Geostat lab seminars

Article presentation

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

by authors: P. M. Riechers & J. P. Crutchfield

H. Yahia, INRIA, September 14th, 2018
1

Article’s
Introduction & Motivation
For any operator $A$ (finite dimensional, i.e. a matrix, or non finite dimensional) define consistently $f(A)$ where

$$f : \mathbb{C} \longrightarrow \mathbb{C}$$

is a function holomorphic or even meromorphic whose poles can interact with $A$’s spectrum. Examples: $e^A$, $\log(A)$ even $A^{-1}$ for non invertible operators (”Drazin inverse”) and develop applications in physics, signal processing and machine learning.
Importance of spectral decomposition in Physics & Machine Learning

- **Spectral decomposition**: splitting a linear operator into independent modes.
Importance of spectral decomposition in Physics & Machine Learning

- **Spectral decomposition**: splitting a linear operator into independent modes.
- Describes the organization of large-scale systems, with predictive theories.
Importance of spectral decomposition in Physics & Machine Learning

- **Spectral decomposition**: splitting a linear operator into independent modes.
- Describes the organization of large-scale systems, with predictive theories.
- Example in quantum mechanics & statistical mechanics: identify the energies eigenvalues of Hamiltonians as basic objects in thermodynamics, transitions among the energy states → heat and work.

Applications and discoveries enabled by spectral decompositions yield a long list including LEDs etc.

Non-linearities mapped to linearities in high dimensional spaces: example of Machine Learning through SVM etc.
Importance of spectral decomposition in Physics & Machine Learning

- **Spectral decomposition**: splitting a linear operator into independent modes.
- Describes the organization of large-scale systems, with predictive theories.
- Example in quantum mechanics & statistical mechanics: identify the energies eigenvalues of Hamiltonians as basic objects in thermodynamics, transitions among the energy states \(\rightarrow\) heat and work.
- Applications and discoveries enabled by spectral decompositions yield a long list including LEDs etc.
Importance of spectral decomposition in Physics & Machine Learning

- **Spectral decomposition**: splitting a linear operator into independent modes.
- Describes the organization of large-scale systems, with predictive theories.
- Example in quantum mechanics & statistical mechanics: identify the energies eigenvalues of Hamiltonians as basic objects in thermodynamics, transitions among the energy states $\rightarrow$ heat and work.
- Applications and discoveries enabled by spectral decompositions yield a long list including LEDs etc.
- Non-linearities mapped to linearities in high dimensional spaces: example of Machine Learning through SVM etc.
A little reminder
Spectrum of an operator

Complex Banach space $X$. If $\phi : X \rightarrow X$ is a continuous bijective linear map, i.e. $\phi$ is bounded and has an inverse $\phi^{-1}$, then it can be shown that $\phi^{-1}$, which is obviously linear, is also bounded.

Definition (Spectrum)
The spectrum of $\phi$ is the set

$$\text{sp}(\phi) = \{ \lambda \in \mathbb{C} \mid \lambda 1_X - \phi \text{ is not invertible} \}.$$  \hspace{1cm} (1)
Spectrum of an operator

If $X$ is finite dimensional,

$$\text{sp}(\phi) = \{ \lambda \in \mathbb{C} \mid \det(\lambda 1_X - \phi) = 0 \}$$  \hspace{1cm} (2)

is a finite set of values $\subset \mathbb{C}$ (eigenvalues) but if $X$ is not finite dimensional, the situation is considerably more complex.
The resolvent set $\rho(\phi)$ is the spectrum’s complementary in $\mathbb{C}$:

$$\rho(\phi) = \{ \lambda \in \mathbb{C} \mid \lambda 1_X - \phi \text{ is invertible} \}.$$ In non-finite dimensional vector spaces, as it is the case in signal processing, the non-invertible character of $\lambda 1_X - \phi$ can come from several different reasons:
Spectrum of an operator

1. \( \ker((\lambda I - \phi)) \neq \{0\} \), i.e. there exists \( x \in X, x \neq 0 \) such that \( \phi(x) = \lambda x \): \( x \) is an eigenvector associated to \( \lambda \), which is then an eigenvalue. If \( X \) has finite dimension, this is the only case possible. The subset \( \text{sp}_p(\phi) \subset \text{sp}(\phi) \) of eigenvalues is called the punctual spectrum of \( \phi \).

