

group theory - week 1

Linear algebra

Georgia Tech PHYS-7143

Homework HW1

due Tuesday 2021-05-25

== 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>The matrix square root</i> | 4 points |


Total of 10 points = 100 % score.

Diagonalizing the matrix: that's the key to the whole thing.

— Governor Arnold Schwarzenegger

2021-05-18 Predrag Week 1 *Linear algebra*


- Light purple text in this PDF is a live hyperlink. If you encounter a ChaosBook.org web login: all copyright-protected references are on a password protected site. What password? If you are a Georgia Tech student, I can help you with that.
- The fastest way to watch any week's lecture videos is by letting YouTube run the


 *course playlist*


Downside is that the playlist plays also all 'extra' videos - you can skip through those, if you are short on time. Or patience.


1.1 Week's videos, reading


- Sect. 1.4 *Matrix-valued functions*


 *Matrices : 2 kinds*

 *Derivative of a matrix function*

 *Exponential, logarithm of a matrix*

 *Determinant is a volume*

 *$\log \det = \text{tr} \log$ (updated Aug 18, 2020)*

 *Multi-matrix functions (optional, for the QM inclined)*


- Sect. 1.5 *A linear diversion*


There are two representations of exponential of constant matrix, the Taylor series (1.7) and the compound interest (Euler product) formulas (1.8).


 *Linear differential equations*


- Sect. 1.6 *Eigenvalues and eigenvectors*

Hamilton-Cayley equation, projection operators (1.27), any matrix function is evaluated by spectral decomposition (1.30). Work through example 1.5.


 *Eigenvalues and eigenvectors*

 *What's the deal with Hamilton-Cayley?*

 *Spectral decomposition*

 *Spectral decomposition and completeness*

 *Right, left eigenvectors*


 *A projection operators workout*


1.2 Other sources (optional)


Mopping up operations are the activities that engage most scientists throughout their careers.


— Thomas Kuhn, *The Structure of Scientific Revolutions*

The subject of linear algebra is a vast and very alive research area, generates innumerable tomes of its own, and is way beyond what we can exhaustively cover here. Some linear operators and matrices reading (optional reading for week 1, not required for this course; whenever the text is colored, you can click on the live hyperlink in the pdf version of these notes):


 Stone and Goldbart [15], *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.

 AWH p. 113 [Functions of Matrices](#). Anything prefixed by AWH, like “Kronecker product [AWH eq. \(2.55\)](#)” refers to the more pedestrian and perhaps easier to read Arfken, Weber & Harris [2] *Mathematical Methods for Physicists: A Comprehensive Guide* ([Georgia Tech students can get it from GaTech Library](#)).


 AWH Section 2.2 *Matrices*

 AWH Example 2.2.6 *Exponential of a diagonal matrix*


 AWH Chapter 2 *Determinants and matrices* ([click here](#)).


 AWH Section 6.1 *Eigenvalue Equations* ([click here](#)).


 AWH Chapter 6 *Eigenvalue problems* ([click here](#)).

 In sect. 1.4 I make matrix functions appear easier than they really are. For an in-depth discussion, consult Golub and Van Loan [8] *Matrix Computations*, chap. 9 *Functions of Matrices* ([click here](#)).

exercise 1.3

 Much more than you ever wanted to know about linear algebra: Axler [3] *Down with determinants!* ([click here](#)).

 Steve Trettel *Linear Algebra and the Periodic Table* is a gentle 53 min tour from vectors to function spaces to quantum mechanics. True, what they teach you as QM is 95% linear algebra, but Trettel does not mention that QM is 95% one amazing experimental fact: \hbar is a nature-given *constant*. Mathematicians...

 Grant Sanderson *Essence of linear algebra* ([3Blue1Brown](#)). Karan Shah likes the geometrical explanations of linear algebra eigen-values / -vectors, recommends it.

Question 1.1. Henriette Roux finds course notes confusing

Q Couldn't you use one single, definitive text for methods taught in the course?

A It's a grad school, so it is research focused - I myself am (re)learning the topics that we are

going through the course, using various sources. My emphasis in this course is on understanding and meaning, not on getting all signs and 2π 's right, and I find reading about the topic from several perspectives helpful. But if you really find one book more comfortable, nearly all topics are covered in Arfken, Weber & Harris [2].

