

Phys 6124 zipped!

World Wide Quest to Tame Math Methods

Han Liang and Predrag Cvitanović

August 25, 2019

mathematical methods - week 1

Linear algebra

Georgia Tech PHYS-6124

Homework HW #1

due Monday, August 26, 2019

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort
== if you are LaTeXing, here is the [source code](#)

Exercise 1.1 <i>Trace-log of a matrix</i>	4 points
Exercise 1.2 <i>Stability, diagonal case</i>	2 points
Exercise 1.3 <i>Time-ordered exponentials</i>	4 points

Total of 10 points = 100 % score. Extra points accumulate, can help you later if you miss a few problems.

edited September 9, 2019

Week 1 syllabus

Monday, August 19, 2019

- Sect. 1.2 *Matrix-valued functions*
[Grigoriev notes](#) pages 6.1, 6.2 (skip 1. *Mechanics inertia tensor*), 6.4, 6.5 up to “right and left eigenvectors.”
- sect. 1.3 *A linear diversion*
 There are two representations of exponential of constant matrix, the Taylor series and the compound interest (Euler product) formulas (1.23). If the matrix (for example, the Hamiltonian) changes with time, the exponential has to be time-ordered. The Taylor series generalizes to the nested integrals formula (1.50), and the Euler product to time-ordered product (1.24). The first is used in formal quantum-mechanical calculations, the second in practical, numerical calculations.
- sect. 1.4 *Eigenvalues and eigenvectors*
 Hamilton-Cayley equation, projection operators (1.34), any matrix function is evaluated by spectral decomposition (1.37). Work through example 1.3.
- Optional reading: Stone & Goldbart [Appendix A](#); Arfken & Weber Arfken and Weber [1] Chapter 3
- Feel free to add your topics suggestions at any time: [suggestions by students and faculty](#).
- [you got a new idea?](#)

1.1 Literature

The subject of linear algebra generates innumerable tomes of its own, and is way beyond what we can exhaustively cover. We have added to the [course homepage](#) linear operators and matrices reading: Stone and Goldbart [9], *Mathematics for Physics: A Guided Tour for Graduate Students*, [Appendix A](#). This is an advanced summary where you will find almost everything one needs to know. More pedestrian and perhaps easier to read is Arfken and Weber [1] *Mathematical Methods for Physicists: A Comprehensive Guide*.

1.2 Matrix-valued functions

What is a matrix?

—Werner Heisenberg (1925)

What is the matrix?

—Keanu Reeves (1999)

Why should a working physicist care about linear algebra? Physicists were blissfully ignorant of group theory until 1920's, but with Heisenberg's sojourn in Helgoland, everything changed. Quantum Mechanics was formulated as

$$\phi(t) = \hat{U}^t \phi(0), \quad \hat{U}^t = e^{-\frac{i}{\hbar} t \hat{H}}, \quad (1.1)$$

where $\phi(t)$ is the quantum wave function t , \hat{U}^t is the unitary quantum evolution operator, and \hat{H} is the Hamiltonian operator. Fine, but what does this equation *mean*? In the first lecture we deconstruct it, make \hat{U}^t computationally explicit as a the time-ordered product (1.25).

The matrices that have to be evaluated are very high-dimensional, in principle infinite dimensional, and the numerical challenges can quickly get out of hand. What made it possible to solve these equations analytically in 1920's for a few iconic problems, such as the hydrogen atom, are the symmetries, or in other words group theory, a subject of another course, our [group theory course](#).

Whenever you are confused about an "operator", think "matrix". Here we recapitulate a few matrix algebra concepts that we found essential. The punch line is (1.40): Hamilton-Cayley equation $\prod (\mathbf{M} - \lambda_i \mathbf{1}) = 0$ associates with each distinct root λ_i of a matrix \mathbf{M} a projection onto i th vector subspace

$$\mathbf{P}_i = \prod_{j \neq i} \frac{\mathbf{M} - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j}.$$

What follows - for this week - is a jumble of Predrag's notes. If you understand the examples, we are on the roll. If not, ask :)

How are we to think of the quantum operator (1.1)

$$\hat{H} = \hat{T} + \hat{V}, \quad \hat{T} = \hat{p}^2/2m, \quad \hat{V} = V(\hat{q}), \quad (1.2)$$

corresponding to a classical Hamiltonian $H = T + V$, where T is kinetic energy, and V is the potential?

Expressed in terms of basis functions, the quantum evolution operator is an infinite-dimensional matrix; if we happen to know the eigenbasis of the Hamiltonian, the problem is solved already. In real life we have to guess that some complete basis set is good starting point for solving the problem, and go from there. In practice we truncate such operator representations to finite-dimensional matrices, so it pays to recapitulate a few relevant facts about matrix algebra and some of the properties of functions of finite-dimensional matrices.