2.

2.1

2.2
Spectrum of an operator

1. \( \text{Ker}(\lambda \mathbb{1}_X - \phi) \neq \{0\} \), i.e. there exists \( x \in X, x \neq 0 \) such that \( \phi(x) = \lambda x \): \( x \) is an **eigenvector** associated to \( \lambda \), which is then an **eigenvalue**. If \( X \) has finite dimension, this is the only case possible. The subset \( \text{sp}_p(\phi) \subset \text{sp}(\phi) \) of eigenvalues is called the **punctual spectrum** of \( \phi \).

2. \( \text{Ker}(\lambda \mathbb{1}_X - \phi) = \{0\} \), however \( \text{Im}(\lambda \mathbb{1}_X - \phi) \neq X \), i.e. \( \lambda \mathbb{1}_X - \phi \) is injective but not surjective. Two subcases:
   2.1
   2.2
Spectrum of an operator

1. Ker(λ1_X − φ) \neq \{0\}, i.e. there exists \( x \in X, x \neq 0 \) such that \( \phi(x) = \lambda x \): \( x \) is an **eigenvector** associated to \( \lambda \), which is then an **eigenvalue**. If \( X \) has finite dimension, this is the only case possible. The subset \( \text{sp}_p(\phi) \subset \text{sp}(\phi) \) of eigenvalues is called the **punctual spectrum** of \( \phi \).

2. Ker(λ1_X − φ) = \{0\}, however \( \text{Im}(\lambda1_X - \phi) \neq X \), i.e. \( \lambda1_X - \phi \) is injective but not surjective. Two subcases:
   2.1 Ker(λ1_X − φ) = \{0\}, \( \text{Im}(\lambda1_X - \phi) \) not dense in \( X \),
   \[
   \text{sp}_{\text{res}}(\phi) = \{x \in X | \text{Ker}(\lambda1_X-\phi) = 0\} \text{ and } \text{Im}(\lambda1_X - \phi) \neq X
   \]
   2.2
Spectrum of an operator

1. $\ker(\lambda \mathbf{1}_X - \phi) \neq \{0\}$, i.e. there exists $x \in X$, $x \neq 0$ such that $\phi(x) = \lambda x$: $x$ is an eigenvector associated to $\lambda$, which is then an eigenvalue. If $X$ has finite dimension, this is the only case possible. The subset $\sp_p(\phi) \subset \sp(\phi)$ of eigenvalues is called the punctual spectrum of $\phi$.

2. $\ker(\lambda \mathbf{1}_X - \phi) = \{0\}$, however $\text{Im}(\lambda \mathbf{1}_X - \phi) \neq X$, i.e. $\lambda \mathbf{1}_X - \phi$ is injective but not surjective. Two subcases:
   2.1 $\ker(\lambda \mathbf{1}_X - \phi) = \{0\}$, $\text{Im}(\lambda \mathbf{1}_X - \phi)$ not dense in $X$,

   $$\sp_{\text{res}}(\phi) = \{x \in X \mid \ker(\lambda \mathbf{1}_X - \phi) = 0\} \text{ and } \overline{\text{Im}(\lambda \mathbf{1}_X - \phi)} \neq X$$

   is the residual spectrum of $\phi$.
   2.2
**Spectrum of an operator**

1. \( \text{Ker}(\lambda 1_X - \phi) \neq \{0\} \), *i.e.* there exists \( x \in X, x \neq 0 \) such that \( \phi(x) = \lambda x \): \( x \) is an **eigenvector** associated to \( \lambda \), which is then an **eigenvalue**. If \( X \) has finite dimension, this is the only case possible. The subset \( \text{sp}_p(\phi) \subset \text{sp}(\phi) \) of eigenvalues is called the **punctual spectrum** of \( \phi \).