1.3 Special projects

Several people have been interested in taking a **special project**, instead of the final in the course. If you propose to work out in detail some group-theory needed for your own research (but you have not taken the time to master the theory), that would be ideal. Random examples of interesting topics (i.e., something that Predrag would like to learn from you:) –

1. The talk by [David Weitz](#) on melting of crystal lattices. Can you do a calculation on a Wigner lattice or a graphene, or a silicon carbide polytype used as a substrate in our graphene lab (ask Claire Berger about it), using our group theory methods as applied to space groups (2- or 3-D lattices)?
2. If you are really wild about string theory, then you can read Giles and Thorn [7] *Lattice approach to string theory*, and write up what you have learned as the project report. The Giles-Thorn (GT) discretization of the worldsheet begins with a representation of the free closed or open string propagator as a light-cone worldsheet path integral defined on a lattice. The sequel Papanastasiou and Thorn [14] *Worldsheet propagator on the lightcone worldsheet lattice* gives in Appendix B 2D lattice Neumann open string, Dirichlet open string, and closed string propagators. *Discrete Green's functions* are explained, for example, by Chung and Yau [4] who give explicitly, in their Theorem 6, a 2-dimensional lattice Green's function for a rectangular region $R^{[\ell_1 \times \ell_2]}$. The paper is cited over 100 times, maybe there is a better, more up-to-date one to read in that list.
3. 3-springs system of sect. 6.4.

I recommend that you take a final, as these are hard and time-consuming projects, and the faculty does not want to overburden you with course work. However, if a project dovetails with your research interests, it might be worth it. Fly it by me.

1.4 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.

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, which start sketching out in the second lecture (and fill in the details in the next 27 lectures).

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 (1.35): Hamilton-Cayley equation $\prod_i (\mathbf{M} - \lambda_i \mathbf{1}) = 0$ 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

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

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, if the inverse exists, 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 a 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 for a function of a single variable 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 of several matrices is not as straightforward. Expression involving several matrices depend on their commutation relations. For example, the Baker-Campbell-Hausdorff 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 $\hbar = 1$.

1.5 A linear diversion

Linear is good, nonlinear is bad.
—Jean Bellissard

(Notes based of [ChaosBook.org/chapters/stability.pdf](https://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:

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

where $J^t = J(t)$ is a short hand for $\exp(tA)$.

1.6 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

$$\det(\mathbf{M} - \lambda\mathbf{1}) = \prod (\lambda_i - \lambda) = 0. \quad (1.21)$$

Given a nonsingular matrix \mathbf{M} , 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

$$\mathbf{M}\mathbf{e}^{(i)} = \lambda_i\mathbf{e}^{(i)}. \quad (1.22)$$

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 order by their real parts, $\text{Re } \lambda_i \geq \text{Re } \lambda_{i+1}$.

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 $\mathbf{x} \in \mathbb{R}^d$

$$\mathbf{x} = x_1 \mathbf{e}^{(1)} + x_2 \mathbf{e}^{(2)} + \dots + x_d \mathbf{e}^{(d)}. \quad (1.23)$$

However, r , the number of distinct eigenvalues, is in general smaller than the dimension of the matrix, $r \leq d$ (see example 1.3).

From (1.22) it follows that

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

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

$$\prod_{i=1}^d (\mathbf{M} - \lambda_i \mathbf{1}) = 0. \quad (1.24)$$

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 (1.23) 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*

$$\mathbf{P}_i = \prod_{j \neq i} \frac{\mathbf{M} - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j}, \quad (1.25)$$

which are *orthogonal* and *complete*:

$$\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_j, \quad (\text{no sum on } j), \quad \sum_{i=1}^r \mathbf{P}_i = \mathbf{1}, \quad (1.26)$$

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} ,

$$(\mathbf{M} - \lambda_j \mathbf{1}) \mathbf{P}_j = \mathbf{P}_j (\mathbf{M} - \lambda_j \mathbf{1}) = 0, \quad (1.27)$$

the columns/rows of \mathbf{P}_j are the right/left eigenvectors $\mathbf{e}^{(j)}$, $\mathbf{e}_{(j)}$ of \mathbf{M} which (provided \mathbf{M} is not of Jordan type, see example 1.3) 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 (1.27) that λ_i is the eigenvalue of \mathbf{M} on \mathbf{P}_i subspace:

$$\mathbf{M} \mathbf{P}_i = \lambda_i \mathbf{P}_i \quad (\text{no sum on } i). \quad (1.28)$$

Using $\mathbf{M} = \mathbf{M} \mathbf{1}$ and completeness relation (1.26) we can rewrite \mathbf{M} as

$$\mathbf{M} = \lambda_1 \mathbf{P}_1 + \lambda_2 \mathbf{P}_2 + \cdots + \lambda_d \mathbf{P}_d. \quad (1.29)$$