Matrix derivatives. The derivative of a matrix is a matrix with elements

$$\mathbf{A}'(x) = \frac{d\mathbf{A}(x)}{dx}, \quad A'_{ij}(x) = \frac{d}{dx} A_{ij}(x). \quad (1.3)$$

Derivatives of products of matrices are evaluated by the chain rule

$$\frac{d}{dx}(\mathbf{A}\mathbf{B}) = \frac{d\mathbf{A}}{dx}\mathbf{B} + \mathbf{A}\frac{d\mathbf{B}}{dx}. \quad (1.4)$$

A matrix and its derivative matrix in general do not commute

$$\frac{d}{dx} \mathbf{A}^2 = \frac{d\mathbf{A}}{dx} \mathbf{A} + \mathbf{A} \frac{d\mathbf{A}}{dx}. \quad (1.5)$$

The derivative of the inverse of a matrix, follows from $\frac{d}{dx}(\mathbf{A}\mathbf{A}^{-1}) = 0$:

$$\frac{d}{dx} \mathbf{A}^{-1} = -\frac{1}{\mathbf{A}} \frac{d\mathbf{A}}{dx} \frac{1}{\mathbf{A}}. \quad (1.6)$$

Matrix functions. A function of a single variable that can be expressed in terms of additions and multiplications generalizes to a matrix-valued function by replacing the variable by the matrix.

In particular, the exponential of a constant matrix can be defined either by its series expansion, or as a limit of an infinite product:

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k, \quad \mathbf{A}^0 = \mathbf{1} \quad (1.7)$$

$$= \lim_{N \rightarrow \infty} \left(\mathbf{1} + \frac{1}{N} \mathbf{A} \right)^N \quad (1.8)$$

The first equation follows from the second one by the binomial theorem, so these indeed are equivalent definitions. That the terms of order $O(N^{-2})$ or smaller do not matter follows from the bound

$$\left(1 + \frac{x - \epsilon}{N} \right)^N < \left(1 + \frac{x + \delta x_N}{N} \right)^N < \left(1 + \frac{x + \epsilon}{N} \right)^N,$$

where $|\delta x_N| < \epsilon$. If $\lim \delta x_N \rightarrow 0$ as $N \rightarrow \infty$, the extra terms do not contribute. A proof for matrices would probably require defining the norm of a matrix (and, more generally, a norm of an operator acting on a Banach space) first. If you know an easy proof, let us know.

Logarithm of a matrix. The logarithm of a matrix is defined by the power series

$$\ln(\mathbf{1} - \mathbf{A}) = -\sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k}. \quad (1.9)$$

log det = tr log matrix identity. Consider now the determinant

$$\det(e^{\mathbf{A}}) = \lim_{N \rightarrow \infty} (\det(\mathbf{1} + \mathbf{A}/N))^N.$$

To the leading order in $1/N$

$$\det(\mathbf{1} + \mathbf{A}/N) = 1 + \frac{1}{N} \text{tr} \mathbf{A} + O(N^{-2}).$$

hence

$$\det e^{\mathbf{A}} = \lim_{N \rightarrow \infty} \left(1 + \frac{1}{N} \text{tr } \mathbf{A} + O(N^{-2}) \right)^N = \lim_{N \rightarrow \infty} \left(1 + \frac{\text{tr } \mathbf{A}}{N} \right)^N = e^{\text{tr } \mathbf{A}} \quad (1.10)$$

Defining $\mathbf{M} = e^{\mathbf{A}}$ we can write this as

$$\ln \det \mathbf{M} = \text{tr } \ln \mathbf{M}, \quad (1.11)$$

a crazy useful identity that you will run into over and over again.

Functions of several matrices. Due to non-commutativity of matrices, generalization of a function of several variables to a function is not as straightforward. Expression involving several matrices depend on their commutation relations. For example, the commutator expansion

$$e^{t\mathbf{A}} \mathbf{B} e^{-t\mathbf{A}} = \mathbf{B} + t[\mathbf{A}, \mathbf{B}] + \frac{t^2}{2}[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \frac{t^3}{3!}[\mathbf{A}, [\mathbf{A}, [\mathbf{A}, \mathbf{B}]]] + \dots \quad (1.12)$$

sometimes used to establish the equivalence of the Heisenberg and Schrödinger pictures of quantum mechanics follows by recursive evaluation of t derivatives

$$\frac{d}{dt} (e^{t\mathbf{A}} \mathbf{B} e^{-t\mathbf{A}}) = e^{t\mathbf{A}} [\mathbf{A}, \mathbf{B}] e^{-t\mathbf{A}}.$$