2. \( \text{Ker}(\lambda 1_X - \phi) = \{0\} \), however \( \text{Im}(\lambda 1_X - \phi) \neq X \), *i.e.* \( \lambda 1_X - \phi \) is injective but not surjective. Two subcases:
   2.1 \( \text{Ker}(\lambda 1_X - \phi) = \{0\} \), \( \text{Im}(\lambda 1_X - \phi) \) not dense in \( X \),
   
   \[ \text{sp}_{\text{res}}(\phi) = \{ x \in X \mid \text{Ker}(\lambda 1_X - \phi) = 0 \} \text{ and } \text{Im}(\lambda 1_X - \phi) \neq X \} \]

   is the **residual spectrum** of \( \phi \).

   2.2 \( \text{Ker}(\lambda 1_X - \phi) = \{0\} \), \( \text{Im}(\lambda 1_X - \phi) \) dense in \( X \),
   
   \[ \text{sp}_{\text{cont}}(\phi) = \{ x \in X \mid \text{Ker}(\lambda 1_X - \phi) = 0 \} \text{ and } \text{Im}(\lambda 1_X - \phi) = X \} \]
Spectrum of an operator

1. Ker($\lambda 1_X - \phi$) $\neq \{0\}$, i.e. there exists $x \in X$, $x \neq 0$ such that $\phi(x) = \lambda x$: $x$ is an eigenvector associated to $\lambda$, which is then an eigenvalue. If $X$ has finite dimension, this is the only case possible. The subset $\text{sp}_p(\phi) \subset \text{sp}(\phi)$ of eigenvalues is called the punctual spectrum of $\phi$.

2. Ker($\lambda 1_X - \phi$) = $\{0\}$, however $\text{Im}(\lambda 1_X - \phi) \neq X$, i.e. $\lambda 1_X - \phi$ is injective but not surjective. Two subcases:

2.1 Ker($\lambda 1_X - \phi$) = $\{0\}$, $\text{Im}(\lambda 1_X - \phi)$ not dense in $X$,

$$\text{sp}_{\text{res}}(\phi) = \{x \in X \mid \text{Ker}(\lambda 1_X - \phi) = 0\} \text{ and } \text{Im}(\lambda 1_X - \phi) \neq X$$

is the residual spectrum of $\phi$.

2.2 Ker($\lambda 1_X - \phi$) = $\{0\}$, $\text{Im}(\lambda 1_X - \phi)$ dense in $X$,

$$\text{sp}_{\text{cont}}(\phi) = \{x \in X \mid \text{Ker}(\lambda 1_X - \phi) = 0\} \text{ and } \text{Im}(\lambda 1_X - \phi) = X$$

is the continuous spectrum of $\phi$. 
Spectrum of an operator

Consequently there is the partition:

$$\text{sp}(\phi) = \text{sp}_p(\phi) \cup \text{sp}_{\text{res}}(\phi) \cup \text{sp}_{\text{cont}}(\phi).$$

However, in the case of self-adjoint operators, the residual spectrum is empty. Let us denote by $\mathcal{L}(X)$ the space of bounded linear operators on the Banach space $X$.

**Theorem**

*Let $\phi \in \mathcal{L}(X)$.*

1. *The spectrum $\text{sp}(\phi)$ is a closed non-empty subset of $\mathbb{C}$.*
2. *The spectrum $\text{sp}(\phi)$ is bounded: $\text{sp}(\phi) \subset \{0 \leq |z| \leq \|\phi\|\}$.*
3. *The map defined on $\rho(\phi) = \mathbb{C} - \text{sp}(\phi)$ by $\lambda \mapsto (\lambda \mathbb{1}_X - \phi)^{-1}$ is analytic.*
Spectrum of an operator

Consequently, the resolvent set of φ is open and \( \sigma_p(\phi) \) is a compact subset of \( \mathbb{C} \). We give here an example without the proofs and computations, which can be found in functional analysis literature. On the Banach space \( X = \ell^1(\mathbb{N}) \), consider the continuous linear map

\[
\phi : \ell^1(\mathbb{N}) \longrightarrow \ell^1(\mathbb{N})
\]

\[
x = (x_1, x_2, \ldots) \longmapsto \phi(x) = (x_2, x_3, \ldots)
\]

