Phys 6124 zipped!

World Wide Quest to Tame Math Methods

Predrag Cvitanović

Jun 16, 2023

Overview

> The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction.
> — Sidney Coleman

I am leaving the course notes here, not so much for the notes themselves –they cannot be understood on their own, without viewing the recorded live lectures– but for the hyperlinks to various source texts you might find useful later on in your research.

We change the topics covered year to year, in hope that they reflect better what a graduate student needs to know. This year's experiment is taking the course online. Let's work together to make it work for everyone in the course.

- *Course outline : An ode in 15 stanzas*
- *Course policy*
- *My teaching philosophy : Bologna*
- *How does one pronounce 'Euler'? 'Cvitanović'?*

After a while you might notice a pattern: Every week we start with something obvious that you already know, let mathematics lead us on, and then veers off and ends someplace amazing and highly non-intuitive.

Contents

mathematical methods - week 1

Linear algebra

Georgia Tech PHYS-6124

Homework HW #1 due Tuesday, August 25, 2020

== 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 *Trace-log of a matrix*	4 points
Exercise 1.2 *Stability, diagonal case*	2 points
Exercise 1.3 *The matrix square root*	4 points
Exercise 1.4 *Exponential of a matrix of Jordan form*	4 bonus points

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

edited October 11, 2022

Week 1 syllabus

Tuesday, August 18, 2020

> Diagonalizing the matrix: that's the key to the whole thing.
> — Governor Arnold Schwarzenegger

Anything prefixed by AWH, like "Kronecker product AWH eq. (2.55)" refers to Arfken, Weber & Harris [2] *Mathematical Methods for Physicists: A Comprehensive Guide* (Georgia Tech students can get it from GaTech Library). Light blue 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.

This week's lectures are related to AWH Chapter 2 *Determinants and matrices* (click here), and Chapter 6 *Eigenvalue problems* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run the

course playlist

- Sect. 1.2 *Matrix-valued functions*
  - AWH p. 113 Functions of Matrices
  - AWH Section *2.2 Matrices*
  - *Matrices : 2 kinds*
  - *Derivative of a matrix function*
  - *Exponential, logarithm of a matrix*
  - AWH Example 2.2.6 *Exponential of a diagonal matrix*
  - *Determinant is a volume*
  - *log det = tr log* (updated Aug 18, 2020)
  - *Multi-matrix functions* (optional, for the QM inclined)
- 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 (1.10). 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, and the Euler product to time-ordered product (1.11). The first is used in formal quantum-mechanical calculations, the second in practical, numerical calculations.
  - *Linear differential equations*
  - *Nonlinear differential equations*
- Sect. 1.4 *Eigenvalues and eigenvectors*
  Hamilton-Cayley equation, projection operators (1.21), any matrix function is evaluated by spectral decomposition (1.24). Work through example 1.3.

- AWH Section 6.1 *Eigenvalue Equations*
- *Eigenvalues and eigenvectors*
- *What's the deal with Hamilton-Cayley?*
- *Spectral decomposition*
- *Spectral decomposition and completeness*
- *Right, left eigenvectors*
- *A projection operators workout*
- *Jordan form*
- AWH p. 324 *Defective matrices* (optional, for QM inclined)
- *Are there Jordan form matrices in physics?* (optional, for QM inclined)

1.1 Other sources

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. A few resources that you might find helpful going forward:

Linear operators and matrices reading (optional reading for week 1, not required for this course):

- Stone and Goldbart [13], *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.

- In sect. 1.2 I make matrix functions appear easier than they really are. For an in-depth discussion, consult Golub and Van Loan [7] *Matrix Computations*, chap. 9 *Functions of Matrices* (click here). exercise 1.3

- Petersen and Pedersen *The Matrix Cookbook* (click here).

- 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.

- Sergey Loyka likes Aplevich [1] *The Essentials of Linear State-space Systems* (2000).

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].

Other online math methods courses we liked:

Ilya Kuprov *Mathematics for Chemists*.

1.2 Matrix-valued functions

What is a matrix?
—Werner Heisenberg (1925)
What is The Matrix?
—-Keanu Reeves (1999)

(optional, for QM inclined)

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

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

where $|\phi(t)\rangle$ is the quantum wave function at time 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 (1.12).