Expanding $\exp(\mathbf{A} + \mathbf{B})$, $\exp \mathbf{A}$, $\exp \mathbf{B}$ to first few orders using (1.7) yields

$$e^{(\mathbf{A}+\mathbf{B})/N} = e^{\mathbf{A}/N} e^{\mathbf{B}/N} - \frac{1}{2N^2} [\mathbf{A}, \mathbf{B}] + O(N^{-3}), \quad (1.13)$$

and the Trotter product formula: if \mathbf{B} , \mathbf{C} and $\mathbf{A} = \mathbf{B} + \mathbf{C}$ are matrices, then

$$e^{\mathbf{A}} = \lim_{N \rightarrow \infty} \left(e^{\mathbf{B}/N} e^{\mathbf{C}/N} \right)^N \quad (1.14)$$

In particular, we can now make sense of the quantum evolution operator (1.1) as a succession of short free flights (kinetic term) interspersed by small acceleration kicks (potential term),

$$e^{-it\hat{H}} = \lim_{N \rightarrow \infty} \left(e^{-i\Delta t \hat{T}} e^{-i\Delta t \hat{V}} \right)^N, \quad \Delta t = t/N, \quad (1.15)$$

where we have set $\hbar = 1$.

1.3 A linear diversion

(Notes based of ChaosBook.org/chapters/stability.pdf)

Linear fields are the simplest vector fields, described by linear differential equations which can be solved explicitly, with solutions that are good for all times. The state space for linear differential equations is $\mathcal{M} = \mathbb{R}^d$, and the equations of motion are written in terms of a vector x and a constant stability matrix A as

$$\dot{x} = v(x) = Ax. \quad (1.16)$$

Solving this equation means finding the state space trajectory

$$x(t) = (x_1(t), x_2(t), \dots, x_d(t))$$

passing through a given initial point x_0 . If $x(t)$ is a solution with $x(0) = x_0$ and $y(t)$ another solution with $y(0) = y_0$, then the linear combination $ax(t) + by(t)$ with $a, b \in \mathbb{R}$ is also a solution, but now starting at the point $ax_0 + by_0$. At any instant in time, the space of solutions is a d -dimensional vector space, spanned by a basis of d linearly independent solutions.

How do we solve the linear differential equation (1.16)? If instead of a matrix equation we have a scalar one, $\dot{x} = \lambda x$, the solution is $x(t) = e^{t\lambda}x_0$. In order to solve the d -dimensional matrix case, it is helpful to rederive this solution by studying what happens for a short time step δt . If time $t = 0$ coincides with position $x(0)$, then

$$\frac{x(\delta t) - x(0)}{\delta t} = \lambda x(0), \quad (1.17)$$

which we iterate m times to obtain Euler's formula for compounding interest

$$x(t) \approx \left(1 + \frac{t}{m}\lambda\right)^m x(0) \approx e^{t\lambda}x(0). \quad (1.18)$$

The term in parentheses acts on the initial condition $x(0)$ and evolves it to $x(t)$ by taking m small time steps $\delta t = t/m$. As $m \rightarrow \infty$, the term in parentheses converges to $e^{t\lambda}$. Consider now the matrix version of equation (1.17):

$$\frac{x(\delta t) - x(0)}{\delta t} = Ax(0). \quad (1.19)$$

A representative point x is now a vector in \mathbb{R}^d acted on by the matrix A , as in (1.16). Denoting by $\mathbf{1}$ the identity matrix, and repeating the steps (1.17) and (1.18) we obtain Euler's formula for the exponential of a matrix:

$$x(t) = J^t x(0), \quad J^t = e^{tA} = \lim_{m \rightarrow \infty} \left(\mathbf{1} + \frac{t}{m}A\right)^m. \quad (1.20)$$

We will find this definition for the exponential of a matrix helpful in the general case, where the matrix $A = A(x(t))$ varies along a trajectory.

Now that we have some feeling for the qualitative behavior of a linear flow, we are ready to return to the nonlinear case. Consider an infinitesimal perturbation of the

initial state, $x_0 + \delta x(x_0, 0)$. How do we compute the exponential (1.20) that describes linearized perturbation $\delta x(x_0, t)$?

$$x(t) = f^t(x_0), \quad \delta x(x_0, t) = J^t(x_0) \delta x(x_0, 0). \quad (1.21)$$

The equations are linear, so we should be able to integrate them—but in order to make sense of the answer, we derive this integration step by step. The Jacobian matrix is computed by integrating the equations of variations

$$\dot{x}_i = v_i(x), \quad \dot{\delta x}_i = \sum_j A_{ij}(x) \delta x_j \quad (1.22)$$

Consider the case of a general, non-stationary trajectory $x(t)$. The exponential of a constant matrix can be defined either by its Taylor series expansion or in terms of the Euler limit (1.20):

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = \lim_{m \rightarrow \infty} \left(\mathbf{1} + \frac{t}{m} A \right)^m. \quad (1.23)$$