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*

$$f(\mathbf{M}) = \sum_i f(\lambda_i) \mathbf{P}_i. \quad (1.30)$$

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 (1.27) 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 (1.26), they are mutually orthogonal,

$$\mathbf{e}_{(i)} \cdot \mathbf{e}^{(j)} = c_j \delta_i^j. \quad (1.31)$$

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

Example 1.1. Linear stability of 2-dimensional flows: For a 2-dimensional flow the eigenvalues λ_1, λ_2 of A are either real, leading to a linear motion along their eigenvectors, $x_j(t) = x_j(0) \exp(t\lambda_j)$, or form a complex conjugate pair $\lambda_1 = \mu + i\omega, \lambda_2 = \mu - i\omega$, leading to a circular or spiral motion in the $[x_1, x_2]$ plane, see example 1.2.

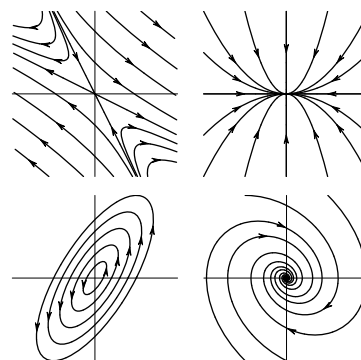


Figure 1.1: Streamlines for several typical 2-dimensional flows: saddle (hyperbolic), in node (attracting), center (elliptic), in spiral.

These two possibilities are refined further into sub-cases depending on the signs of the real part. In the case of real $\lambda_1 > 0, \lambda_2 < 0$, x_1 grows exponentially with time, and x_2 contracts exponentially. This behavior, called a saddle, is sketched in figure 1.1, as are the remaining possibilities: in/out nodes, inward/outward spirals, and the center. The magnitude of out-spiral $|x(t)|$ diverges exponentially when $\mu > 0$, and in-spiral contracts into $(0, 0)$ when $\mu < 0$; whereas, the phase velocity ω controls its oscillations.

If eigenvalues $\lambda_1 = \lambda_2 = \lambda$ are degenerate, the matrix might have two linearly independent eigenvectors, or only one eigenvector, see example 1.3. We distinguish two cases: (a) A can be brought to diagonal form and (b) A can be brought to Jordan

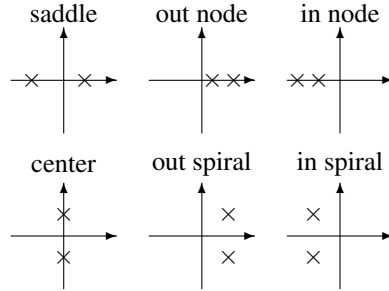


Figure 1.2: Qualitatively distinct types of exponents $\{\lambda^{(1)}, \lambda^{(2)}\}$ of a $[2 \times 2]$ Jacobian matrix. Here the eigenvalues of the Jacobian matrix are *multipliers* $\Lambda^{(j)}$, and the *exponents* are defined as the deformation rates $\lambda^{(j)} = \log(\Lambda^{(j)})/t$.

form, which (in dimension 2 or higher) has zeros everywhere except for the repeating eigenvalues on the diagonal and some 1's directly above it. For every such Jordan $[d_\alpha \times d_\alpha]$ block there is only one eigenvector per block.

We sketch the full set of possibilities in figures 1.1 and 1.2.

Example 1.2. Complex eigenvalues: in-out spirals. As M has only real entries, it will in general have either real eigenvalues, or complex conjugate pairs of eigenvalues. Also the corresponding eigenvectors can be either real or complex. All coordinates used in defining a dynamical flow are real numbers, so what is the meaning of a complex eigenvector?

If λ_k, λ_{k+1} eigenvalues that lie within a diagonal $[2 \times 2]$ sub-block $M' \subset M$ form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$, the corresponding complex eigenvectors can be replaced by their real and imaginary parts, $\{e^{(k)}, e^{(k+1)}\} \rightarrow \{\text{Re } e^{(k)}, \text{Im } e^{(k)}\}$. In this 2-dimensional real representation, $M' \rightarrow A$, the block A is a sum of the rescaling \times identity and the generator of $SO(2)$ rotations in the $\{\text{Re } e^{(1)}, \text{Im } e^{(1)}\}$ plane.