\( \phi \) is linear and continuous, \( \|\phi\| = 1 \). Remembering that the topological dual space of \( X \) is identified with the space \( X' \sim \ell^\infty(\mathbb{N}) \), \( \phi' \)'s topological adjoint is the continuous linear map

\[
\phi' : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N})
\]

\[
x = (x_1, x_2, \ldots) \longmapsto \phi'(x) = (0, x_1, x_2, x_3, \ldots).
\]
Spectrum of an operator

Then one can show that:

- \( \text{sp}(\phi) = \{ 0 \leq |z| \leq 1 \} \),
- \( \text{sp}_p(\phi) = \{ 0 \leq |z| < 1 \} \),
- \( \text{sp}_{\text{res}}(\phi) = \emptyset \),
- \( \text{sp}_{\text{cont}}(\phi) = \{|z| = 1\} \),

and

- \( \text{sp}(\phi') = \{ 0 \leq |z| \leq 1 \} \),
- \( \text{sp}_p(\phi') = \emptyset \),
- \( \text{sp}_{\text{res}}(\phi') = \{ 0 \leq |z| < 1 \} \),
- \( \text{sp}_{\text{cont}}(\phi') = \emptyset \).
Spectrum of an operator

Consequently there is the partition:

\[ \text{sp}(\phi) = \text{sp}_p(\phi) \cup \text{sp}_{\text{res}}(\phi) \cup \text{sp}_{\text{cont}}(\phi). \]

**Theorem**

*Let* \( \phi \in \mathcal{L}(X) \).

1. *The spectrum* \( \text{sp}(\phi) \) *is a closed non-empty subset of* \( \mathbb{C} \).
2. *The spectrum* \( \text{sp}(\phi) \) *is bounded*: \( \text{sp}(\phi) \subset \{0 \leq |z| \leq \|\phi\|\} \).
3. *The map defined on* \( \rho(\phi) = \mathbb{C} - \text{sp}(\phi) \) *by*
   \[ \lambda \mapsto (\lambda \mathbb{1}_X - \phi)^{-1} \] *is analytic.*

The last sentence invites us to recall and enlarge the elementary definition of analyticity to vector valued functions.
Analytic functions

The complex variable calculus (definition of analytic functions, Cauchy formula, theorem of residues etc.) defined usually for functions

\[ f : \mathbb{C} \rightarrow \mathbb{C} \]

\[ f(z) = \sum_{n \geq 0} a_n(z - z_0)^n, \quad a_n \in \mathbb{C} \]

is extended, without any modifications, to functions

\[ f : \mathbb{C} \rightarrow X \]

where \( X \) is any complex Banach space, in particular \( X = \mathcal{L}(E) \), a space of operators.
Analytic functions

In particular, if \( A \) is any matrix or any operator and \( \lambda \) an eigenvalue of \( A \), the residue

\[
A_\lambda = \text{Res}((zI - A)^{-1}), \lambda
\]

is no more a complex number, but an operator, called a projection operator. They are orthonormal:

\[
A_\lambda A_\xi = \delta_{\lambda \xi} A_\lambda
\]

and sum to \( I \), the identity (when \( A \) is a matrix but can be extended):

\[
I = \sum_{\lambda \in \text{sp}(A)} A_\lambda.
\]
The finite dimensional case

In the finite dimensional case ($A$ is a matrix) the spectrum is

$$\Lambda_A = \{ \lambda \in \mathbb{C} \mid \det(\lambda I - A) = 0 \}$$

Elements of $\Lambda_A$ are roots of the characteristic polynomial

$$P_A(X) = (X - \lambda_1)^{a_1} \cdots (X - \lambda_p)^{a_p}$$

$A$’s minimal polynomial is

$$Q_A(X) = (X - \lambda_1)^{\nu_1} \cdots (X - \lambda_p)^{\nu_p}$$

$\nu_\lambda$ is the index of eigenvalue $\lambda$. The geometric multiplicity is

$$g_\lambda = \dim(\text{Ker}(\lambda I - A)) \leq a_\lambda.$$
3

Meromorphic

functional calculus
The easy case: holomorphic functional calculus

When

\[ f : \mathbb{C} \rightarrow \mathbb{C} \]

is holomorphic

\[
\begin{align*}
f(A) &= \frac{1}{2\pi i} \int_{C_{\Lambda}} f(z)(zI - A)^{-1} \, dz \\
&= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_{\lambda} - 1} \frac{f^{(m)}(\lambda)}{m!} (A - \lambda I)^m A_{\lambda}
\end{align*}
\]