Taylor expanding is fine if A is a constant matrix. However, only the second, tax-accountant's discrete step definition of an exponential is appropriate for the task at hand. For dynamical systems, the local rate of neighborhood distortion $A(x)$ depends on where we are along the trajectory. The linearized neighborhood is deformed along the flow, and the m discrete time-step approximation to J^t is therefore given by a generalization of the Euler product (1.23):

$$\begin{aligned} J^t &= \lim_{m \rightarrow \infty} \prod_{n=m}^1 (\mathbf{1} + \delta t A(x_n)) = \lim_{m \rightarrow \infty} \prod_{n=m}^1 e^{\delta t A(x_n)} \\ &= \lim_{m \rightarrow \infty} e^{\delta t A(x_m)} e^{\delta t A(x_{m-1})} \dots e^{\delta t A(x_2)} e^{\delta t A(x_1)}, \end{aligned} \quad (1.24)$$

where $\delta t = (t - t_0)/m$, and $x_n = x(t_0 + n\delta t)$. Indexing of the products indicates that the successive infinitesimal deformation are applied by multiplying from the left. The $m \rightarrow \infty$ limit of this procedure is the formal integral

$$J_{ij}^t(x_0) = \left[\mathbf{T} e^{\int_0^t d\tau A(x(\tau))} \right]_{ij}, \quad (1.25)$$

where \mathbf{T} stands for time-ordered integration, *defined* as the continuum limit of the successive multiplications (1.24). This integral formula for J^t is the finite time companion of the differential definition

$$\dot{J}(t) = A(t)J(t), \quad (1.26)$$

with the initial condition $J(0) = \mathbf{1}$. The definition makes evident important properties of Jacobian matrices, such as their being multiplicative along the flow,

$$J^{t+t'}(x) = J^{t'}(x') J^t(x), \quad \text{where } x' = f^{t'}(x_0), \quad (1.27)$$

which is an immediate consequence of the time-ordered product structure of (1.24). However, in practice J is evaluated by integrating differential equation (1.26) along with the ODEs (3.6) that define a particular flow.

exercise 1.3

1.4 Eigenvalues and eigenvectors

10. Try to leave out the part that readers tend to skip.
— Elmore Leonard's Ten Rules of Writing.

Eigenvalues of a $[d \times d]$ matrix \mathbf{M} are the roots of its characteristic polynomial

$$\det(\mathbf{M} - \lambda \mathbf{1}) = \prod (\lambda_i - \lambda) = 0. \quad (1.28)$$

Given a nonsingular matrix \mathbf{M} , $\det \mathbf{M} \neq 0$, with all $\lambda_i \neq 0$, acting on d -dimensional vectors \mathbf{x} , we would like to determine *eigenvectors* $\mathbf{e}^{(i)}$ of \mathbf{M} on which \mathbf{M} acts by scalar multiplication by eigenvalue λ_i

$$\mathbf{M} \mathbf{e}^{(i)} = \lambda_i \mathbf{e}^{(i)}. \quad (1.29)$$

If $\lambda_i \neq \lambda_j$, $\mathbf{e}^{(i)}$ and $\mathbf{e}^{(j)}$ are linearly independent. There are at most d distinct eigenvalues, which we assume have been computed by some method, and ordered by their real parts, $\operatorname{Re} \lambda_i \geq \operatorname{Re} \lambda_{i+1}$.

If all eigenvalues are distinct $\mathbf{e}^{(j)}$ are d linearly independent vectors which can be used as a (non-orthogonal) basis for any d -dimensional vector $\mathbf{x} \in \mathbb{R}^d$

$$\mathbf{x} = x_1 \mathbf{e}^{(1)} + x_2 \mathbf{e}^{(2)} + \dots + x_d \mathbf{e}^{(d)}. \quad (1.30)$$

From (1.29) it follows that

$$(\mathbf{M} - \lambda_i \mathbf{1}) \mathbf{e}^{(j)} = (\lambda_j - \lambda_i) \mathbf{e}^{(j)},$$

matrix $(\mathbf{M} - \lambda_j \mathbf{1})$ annihilates $\mathbf{e}^{(j)}$, the product of all such factors annihilates any vector, and the matrix \mathbf{M} satisfies its characteristic equation

$$\prod_{i=1}^d (\mathbf{M} - \lambda_i \mathbf{1}) = 0. \quad (1.31)$$

This humble fact has a name: the Hamilton-Cayley theorem. If we delete one term from this product, we find that the remainder projects \mathbf{x} from (1.30) onto the corresponding eigenspace:

$$\prod_{j \neq i} (\mathbf{M} - \lambda_j \mathbf{1}) \mathbf{x} = \prod_{j \neq i} (\lambda_i - \lambda_j) x_i \mathbf{e}^{(i)}.$$