It would not be fair to students to expect a prior exposure to Heisenberg's matrix quantum mechanics (1.1), so if you do not 'get' the QM comments of this section, it's OK. It is not needed for what follows, and I'll do it in the class only if you request me to do 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, 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 (1.27): 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 ith vector subspace

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

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)

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

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.

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

$$\dot{x} = v(x) = Ax\,. \tag{1.3}$$

Solving this equation means finding the state space trajectory

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

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 (1.3)? 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

$$\frac{x(\delta t) - x(0)}{\delta t} = \lambda x(0)\,, \tag{1.4}$$

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

$$x(t) \approx \left(1 + \frac{t}{m}\lambda\right)^m x(0) \approx e^{t\lambda}x(0)\,. \tag{1.5}$$

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 $m \to \infty$, the term in parentheses converges to $e^{t\lambda}$. Consider now the matrix version of equation (1.4):

$$\frac{x(\delta t) - x(0)}{\delta t} = A\, x(0)\,. \tag{1.6}$$

A representative point x is now a vector in $\mathbb{R}^d$ acted on by the matrix A, as in (1.3). Denoting by $\mathbf{1}$ the identity matrix, and repeating the steps (1.4) and (1.5) we obtain Euler's formula

$$x(t) = J^t x(0)\,, \qquad J^t = e^{tA} = \lim_{m\to\infty}\left(\mathbf{1} + \frac{t}{m}A\right)^m\,. \tag{1.7}$$

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 (1.7) that describes linearized perturbation $\delta x(x_0, t)$?

$$x(t) = f^t(x_0)\,, \quad \delta x(x_0, t) \;=\; J^t(x_0)\, \delta x(x_0, 0)\,. \tag{1.8}$$

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

$$\dot{x}_i = v_i(x)\,, \quad \dot{\delta x}_i = \sum_j A_{ij}(x)\delta x_j \tag{1.9}$$

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 (1.7):

$$e^{tA} \;=\; \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \;=\; \lim_{m\to\infty}\left(\mathbf{1} + \frac{t}{m}A\right)^m\,. \tag{1.10}$$

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 (1.10):

$$\begin{aligned} J^t &= \lim_{m\to\infty} \prod_{n=m}^{1} (\mathbf{1} + \delta t A(x_n)) = \lim_{m\to\infty} \prod_{n=m}^{1} e^{\delta t\, A(x_n)} \\ &= \lim_{m\to\infty} e^{\delta t\, A(x_m)} e^{\delta t\, A(x_{m-1})} \cdots e^{\delta t\, A(x_2)} e^{\delta t\, A(x_1)}\,, \end{aligned} \tag{1.11}$$

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 $m \to \infty$ limit of this procedure is the formal integral

$$J^t_{ij}(x_0) = \left[\mathbf{T} e^{\int_0^t d\tau A(x(\tau))}\right]_{ij} , \tag{1.12}$$

where $\mathbf{T}$ stands for time-ordered integration, *defined* as the continuum limit of the successive multiplications (1.11). This integral formula for J^t is the finite time companion of the differential definition

$$\dot{J}(t) = A(t)J(t), \tag{1.13}$$

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

$$J^{t+t'}(x) = J^{t'}(x')\, J^t(x), \qquad \text{where } x' = f^t(x_0)\,, \tag{1.14}$$

which is an immediate consequence of the time-ordered product structure of (1.11). However, in practice J is evaluated by integrating differential equation (1.13) along with the ODEs (3.6) that define a particular flow.

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

$$\det(\mathbf{M} - \lambda \mathbf{1}) = \prod(\lambda_i - \lambda) = 0\,. \tag{1.15}$$

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

$$\mathbf{M}\,\mathbf{e}^{(i)} = \lambda_i \mathbf{e}^{(i)}\,. \tag{1.16}$$

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, $\mathrm{Re}\,\lambda_i \geq \mathrm{Re}\,\lambda_{i+1}$.