(the proof is not easy).
The easy case: holomorphic functional calculus

Figure: Holomorphic case.
The easy case: holomorphic functional calculus

\[ f(A) = \frac{1}{2\pi i} \int_{C_{\Lambda A}} f(z)(zl - A)^{-1} \, dz \]

\[ = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu\lambda - 1} \frac{f^{(m)}(\lambda)}{m!} (A - \lambda I)^m A_\lambda \]

Application: \( \log(A) \) easily defined, with \( e^{\log(A)} = A \).
Hypothesis: countable point spectrum, \( f \) meromorphic with poles at certain points of \( \Lambda_A \).

\[
f(A) = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \int_{C_\lambda} f(z)(zl - A)^{-1} \, dz
\]

\[
= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_{\lambda}-1} A_{\lambda}(A - \lambda l)^m \frac{1}{2\pi i} \int_{C_\lambda} \frac{f(z)}{(z - \lambda)^{m+1}} \, dz
\]

\( C_\lambda \) small contour around \( \lambda \in \Lambda_A \subset \mathbb{C} \).
Meromorphic case

Figure: Meromorphic case.
Meromorphic case

\[ f(A) = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \int_{C_\lambda} f(z)(zI-A)^{-1} \, dz = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \int_{C_\lambda} f(z)R(z, A) \, dz \]

\[ R(z, A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{a_\lambda-1} \frac{1}{(z-\lambda)^{m+1}} A_\lambda^m \]

\[ A_\lambda^m \text{ matrices with constant entries (without } z \text{). One has} \]

\[ A_\lambda = A_\lambda^0 = \frac{1}{2\pi i} \int_{C_\lambda} R(z, A) \, dz. \]
Some examples

\[ f(z) = z^0 = 1 \quad \rightarrow \quad I = \sum_{\lambda \in \Lambda_A} A_\lambda \]

\[ f(z) = z \quad \rightarrow \quad A = \sum_{\lambda \in \Lambda_A} (\lambda A_{\lambda 0} + A_{\lambda 1}) \quad \text{Jordan decomposition} \]

If \( A \) is diagonalizable

\[ A = \sum_{\lambda \in \Lambda_A} \lambda A_\lambda. \]

One can show: \( A_{\lambda m} = A_{\lambda}(A - \lambda I)^m \).
\[ f(z) = z^L, \quad L > 0 \]

If \( L \geq \nu\lambda - 1 \) for all \( \lambda \in \Lambda_A \)

\[
A^L = \sum_{\lambda \in \Lambda_A} \left( \sum_{\lambda \neq 0}^{\nu\lambda - 1} \binom{L}{m} \lambda^{L-m} A_{\lambda m} \right)
\]
\( f(z) = z^L, \ L < 0: \text{ Drazin inverse} \)

\[
A^{-|L|} = \sum_{\lambda \in \Lambda_A} \sum_{\lambda \neq 0} \sum_{m=0}^{\nu_{\lambda} - 1} (-1)^m \binom{|L| + m - 1}{m} \lambda^{-|L| - m} A_{\lambda,m}
\]

the Drazin inverse corresponds to \( L = -1 \) (denoted \( A^D \)). It is defined even when \( A \) is not invertible, and is not equal to the Moore-Penrose pseudo-inverse. If \( A \) is invertible, \( A^D = A^{-1} \). If \( T \) is a stochastic matrix \( (I - T + (I - T_0))^{-1} = (I - T)^D + (I - T_0) \) and \( (I - T + T_1)^{-1} = \) fundamental matrix.
4