Dividing through by the $(\lambda_i - \lambda_j)$ factors yields the *projection operators*

$$\mathbf{P}_i = \prod_{j \neq i} \frac{\mathbf{M} - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j}, \quad (1.32)$$

which are *orthogonal* and *complete*:

$$\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_j, \quad (\text{no sum on } j), \quad \sum_{i=1}^r \mathbf{P}_i = \mathbf{1}, \quad (1.33)$$

with the dimension of the i th subspace given by $d_i = \text{tr } \mathbf{P}_i$. For each distinct eigenvalue λ_i of \mathbf{M} ,

$$(\mathbf{M} - \lambda_j \mathbf{1}) \mathbf{P}_j = \mathbf{P}_j (\mathbf{M} - \lambda_j \mathbf{1}) = 0, \quad (1.34)$$

the columns/rows of \mathbf{P}_i are the right/left eigenvectors $\mathbf{e}^{(k)}$, $\mathbf{e}_{(k)}$ of \mathbf{M} which (provided \mathbf{M} is not of Jordan type, see example 1.1) span the corresponding linearized subspace.

The main take-home is that once the distinct non-zero eigenvalues $\{\lambda_i\}$ are computed, projection operators are polynomials in \mathbf{M} which need no further diagonalizations or orthogonalizations. It follows from the characteristic equation (1.34) that λ_i is the eigenvalue of \mathbf{M} on \mathbf{P}_i subspace:

$$\mathbf{M} \mathbf{P}_i = \lambda_i \mathbf{P}_i \quad (\text{no sum on } i). \quad (1.35)$$

Using $\mathbf{M} = \mathbf{M} \mathbf{1}$ and completeness relation (1.33) we can rewrite \mathbf{M} as

$$\mathbf{M} = \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \cdots + \lambda_d \mathbf{P}_d. \quad (1.36)$$

Any matrix function $f(\mathbf{M})$ takes the scalar value $f(\lambda_i)$ on the \mathbf{P}_i subspace, $f(\mathbf{M}) \mathbf{P}_i = f(\lambda_i) \mathbf{P}_i$, and is thus easily evaluated through its *spectral decomposition*

$$f(\mathbf{M}) = \sum_i f(\lambda_i) \mathbf{P}_i. \quad (1.37)$$

This, of course, is the reason why anyone but a fool works with irreducible reps: they reduce matrix (AKA “operator”) evaluations to manipulations with numbers.

By (1.34) every column of \mathbf{P}_i is proportional to a right eigenvector $\mathbf{e}^{(i)}$, and its every row to a left eigenvector $\mathbf{e}_{(i)}$. In general, neither set is orthogonal, but by the idempotence condition (1.33), they are mutually orthogonal,

$$\mathbf{e}_{(i)} \cdot \mathbf{e}^{(j)} = c \delta_i^j. \quad (1.38)$$

The non-zero constant c is convention dependent and not worth fixing, unless you feel nostalgic about Clebsch-Gordan coefficients. We shall set $c = 1$. Then it is convenient to collect all left and right eigenvectors into a single matrix.

Example 1.1. Degenerate eigenvalues. *While for a matrix with generic real elements all eigenvalues are distinct with probability 1, that is not true in presence of symmetries, or spacial parameter values (bifurcation points). What can one say about situation where d_α eigenvalues are degenerate, $\lambda_\alpha = \lambda_i = \lambda_{i+1} = \cdots = \lambda_{i+d_\alpha-1}$? Hamilton-Cayley (1.31) now takes form*

$$\prod_{\alpha=1}^r (\mathbf{M} - \lambda_\alpha \mathbf{1})^{d_\alpha} = 0, \quad \sum_{\alpha} d_\alpha = d. \quad (1.39)$$

We distinguish two cases:

M can be brought to diagonal form. The characteristic equation (1.39) can be replaced by the minimal polynomial,

$$\prod_{\alpha=1}^r (\mathbf{M} - \lambda_{\alpha} \mathbf{1}) = 0, \quad (1.40)$$

where the product includes each distinct eigenvalue only once. Matrix \mathbf{M} acts multiplicatively

$$\mathbf{M} \mathbf{e}^{(\alpha,k)} = \lambda_i \mathbf{e}^{(\alpha,k)}, \quad (1.41)$$

on a d_{α} -dimensional subspace spanned by a linearly independent set of basis eigenvectors $\{\mathbf{e}^{(\alpha,1)}, \mathbf{e}^{(\alpha,2)}, \dots, \mathbf{e}^{(\alpha,d_{\alpha})}\}$. This is the easy case. Luckily, if the degeneracy is due to a finite or compact symmetry group, relevant \mathbf{M} matrices can always be brought to such Hermitian, diagonalizable form.

M can only be brought to upper-triangular, Jordan form. This is the messy case, so we only illustrate the key idea in example 1.2.