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

$$\mathbf{x} = x_1\,\mathbf{e}^{(1)} + x_2\,\mathbf{e}^{(2)} + \cdots + x_d\,\mathbf{e}^{(d)}\,. \tag{1.17}$$

From (1.16) 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

$$\prod_{i=1}^{d}(\mathbf{M}-\lambda_i\mathbf{1})=0\,. \tag{1.18}$$

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 (1.17) 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*

$$P_i=\prod_{j\neq i}\frac{\mathbf{M}-\lambda_j\mathbf{1}}{\lambda_i-\lambda_j}\,, \tag{1.19}$$

which are *orthogonal* and *complete*:

$$P_iP_j=\delta_{ij}P_j\,,\quad(\text{no sum on } j)\,,\qquad\sum_{i=1}^{r}P_i=\mathbf{1}\,, \tag{1.20}$$

with the dimension of the ith subspace given by $d_i=\mathrm{tr}\,P_i$. For each distinct eigenvalue λ_i of $\mathbf{M}$,

$$(\mathbf{M}-\lambda_j\mathbf{1})P_j=P_j(\mathbf{M}-\lambda_j\mathbf{1})=0\,, \tag{1.21}$$

the colums/rows of 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 (1.21) that λ_i is the eigenvalue of $\mathbf{M}$ on P_i subspace:

$$\mathbf{M}\,P_i=\lambda_iP_i\qquad(\text{no sum on } i)\,. \tag{1.22}$$

Using $\mathbf{M}=\mathbf{M}\,\mathbf{1}$ and completeness relation (1.20) we can rewrite $\mathbf{M}$ as

$$\mathbf{M}=\lambda_1P_1+\lambda_2P_2+\cdots+\lambda_dP_d\,. \tag{1.23}$$

Any matrix function $f(\mathbf{M})$ takes the scalar value $f(\lambda_i)$ on the P_i subspace, $f(\mathbf{M})\,P_i=f(\lambda_i)\,P_i$, and is thus easily evaluated through its *spectral decomposition* (see AWH Exercise 3.5.34)

remark 1.1

$$f(\mathbf{M})=\sum_i f(\lambda_i)P_i\,. \tag{1.24}$$

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 (1.21) every column of 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 (1.20), they are mutually orthogonal,

$$\mathbf{e}_{(i)} \cdot \mathbf{e}^{(j)} = c\, \delta_i^j \,. \tag{1.25}$$

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 (1.18) now takes form*

$$\prod_{\alpha=1}^{r} (\mathbf{M} - \lambda_\alpha \mathbf{1})^{d_\alpha} = 0 \,, \qquad \sum_\alpha d_\alpha = d \,. \tag{1.26}$$

We distinguish two cases:

$\mathbf{M}$ can be brought to diagonal form. *The characteristic equation (1.26) can be replaced by the minimal polynomial,*

$$\prod_{\alpha=1}^{r} (\mathbf{M} - \lambda_\alpha \mathbf{1}) = 0 \,, \tag{1.27}$$

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

$$\mathbf{M}\, \mathbf{e}^{(\alpha,k)} = \lambda_i \mathbf{e}^{(\alpha,k)} \,, \tag{1.28}$$

on a d_α-dimensional subspace spanned by a linearly independent set of basis eigenvectors $\{\mathbf{e}^{(\alpha,1)}, \mathbf{e}^{(\alpha,2)}, \cdots, \mathbf{e}^{(\alpha,d_\alpha)}\}$. 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.

$\mathbf{M}$ 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.* *(optional, for QM inclined)*

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

$$\lambda_{1,2} = \frac{1}{2} \mathrm{tr}\, \mathbf{M} \pm \frac{1}{2} \sqrt{(\mathrm{tr}\, \mathbf{M})^2 - 4 \det \mathbf{M}} \tag{1.29}$$

are the roots of the characteristic (secular) equation (1.15):

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

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

$$P_1=\frac{\mathbf{M}-\lambda_2\mathbf{1}}{\lambda_1-\lambda_2}\,,\qquad P_2=\frac{\mathbf{M}-\lambda_1\mathbf{1}}{\lambda_2-\lambda_1}\,,\qquad \lambda_1\neq\lambda_2\,. \tag{1.30}$$

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 $[2\times2]$ matrix the Jordan form is* *(optional, for QM inclined)*

$$\mathbf{M}=\left[\begin{array}{cc}\lambda&1\\0&\lambda\end{array}\right],\qquad \mathbf{e}^{(1)}=\begin{bmatrix}1\\0\end{bmatrix},\ \ \mathbf{v}^{(2)}=\begin{bmatrix}0\\1\end{bmatrix}.$$

