 18, 27

[25] F. C. Binder, J. Thompson, and Mile Gu. A practical, unitary simulator for non-Markovian complex processes. arXiv.org:1709.02375. 3

[26] P. M. Riechers and J. P. Crutchfield. Fluctuations when driving between nonequilibrium steady states. *J. Stat. Phys.*, 168(4):873–918, 2017. 3, 20, 21

[27] A. B. Boyd, D. Mandal, P. M. Riechers, and J. P. Crutchfield. Transient dissipation and structural costs of physical information transduction. *Phys. Rev. Lett.*, 118:220602, 2017. 3

[28] B. Latni. *Signal Processing and Linear Systems*. Oxford University Press, New York, New York, 1998. 3, 7

[29] M. L. Boas. *Mathematical Methods in the Physical Sciences*, volume 2. Wiley and Sons, New York, New York, 1966. 3, 18

[30] N. Dunford. Spectral theory I. Convergence to projections. *Trans. Am. Math. Soc.*, 54(2):pp. 185–217, 1943. 4, 5

[31] M. Atiyah, R. Bott, and V. K. Patodi. On the heat equation and the index theorem. *Inventiones Mathematicae*, 19(4):279–330, 1973. 4

[32] C. S. Kubrusly. *Spectral theory of operators on Hilbert spaces*. Springer Science & Business Media, 2012. 4

[33] M. Haase. Spectral mapping theorems for holomorphic functional calculi. *J. London Math. Soc.*, 71(3):723–739, 2005. 5

[34] H. A. Gindler. An operational calculus for meromorphic functions. *Nagoya Math. J.*, 26:31–38, 1966. 6, 11

[35] B. Nagy. On an operational calculus for meromorphic functions. *Acta Math. Acad. Sci. Hungarica*, 33(3):379–390, 1979. 6, 11

[36] E. H. Moore. On the reciprocal of the general algebraic matrix. *Bull. Am. Math. Soc.*, 26, 1920. 10

[37] R. Penrose. A generalized inverse for matrices. *Math. Proc. Cambridge Phil. Soc.*, 51, 1955. 10

[38] A. Ben-Israel and T. N. E. Greville. *Generalized Inverses: Theory and Applications*. CMS Books in Mathematics. Springer, New York, New York, 2003. 10

[39] J. J. Koliha. A generalized Drazin inverse. *Glasgow Mathematical Journal*, 38(3):367–381, 1996. 10

[40] J. J. Koliha and T. D. Tran. The Drazin inverse for closed linear operators and the asymptotic convergence of C 0-semigroups. *J. Operator Th.*, pages 323–336, 2001. 10

[41] U. G. Rothblum. A representation of the Drazin inverse and characterizations of the index. *SIAM J. App. Math.*, 31(4):646–648, 1976. 10

[42] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*, volume 356. D. Van Nostrand, New York, New York, 1960. 11

[43] N. Dunford and J. T. Schwartz. *Linear Operators*. Interscience Publishers, New York, 1967. 11

[44] T. Bermúdez. Meromorphic functional calculus and local spectral theory. *Rocky Mountain J. Math.*, 29(2):437–447, 1999. 11

[45] J. J. Sakurai and J. J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, San Francisco, California, 2011. 13

[46] S. J. Axler. *Linear Algebra Done Right*, volume 2. Springer, New York, New York, 1997. 15

[47] M. Li and P. M. B. Vitanyi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York, 1993. 18

[48] J. P. Crutchfield. Between order and chaos. *Nature Physics*, 8(January):17–24, 2012. 18

[49] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney. Time's barbed arrow: Irreversibility, crypticity, and stored information. *Phys. Rev. Lett.*, 103(9):094101, 2009. 18

[50] C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield. Prediction, retrodiction, and the amount of information stored in the present. *J. Stat. Phys.*, 136(6):1005–1034, 2009. 18

[51] V. S. Barbu and N. Limnios. *Semi-Markov Chains and Hidden semi-Markov Models toward Applications: Their Use in Reliability and DNA Analysis*, volume 191. Springer, New York, 2008. 18, 20

[52] W. L. Smith. Renewal theory and its ramifications. *J. Roy. Stat. Soc. B*, 20(2):243–302, 1958.

[53] W. Gerstner and W. Kistler. Statistics of spike trains. In *Spiking Neuron Models*. Cambridge University Press, Cambridge, United Kingdom, 2002.

[54] F. Beichelt. *Stochastic Processes in Science, Engineering and Finance*. Chapman and Hall, New York, 2006. 18

[55] S. Marzen and J. P. Crutchfield. Informational and causal architecture of discrete-time renewal processes. *Entropy*, 17(7):4891–4917, 2015. 20

[56] S. Marzen and J. P. Crutchfield. Informational and causal architecture of continuous-time renewal processes. *J. Stat. Phys.*, 168(a):109–127, 2017.

[57] S. Marzen, M. R. DeWeese, and J. P. Crutchfield. Time resolution dependence of information measures for spiking neurons: Scaling and universality. *Front. Comput. Neurosci.*, 9:109, 2015.