Example 1.2. Decomposition of 2-dimensional vector spaces: Enumeration of every possible kind of linear algebra eigenvalue / eigenvector combination is beyond what we can reasonably undertake here. However, enumerating solutions for the simplest case, a general $[2 \times 2]$ non-singular matrix

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}.$$

takes us a long way toward developing intuition about arbitrary finite-dimensional matrices. The eigenvalues

$$\lambda_{1,2} = \frac{1}{2} \text{tr } \mathbf{M} \pm \frac{1}{2} \sqrt{(\text{tr } \mathbf{M})^2 - 4 \det \mathbf{M}} \quad (1.42)$$

are the roots of the characteristic (secular) equation (1.28):

$$\begin{aligned} \det(\mathbf{M} - \lambda \mathbf{1}) &= (\lambda_1 - \lambda)(\lambda_2 - \lambda) \\ &= \lambda^2 - \text{tr } \mathbf{M} \lambda + \det \mathbf{M} = 0. \end{aligned}$$

Distinct eigenvalues case has already been described in full generality. The left/right eigenvectors are the rows/columns of projection operators

$$P_1 = \frac{\mathbf{M} - \lambda_2 \mathbf{1}}{\lambda_1 - \lambda_2}, \quad P_2 = \frac{\mathbf{M} - \lambda_1 \mathbf{1}}{\lambda_2 - \lambda_1}, \quad \lambda_1 \neq \lambda_2. \quad (1.43)$$

Degenerate eigenvalues. If $\lambda_1 = \lambda_2 = \lambda$, we distinguish two cases: (a) \mathbf{M} can be brought to diagonal form. This is the easy case. (b) \mathbf{M} can be brought to Jordan form, with zeros everywhere except for the diagonal, and some 1's directly above it; for a $[2 \times 2]$ matrix the Jordan form is

$$\mathbf{M} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad \mathbf{e}^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}^{(2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

$\mathbf{v}^{(2)}$ helps span the 2-dimensional space, $(\mathbf{M} - \lambda)^2 \mathbf{v}^{(2)} = 0$, but is not an eigenvector, as $\mathbf{M}\mathbf{v}^{(2)} = \lambda \mathbf{v}^{(2)} + \mathbf{e}^{(1)}$. For every such Jordan $[d_\alpha \times d_\alpha]$ block there is only one eigenvector per block. Noting that

$$\mathbf{M}^m = \begin{bmatrix} \lambda^m & m\lambda^{m-1} \\ 0 & \lambda^m \end{bmatrix},$$

we see that instead of acting multiplicatively on \mathbb{R}^2 , Jacobian matrix $J^t = \exp(t\mathbf{M})$

$$e^{t\mathbf{M}} \begin{pmatrix} u \\ v \end{pmatrix} = e^{t\lambda} \begin{pmatrix} u + tv \\ v \end{pmatrix} \quad (1.44)$$

picks up a power-law correction. That spells trouble (logarithmic term $\ln t$ if we bring the extra term into the exponent).

Example 1.3. Projection operator decomposition in 2 dimensions: Let's illustrate how the distinct eigenvalues case works with the $[2 \times 2]$ matrix

$$\mathbf{M} = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}.$$

Its eigenvalues $\{\lambda_1, \lambda_2\} = \{5, 1\}$ are the roots of (1.42):

$$\det(\mathbf{M} - \lambda \mathbf{1}) = \lambda^2 - 6\lambda + 5 = (\lambda - 5)(\lambda - 1) = 0.$$

That \mathbf{M} satisfies its secular equation (Hamilton-Cayley theorem) can be verified by explicit calculation:

$$\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}^2 - 6 \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} + 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Associated with each root λ_i is the projection operator (1.43)

$$P_1 = \frac{1}{4}(\mathbf{M} - \mathbf{1}) = \frac{1}{4} \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix} \quad (1.45)$$

$$P_2 = \frac{1}{4}(\mathbf{M} - 5 \cdot \mathbf{1}) = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ -3 & 3 \end{bmatrix}. \quad (1.46)$$

Matrices \mathbf{P}_i are orthonormal and complete, The dimension of the i th subspace is given by $d_i = \text{tr } \mathbf{P}_i$; in case at hand both subspaces are 1-dimensional. From the characteristic equation it follows that \mathbf{P}_i satisfies the eigenvalue equation $\mathbf{M}\mathbf{P}_i = \lambda_i \mathbf{P}_i$. Two consequences are immediate. First, we can easily evaluate any function of \mathbf{M} by spectral decomposition, for example

$$\mathbf{M}^7 - 3 \cdot \mathbf{1} = (5^7 - 3)\mathbf{P}_1 + (1 - 3)\mathbf{P}_2 = \begin{bmatrix} 58591 & 19531 \\ 58593 & 19529 \end{bmatrix}.$$

Second, as \mathbf{P}_i satisfies the eigenvalue equation, its every column is a right eigenvector, and every row a left eigenvector. Picking first row/column we get the eigenvectors:

$$\begin{aligned} \{\mathbf{e}^{(1)}, \mathbf{e}^{(2)}\} &= \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \end{bmatrix} \right\} \\ \{\mathbf{e}_{(1)}, \mathbf{e}_{(2)}\} &= \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}, \end{aligned}$$

with overall scale arbitrary. The matrix is not symmetric, so $\{e^{(j)}\}$ do not form an orthogonal basis. The left-right eigenvector dot products $e^{(j)} \cdot e^{(k)}$, however, are orthogonal as in (1.38), by inspection. (Continued in example 2.2.)