Computation of projection operators
An unpublished formula

When $\nu_\lambda = 1$

$$A_\lambda = \prod_{\xi \in A, \xi \neq \lambda} \left( \frac{A - \xi I}{\lambda - \xi} \right)^{\nu_\xi}$$

so when the operator is diagonalizable

$$A_\lambda = \prod_{\xi \in A, \xi \neq \lambda} \left( \frac{A - \xi I}{\lambda - \xi} \right)$$
Normal operators

These are the operators such that $AA^* = A^*A$. Examples: Hermitian, orthogonal, symmetric, skew-symmetric etc. For such operators there is a decomposition $A = U\Lambda_A \bar{U}^t$ with $U$ orthogonal. Supposing $\Lambda_A = \{\lambda_1, \cdots, \lambda_M\}$,

$$A_\lambda = \sum_{j=1}^{M} \delta_{\lambda,\lambda_j} u_j \otimes \bar{u}_j.$$
Diagonalizable operators

Using bra-ket notation

\[ A_\lambda = \sum_{j=1}^{M} \delta_{\lambda,\lambda_j} |\lambda_j\rangle \langle \lambda_j| \]
Any matrix

Use Jordan decomposition and apply the previous on each Jordan block.
5

Applications
Poisson point processes

Figure: Explicit Markov-chain representation of a truncated Poisson process.
Poisson point processes

We consider a truncated Poisson process at step $N$ and let $N \to +\infty$. The transition matrix between states is

$$G = \begin{pmatrix}
-r & r \\
-r & r \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
-r & r \\
r & r \\
\end{pmatrix}$$

$g_{ij} = \text{transition to state } j \text{ given that the system is in state (count) } i$. We have $\Lambda_G = \{-r\}$, $a_{-r} = N + 1$, $g_{-r} = 1$, $\nu_{-r} = N + 1$ and $G_{-r} = I$: the Poisson process is non-diagonalizable.
Poisson point processes

The net counter state-to-state transition operator from initial time 0 to time $t$ is

$$T(t) = e^{tG} = \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_{\lambda} - 1} G_\lambda (G - \lambda I)^m \left( \frac{1}{2\pi i} \int_{C_\lambda} \frac{e^{tz}}{(z - \lambda)^m} \, dz \right).$$

Develop and compute... and retrieve easily the probability that the counter is incremented by $n$ in a time interval $t$ is independent of the initial count and is given by

$$(rt)^n e^{-rt} \frac{n!}{n!}.$$

This simple result is obtained directly rather than a limit of a binomial distribution.
Inhomogeneous Poisson process

Generalization of the previous model to time inhomogeneous Poisson process i.e. the transition rate $r$ no more constant but a function of time: $r(t)$. The net counter state-to-state transition from time $t_0$ to time $t_f$ is:

$$T_{t_0,t_f} = e^{\int_{t_0}^{t_f} G(t) dt}$$

Applying the meromorphic calculus, one gets the probability that the count is incremented by $n$ during the time interval $\Delta t$ is

$$\frac{(\langle r \rangle \Delta t)^n e^{-\langle r \rangle \Delta t}}{n!},$$

an important result in Poisson point processes.
Stochastic thermodynamics

Consider a singular operator $\mathcal{L}$ that is not necessarily normal nor diagonalizable. Via the meromorphic calculus

$$\int_0^\tau e^{t\mathcal{L}} dt = \sum_{\lambda \in \Lambda_{\mathcal{L}}} \sum_{m=0}^{\nu_{\lambda}-1} \mathcal{L}_{\lambda,m} \frac{1}{2\pi i} \int_{C_{\lambda}} \frac{\int_0^\tau e^{tz} dt}{(z - \lambda)^{m+1}} dz$$

$$= \left( \sum_{m=0}^{\nu_0-1} \frac{\tau^{m+1}}{(m+1)!} \mathcal{L}_{0,m} \right) + \mathcal{L}^D (e^{\tau \mathcal{L} - I}).$$