[58] S. Marzen and J. P. Crutchfield. Statistical signatures of structural organization: The case of long memory in renewal processes. *Phys. Lett. A*, 380(17):1517–1525, 2016. 20, 28

[59] S. Marzen and J. P. Crutchfield. Structure and randomness of continuous-time discrete-event processes. *J. Stat. Physics*, 169(2):303–315, 2016. 20

[60] M. S. Green. Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids. *J. Chem. Physics*, 22(3):398–413, 1954. 21

[61] R. Zwanzig. Time-correlation functions and transport coefficients in statistical mechanics. *Ann. Rev. Phys. Chemistry*, 16(1):67–102, 1965. 21

[62] D. Mandal and C. Jarzynski. Analysis of slow transitions between nonequilibrium steady states. *arXiv:1507.06269*. 21

[63] S. Lahiri, J. Sohl-Dickstein, and S. Ganguli. A universal tradeoff between power, precision and speed in physical communication. *arXiv:1603.07758*. 21

[64] P. Löwdin. On operators, superoperators, Hamiltonians, and Liouvillians. *Intl. J. Quant. Chem.*, 22(S16):485–560, 1982. 21

[65] G. Lindblad. On the generators of quantum dynamical semigroups. *Comm. Math. Physics*, 48(2):119–130, 1976. 21

[66] S. M. Barnett and S. Stenholm. Spectral decomposition of the Lindblad operator. *J. Mod. Optics*, 47(14-15):2869–2882, 2000. 21

[67] T. Petrosky and I. Prigogine. The Liouville space extension of quantum mechanics. *Adv. Chem. Phys*, 99:1–120, 1997. 21

[68] Y. Oono and M. Paniconi. Steady state thermodynamics. *Prog. Theo. Phys. Supp.*, 130:29–44, 1998. 22

[69] P. Stoica and R. L. Moses. *Spectral Analysis of Signals*. Pearson Prentice Hall, Upper Saddle River, New Jersey, 2005. 22

[70] H. Stark, W. R. Bennett, and M. Arm. Design considerations in power spectra measurements by diffraction of coherent light. *Appl. Opt.*, 8(11):2165–2172, Nov 1969. 22

[71] R. G. James, J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. Many roads to synchrony: Natural time scales and their algorithms. *Phys. Rev E*, 89:042135, 2014. 23

[72] E. Narevicius, P. Serra, and N. Moiseyev. Critical phenomena associated with self-orthogonality in non-Hermitian quantum mechanics. *EPL (Europhys. Lett.)*, 62(6):789, 2003. 25

[73] A. V. Sokolov, A. A. Andrianov, and F. Cannata. Non-Hermitian quantum mechanics of non-diagonalizable Hamiltonians: Puzzles with self-orthogonal states. *J. Physics A*, 39(32):10207, 2006.

[74] A. Mostafazadeh. Spectral singularities of complex scattering potentials and infinite reflection and transmission coefficients at real energies. *Phys. Rev. Lett.*, 102:220402, 2009. 25

[75] J. D. Farmer, J. P. Crutchfield, H. Froehling, N. H. Packard, and R. S. Shaw. Power spectra and mixing properties of strange attractors. *Ann. New York Acad. Sci.*, 357:453, 1980. 26

[76] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20(2):130–141, 1963. 26

[77] D. Ruelle and F. Takens. On the nature of turbulence. *Comm. Math. Physics*, 20(3):167–192, 1971. 26

[78] M. C. Mackey. *Time's Arrow: The Origins of Thermodynamic Behavior*. Springer, New York, 1992. 26

[79] R. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, third edition, 1999. 26

[80] B. Gardas, S. Deffner, and A. Saxena. Non-hermitian quantum thermodynamics. *Sci. Reports*, 6:23408, 2016. 27

[81] N. Berkovits and D. Witten. Conformal supergravity in twistor-string theory. *J. High Energy Physics*, 2004(08):009, 2004. 27

[82] M. Newman. *Networks: an introduction*. Oxford University Press, Oxford, United Kingdom, 2010. 27

[83] T. Nishikawa and A. E. Motter. Synchronization is optimal in nondiagonalizable networks. *Phys. Rev. E*, 73:065106, Jun 2006. 27

[84] F. R. K. Chung. *Spectral Graph Theory*, volume 92. American Mathematical Soc., Providence, Rhode Island, 1997. 27

[85] L. N. Trefethen. Pseudospectra of linear operators. *SIAM Review*, 39(3):383–406, 1997. 28

[86] L. N. Trefethen and M. Embree. *Spectra and pseudospectra: The behavior of nonnormal matrices and operators*. Princeton University Press, Princeton, New Jersey, 2005. 28

[87] J. P. Crutchfield and K. Young. Computation at the onset of chaos. In W. Zurek, editor, *Entropy, Complexity, and the Physics of Information*, volume VIII of *SFI Studies in the Sciences of Complexity*, pages 223 – 269, Reading, Massachusetts, 1990. Addison-Wesley. 28

[88] N. Travers and J. P. Crutchfield. Infinite excess entropy processes with countable-state generators. *Entropy*, 16:1396–1413, 2014.

[89] J. P. Crutchfield and S. E. Marzen. Signatures of infinity: Nonergodicity and resource scaling in prediction, complexity, and learning. *Phys. Rev. E*, 91:050106, 2015. 28