$$A = \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix} = \mu \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Trajectories of $\dot{\mathbf{x}} = A\mathbf{x}$, given by $\mathbf{x}(t) = J^t \mathbf{x}(0)$, where (omitting $e^{(3)}, e^{(4)}, \dots$ eigen-directions)

$$J^t = e^{tA} = e^{t\mu} \begin{bmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{bmatrix}, \tag{1.32}$$

spiral in/out around $(x, y) = (0, 0)$, see figure 1.1, with the rotation period T and the radial expansion /contraction multiplier along the $e^{(j)}$ eigen-direction per a turn of the spiral:

$$T = 2\pi/\omega, \quad \Lambda_{radial} = e^{T\mu}. \tag{1.33}$$

exercise 1.4

We learn that the typical turnover time scale in the neighborhood of the equilibrium $(x, y) = (0, 0)$ is of order $\approx T$ (and not, let us say, $1000T$, or $10^{-2}T$).

Example 1.3. 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, $\lambda_\alpha = \lambda_i = \lambda_{i+1} = \dots = \lambda_{i+d_\alpha-1}$?
Hamilton-Cayley (1.24) now takes form

$$\prod_{\alpha=1}^r (\mathbf{M} - \lambda_\alpha \mathbf{1})^{d_\alpha} = 0, \quad \sum_{\alpha} d_\alpha = d. \quad (1.34)$$

We distinguish two cases:

M can be brought to diagonal form. The characteristic equation (1.34) can be replaced by the minimal polynomial,

$$\prod_{\alpha=1}^r (\mathbf{M} - \lambda_\alpha \mathbf{1}) = 0, \quad (1.35)$$

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

$$\mathbf{M} \mathbf{e}^{(\alpha,k)} = \lambda_i \mathbf{e}^{(\alpha,k)}, \quad (1.36)$$

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 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 1.4.

Example 1.4. 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

$$\lambda_{1,2} = \frac{1}{2} \text{tr } \mathbf{M} \pm \frac{1}{2} \sqrt{(\text{tr } \mathbf{M})^2 - 4 \det \mathbf{M}} \quad (1.37)$$

are the roots of the characteristic (secular) equation (1.21):

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

Distinct eigenvalues case has already been described in full generality. The left/right eigenvectors are the rows/columns of projection operators (see example 1.5)

$$P_1 = \frac{\mathbf{M} - \lambda_2 \mathbf{1}}{\lambda_1 - \lambda_2}, \quad P_2 = \frac{\mathbf{M} - \lambda_1 \mathbf{1}}{\lambda_2 - \lambda_1}, \quad \lambda_1 \neq \lambda_2. \quad (1.38)$$

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 $J^t = \exp(t\mathbf{M})$

$$e^{t\mathbf{M}} \begin{pmatrix} u \\ v \end{pmatrix} = e^{t\lambda} \begin{pmatrix} u + tv \\ v \end{pmatrix} \quad (1.39)$$

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

Example 1.5. Projection operator decomposition in 2 dimensions: Let's illustrate how the distinct eigenvalues case works with the $[2 \times 2]$ matrix [11]

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

Its eigenvalues $\{\lambda_1, \lambda_2\} = \{5, 1\}$ are the roots (1.37):

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

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 (1.38)

$$\mathbf{P}_1 = \frac{1}{4}(\mathbf{M} - \mathbf{1}) = \frac{1}{4} \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix} \quad (1.40)$$

$$\mathbf{P}_2 = -\frac{1}{4}(\mathbf{M} - 5 \cdot \mathbf{1}) = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ -3 & 3 \end{bmatrix}. \quad (1.41)$$

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 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:

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

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 (1.31), by inspection. (Continued in example ??.)

Example 1.6. 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: $A^n = (FDF^{-1})(FDF^{-1}) \dots (FDF^{-1}) = FD^n F^{-1}$. Inserting this into the Taylor series for e^x gives $e^{At} = Fe^{Dt} F^{-1}$.

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 $e^{At} = Ue^{Bt}U^{-1}$. 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)

1.6.1 Yes, but how do you really do it?

As \mathbf{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?