Example 1.4. Computing matrix exponentials. If A is diagonal (the system is uncoupled), then e^{tA} is given by

$$\exp \begin{pmatrix} \lambda_1 t & & & \\ & \lambda_2 t & & \\ & & \ddots & \\ & & & \lambda_d t \end{pmatrix} = \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_d t} \end{pmatrix}.$$

If A is diagonalizable, $A = FDF^{-1}$, where D is the diagonal matrix of the eigenvalues of A and F is the matrix of corresponding eigenvectors, the result is simple: $A^n = (FDF^{-1})(FDF^{-1}) \dots (FDF^{-1}) = F D^n F^{-1}$. Inserting this into the Taylor series for e^x gives $e^{At} = F e^{Dt} F^{-1}$.

But A may not have d linearly independent eigenvectors, making F singular and forcing us to take a different route. To illustrate this, consider $[2 \times 2]$ matrices. For any linear system in \mathbb{R}^2 , there is a similarity transformation

$$B = U^{-1}AU,$$

where the columns of U consist of the generalized eigenvectors of A such that B has one of the following forms:

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad B = \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix}.$$

These three cases, called normal forms, correspond to A having (1) distinct real eigenvalues, (2) degenerate real eigenvalues, or (3) a complex pair of eigenvalues. It follows that

$$e^{Bt} = \begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{bmatrix}, \quad e^{Bt} = e^{\lambda t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}, \quad e^{Bt} = e^{\mu t} \begin{bmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{bmatrix},$$

and $e^{At} = U e^{Bt} U^{-1}$. What we have done is classify all $[2 \times 2]$ matrices as belonging to one of three classes of geometrical transformations. The first case is scaling, the second is a shear, and the third is a combination of rotation and scaling. The generalization of these normal forms to \mathbb{R}^d is called the Jordan normal form. (J. Halcrow)

1.4.1 Yes, but how do you really do it?

As M has only real entries, it will in general have either real eigenvalues (over-damped oscillator, for example), or complex conjugate pairs of eigenvalues (under-damped oscillator, for example). That is not surprising, but also the corresponding eigenvectors can be either real or complex. All coordinates used in defining the flow are real numbers, so what is the meaning of a complex eigenvector?

If two eigenvalues form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$, they are in a sense degenerate: while a real λ_k characterizes a motion along a line, a complex λ_k characterizes a spiralling motion in a plane. We determine this plane by

replacing the corresponding complex eigenvectors by their real and imaginary parts, $\{\mathbf{e}^{(k)}, \mathbf{e}^{(k+1)}\} \rightarrow \{\operatorname{Re} \mathbf{e}^{(k)}, \operatorname{Im} \mathbf{e}^{(k)}\}$, or, in terms of projection operators:

$$\mathbf{P}_k = \frac{1}{2}(\mathbf{R} + i\mathbf{Q}), \quad \mathbf{P}_{k+1} = \mathbf{P}_k^*,$$

where $\mathbf{R} = \mathbf{P}_k + \mathbf{P}_{k+1}$ is the subspace decomposed by the k th complex eigenvalue pair, and $\mathbf{Q} = (\mathbf{P}_k - \mathbf{P}_{k+1})/i$, both matrices with real elements. Substitution

$$\begin{bmatrix} \mathbf{P}_k \\ \mathbf{P}_{k+1} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{Q} \end{bmatrix},$$

suggest introduction of a $\det U = 1$, special unitary matrix

$$U = \frac{e^{i\pi/2}}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \quad (1.47)$$

which brings the $\lambda_k \mathbf{P}_k + \lambda_{k+1} \mathbf{P}_{k+1}$ complex eigenvalue pair in the spectral decomposition into the real form,

$$U^T \begin{bmatrix} \mu + i\omega & 0 \\ 0 & \mu - i\omega \end{bmatrix} U = \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix}.$$

$$[\mathbf{P}_k, \mathbf{P}_{k+1}] \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^* \end{bmatrix} \begin{bmatrix} \mathbf{P}_k \\ \mathbf{P}_{k+1} \end{bmatrix} = [\mathbf{R}, \mathbf{Q}] \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{Q} \end{bmatrix}, \quad (1.48)$$

where we have dropped the superscript $^{(k)}$ for notational brevity.

To summarize, spectrally decomposed matrix \mathbf{M} acts along lines on subspaces corresponding to real eigenvalues, and as a $[2 \times 2]$ rotation in a plane on subspaces corresponding to complex eigenvalue pairs.