When $\mathcal{L}$ is a singular operator called $G$ that is a transition rate operator exhibiting a single stationary distribution the previous equation simplifies to

$$\int_0^\tau e^{tG} dt = \tau |0_G\rangle \langle 0_G| + G^D (e^{\tau G - I}).$$
**Stochastic thermodynamics**

Since $G$ is a transition-rate operator, the above integral corresponds to integrated time evolution. The Drazin inverse $G^D$ concentrates on the transient contribution beyond the persistent stationary background. The Green-Kubo relation for a transport coefficient $\kappa$ is:

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

(s.s. = steady state) where $A(0)$ and $A(t)$ are some observable of the stationary stochastic dynamical system at time 0 and time $t$ respectively. Using the previous equation, one finds:

$$\kappa = -\langle AG^D A \rangle_{\text{s.s.}}.$$
Stochastic thermodynamics

In the special case where the transition operator is diagonalizable one finds

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

Let \((X_n)_{n \in \mathbb{N} - \{0\}}\) be a signal and

\[
P(\omega) = \lim_{N \to +\infty} \frac{1}{N} \langle | \sum_{n=1}^{N} X_n e^{-i\omega n} |^2 \rangle
\]

its power spectrum. If the process is wide-sense stationary, the power spectrum is obtained from the signal’s autocorrelation \(\gamma(\tau)\):

\[
P(\omega) = \lim_{N \to +\infty} \frac{1}{N} \sum_{\tau=-N}^{N} (N - |\tau|) \gamma(\tau) e^{-i\omega \tau}.
\]

Question: how to compute the power spectrum given a model of the signal’s generator in the form of a Hidden Markov model?
Power spectra

Figure: Bayes network for a state-emitting hidden Markov model graphically depicts the structure of conditional independence among random variables for the latent state \((S_n)_{n \in \mathbb{Z}}\) and the random variables \((X_n)_{n \in \mathbb{Z}}\) for the observation at each time \(n\).
State-emitting hidden Markov model (HMM) $\mathcal{M} = (S, A, \mathcal{P}, T)$ that generates the stationary stochastic process $(X_n)_{n \in \mathbb{Z}}$ according to the following: $S$ is the finite set of latent states of the hidden Markov chain and $A \in \mathbb{C}$ is the observable alphabet. $S_t$ is the random variable for the hidden state at time $t$ that takes on values $s \in S$. $X_t$ is the random variable for the observation at time $t$ that takes on values $x \in A$. Given the latent state at time $t$, the possible observations are distributed according to the conditional probability density functions: $\mathcal{P} = (p(X_t = x \mid S_t = s))_{s \in S}$, abbreviated $p(x \mid s)$ 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 $T_{i,j} = \Pr(S_{t+1} = s_j \mid S_t = s_i)$ which give the probability of transitions from latent state $s_i$ to $s_j$. 
Power spectra

Using meromorphic calculus, the authors show that if $\Omega$ is the $|S| \times |S|$ matrix

$$\Omega = \sum_{s \in S} \langle x \rangle p(x|s) \delta_s \langle \delta_s | T$$

the continuous part of the power spectrum is

$$P_c(\omega) = \langle |x|^2 \rangle + \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu-1} 2\text{Re} \frac{\langle \pi | \tilde{\Omega} T_{\lambda,m} \Omega | 1 \rangle}{(e^{i\omega} - \lambda)^{m+1}}$$

with $\langle \pi | = \langle 1_T |$ is the stationary distribution induced by $T$ over the latent states and $|1\rangle = |1_T\rangle$ is a column vector of all ones. Note that $\langle \pi | \delta_s \rangle = \Pr(s)$. 
The discrete portion of the spectrum is

\[ P_d(\omega) = \sum_{k=-\infty}^{+\infty} \sum_{\lambda \in \Lambda_T, |\lambda|=1} 2\pi \delta(\omega - \omega_\lambda + 2k\pi) \text{Re}(\lambda^{-1} \langle \pi | \bar{\Omega} T_\lambda \Omega | 1 \rangle) \]

where \( \omega_\lambda \) is related to \( \lambda \) 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 \( \nu_\lambda = 1 \).