Due to the reality of \mathbf{M} , complex eigenvalues form complex conjugate pairs,

$$\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\},$$

and the sum of terms in the spectral decomposition (1.29) of \mathbf{M} is real,

$$\begin{aligned} \mathbf{M} &= \cdots + (\mu_k + i\omega_k)\mathbf{P}_k + (\mu_k - i\omega_k)\mathbf{P}_{k+1} + \cdots \\ &= \cdots + \mu_k\mathbf{R}_k + \omega_k\mathbf{Q}_k + \cdots, \end{aligned} \quad (1.42)$$

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

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

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

they are in a sense degenerate:

is also a projection operator pair, but this time

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, $\{\mathbf{e}^{(k)}, \mathbf{e}^{(k+1)}\} \rightarrow \{\text{Re } \mathbf{e}^{(k)}, \text{Im } \mathbf{e}^{(k)}\}$, or, in terms of projection operators: 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

$$U = \frac{e^{i\pi/2}}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \quad (1.43)$$

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,

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

$$[\mathbf{P}_k, \mathbf{P}_{k+1}] \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^* \end{bmatrix} \begin{bmatrix} \mathbf{P}_k \\ \mathbf{P}_{k+1} \end{bmatrix} = [\mathbf{R}, \mathbf{Q}] \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{Q} \end{bmatrix}, \quad (1.44)$$

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. 1.6.1 is taken from refs. [5, 6]. Who wrote this down first we do not know, lineage certainly goes all the way back to Lagrange polynomials [13], but projection operators tend to get drowned in sea of algebraic details. Arfken and Weber [1] ascribe spectral decomposition (1.30) to Sylvester. Halmos [9] is a good early reference - but we like Harter's exposition [10–12] best, for its multitude of specific examples and physical illustrations. In particular, by the time we get to (1.27) we have tacitly assumed full diagonalizability of matrix 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 1.4. 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] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. (Academic, New York, 2013).
- [3] S. Axler, "Down with determinants!", *Amer. Math. Monthly* **102**, 139–154 (1995).
- [4] F. Chung and S.-T. Yau, "Discrete Green's functions", *J. Combin. Theory A* **91**, 191–214 (2000).
- [5] P. Cvitanović, "Group theory for Feynman diagrams in non-Abelian gauge theories", *Phys. Rev. D* **14**, 1536–1553 (1976).
- [6] P. Cvitanović, *Classical and exceptional Lie algebras as invariance algebras*, Oxford Univ. preprint 40/77, unpublished., 1977.
- [7] R. Giles and C. B. Thorn, "Lattice approach to string theory", *Phys. Rev. D* **16**, 366–386 (1977).
- [8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. (J. Hopkins Univ. Press, Baltimore, MD, 2013).
- [9] P. R. Halmos, *Finite-Dimensional Vector Spaces* (Princeton Univ. Press, Princeton NJ, 1948).
- [10] W. G. Harter, "Algebraic theory of ray representations of finite groups", *J. Math. Phys.* **10**, 739–752 (1969).
- [11] W. G. Harter, *Principles of Symmetry, Dynamics, and Spectroscopy* (Wiley, New York, 1993).
- [12] 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).
- [13] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed. (Prentice-Hall, Englewood Cliffs NJ, 1971).

EXERCISES

- [14] G. Papathanasiou and C. B. Thorn, “Worldsheet propagator on the lightcone worldsheet lattice”, *Phys. Rev. D* **87**, 066005 (2013).
- [15] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 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

$$J^t = e^{tA} = \mathbf{U}^{-1} e^{tA_D} \mathbf{U}, \quad A_D = \mathbf{U} \mathbf{A} \mathbf{U}^{-1}. \quad (1.45)$$

- 1.3. **The matrix square root.** Consider matrix

$$A = \begin{bmatrix} 4 & 10 \\ 0 & 9 \end{bmatrix}.$$

Generalize the square root function $f(x) = x^{1/2}$ to a square root $f(A) = A^{1/2}$ of a matrix A .

- a) Which one(s) of these is/are the square root of A

$$\begin{bmatrix} 2 & 2 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} -2 & 10 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} -2 & -2 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} 2 & -10 \\ 0 & -3 \end{bmatrix} ?$$

b) Assume that the eigenvalues of a $[d \times d]$ matrix are all distinct. How many square root matrices does such matrix have?

c) Given a $[2 \times 2]$ matrix A with a distinct pair of eigenvalues $\{\lambda_1, \lambda_2\}$, write down a formula that generates all square root matrices $A^{1/2}$. Hint: one can do this using the 2 projection operators associates with the matrix A . 2 points

- 1.4. **Real representation of complex eigenvalues.** (Verification of example 1.2.) 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 $\cdots + \lambda_k \mathbf{P}_k + \lambda_k^* \mathbf{P}_{k+1} + \cdots$ complex eigenvalue pair in the spectral decomposition (1.29) is now replaced by a real $[2 \times 2]$ matrix

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

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