Commentary

Remark 1.1. Projection operators. The construction of projection operators given in sect. 1.4.1 is taken from refs. [2, 3]. Who wrote this down first we do not know, lineage certainly goes all the way back to Lagrange polynomials [8], but projection operators tend to get drowned in sea of algebraic details. Arfken and Weber [1] ascribe spectral decomposition (1.37) to Sylvester. Halmos [4] is a good early reference - but we like Harter's exposition [5–7] best, for its multitude of specific examples and physical illustrations. In particular, by the time we get to (1.34) we have tacitly assumed full diagonalizability of matrix \mathbf{M} . That is the case for the compact groups we will study here (they are all subgroups of $U(n)$) but not necessarily in other applications. A bit of what happens then (nilpotent blocks) is touched upon in example 1.2. Harter in his lecture Harter's [lecture 5](#) (starts about min. 31 into the lecture) explains this in great detail - its well worth your time.

References

- [1] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists: A Comprehensive Guide*, 6th ed. (Academic, New York, 2005).

- [2] P. Cvitanović, “Group theory for Feynman diagrams in non-Abelian gauge theories”, *Phys. Rev. D* **14**, 1536–1553 (1976).
- [3] P. Cvitanović, *Classical and exceptional Lie algebras as invariance algebras*, Oxford Univ. preprint 40/77, unpublished., 1977.
- [4] P. R. Halmos, *Finite-Dimensional Vector Spaces* (Princeton Univ. Press, Princeton NJ, 1948).
- [5] W. G. Harter, “Algebraic theory of ray representations of finite groups”, *J. Math. Phys.* **10**, 739–752 (1969).
- [6] W. G. Harter, *Principles of Symmetry, Dynamics, and Spectroscopy* (Wiley, New York, 1993).
- [7] W. G. Harter and N. dos Santos, “Double-group theory on the half-shell and the two-level system. I. Rotation and half-integral spin states”, *Amer. J. Phys.* **46**, 251–263 (1978).
- [8] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed. (Prentice-Hall, Englewood Cliffs NJ, 1971).
- [9] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge, 2009).

Exercises

- 1.1. **Trace-log of a matrix.** Prove that

$$\det M = e^{\text{tr} \ln M}.$$

for an arbitrary nonsingular finite dimensional matrix M , $\det M \neq 0$.

- 1.2. **Stability, diagonal case.** Verify that for a diagonalizable matrix A the exponential is also diagonalizable

$$J^t = e^{tA} = \mathbf{U}^{-1} e^{tA_D} \mathbf{U}, \quad A_D = \mathbf{U} \mathbf{A} \mathbf{U}^{-1}. \quad (1.49)$$

- 1.3. **Time-ordered exponentials.** Given a time dependent matrix $A(t)$, show that the time-ordered exponential

$$J(t) = \mathbf{T} e^{\int_0^t d\tau A(\tau)}$$

may be written as

$$J(t) = \mathbf{1} + \sum_{m=1}^{\infty} \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{m-1}} dt_m A(t_1) A(t_2) \cdots A(t_m). \quad (1.50)$$

(Hint: for a warmup, consider summing elements of a finite-dimensional symmetric matrix $S = S^\top$. Use the symmetry to sum over each matrix element once; (1.50) is a continuous limit generalization, for an object symmetric in m variables. If you find this hint confusing, ignore it.) Verify, by using this representation, that $J(t)$ satisfies the equation

$$\dot{J}(t) = A(t)J(t),$$

with the initial condition $J(0) = 1$.

- 1.4. **Real representation of complex eigenvalues.** (Verification of example 3.2.) λ_k, λ_{k+1} eigenvalues form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$. Show that

- (a) corresponding projection operators are complex conjugates of each other,

$$\mathbf{P} = \mathbf{P}_k, \quad \mathbf{P}^* = \mathbf{P}_{k+1},$$

where we denote \mathbf{P}_k by \mathbf{P} for notational brevity.

- (b) \mathbf{P} can be written as

$$\mathbf{P} = \frac{1}{2}(\mathbf{R} + i\mathbf{Q}),$$

where $\mathbf{R} = \mathbf{P}_k + \mathbf{P}_{k+1}$ and \mathbf{Q} are matrices with real elements.

- (c)

$$\begin{bmatrix} \mathbf{P}_k \\ \mathbf{P}_{k+1} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{Q} \end{bmatrix}.$$

- (d) The $\dots + \lambda_k \mathbf{P}_k + \lambda_k^* \mathbf{P}_{k+1} + \dots$ complex eigenvalue pair in the spectral decomposition (1.36) is now replaced by a real $[2 \times 2]$ matrix

$$\dots + \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{Q} \end{bmatrix} + \dots$$

or whatever you find the clearest way to write this real representation.