$\mathbf{v}^{(2)}$ *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*

$$\mathbf{M}^m=\left[\begin{array}{cc}\lambda^m&m\lambda^{m-1}\\0&\lambda^m\end{array}\right],$$

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

$$e^{t\mathbf{M}}\begin{pmatrix}u\\v\end{pmatrix}=e^{t\lambda}\begin{pmatrix}u+tv\\v\end{pmatrix} \tag{1.31}$$

picks up a power-low 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 $[2\times2]$ matrix*

$$\mathbf{M}=\left[\begin{array}{cc}4&1\\3&2\end{array}\right].$$

Its eigenvalues $\{\lambda_1,\lambda_2\}=\{5,1\}$ are the roots of (1.29):

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

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

$$\left[\begin{array}{cc}4&1\\3&2\end{array}\right]^2-6\left[\begin{array}{cc}4&1\\3&2\end{array}\right]+5\left[\begin{array}{cc}1&0\\0&1\end{array}\right]=\left[\begin{array}{cc}0&0\\0&0\end{array}\right].$$

Associated with each root λ_i is the projection operator *(1.30)*

$$\begin{aligned} P_1 &= \frac{1}{4}(\mathbf{M}-\mathbf{1}) = \frac{1}{4}\left[\begin{array}{cc} 3 & 1 \\ 3 & 1 \end{array}\right] && (1.32) \\ P_2 &= \frac{1}{4}(\mathbf{M}-5\cdot\mathbf{1}) = \frac{1}{4}\left[\begin{array}{cc} 1 & -1 \\ -3 & 3 \end{array}\right]. && (1.33) \end{aligned}$$

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

$$\mathbf{M}^7 - 3\cdot\mathbf{1} = (5^7-3)P_1 + (1-3)P_2 = \left[\begin{array}{cc} 58591 & 19531 \\ 58593 & 19529 \end{array}\right].$$

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:

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

with overall scale arbitrary. The matrix is not symmetric, so $\{\mathbf{e}^{(j)}\}$ do not form an orthogonal basis. The left-right eigenvector dot products $\mathbf{e}_{(j)}\cdot\mathbf{e}^{(k)}$, however, are orthogonal as in (1.25), by inspection.

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

But A may not have d linearly independant eigenvectors, making F singular and forcing us to take a different route. To illustrate this, consider $[2\times2]$ 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:

$$B = \left[\begin{array}{cc} \lambda & 0 \\ 0 & \mu \end{array}\right], \qquad B = \left[\begin{array}{cc} \lambda & 1 \\ 0 & \lambda \end{array}\right], \qquad B = \left[\begin{array}{cc} \mu & -\omega \\ \omega & \mu \end{array}\right].$$

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

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

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)*

Example 1.5. ***Determinants and traces.***

The usual textbook expression for a determinant is the sum of products of all permutations

$$\det M = \sum_{\{\pi\}} (-1)^{\pi} M_{1,\pi_1} M_{2,\pi_2} \cdots M_{m,\pi_m} \tag{1.34}$$

where M is a $[m\times m]$ matrix, $\{\pi\}$ denotes the set of permutations of m symbols, π_k is the permutation π applied to k, and $(-1)^{\pi} = \pm 1$ is the parity of permutation π. For example, for a $[2\times 2]$ matrix, the permutations are $\{\pi_m\} = \{(1)(2), (12)\}$, so

$$\det M = M_{11}M_{22} - M_{12}M_{21}\,, \tag{1.35}$$

for a $[3\times 3]$ matrix

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}$$

there are $6 = 3!$ permutations,

$$\begin{aligned} \det M &= M_{11}M_{22}M_{33} - M_{11}M_{23}M_{32} - M_{12}M_{21}M_{33} + M_{12}M_{23}M_{31} \\ &\quad + M_{13}M_{21}M_{32} - M_{13}M_{22}M_{31}\,, \end{aligned} \tag{1.36}$$

and so on. Not very illuminating.

But if $M = T - \lambda\mathbf{1}$, evaluation of the $[2\times 2]$ case,

$$\det(T - \lambda\mathbf{1}) = (T_{11} - \lambda)(T_{22} - \lambda) - M_{12}M_{21} = \lambda^2 + (\mathrm{tr}\, T)\lambda + \det T\,, \tag{1.37}$$

used in (1.29), offers a hint of better things to come. This way of computing determinants is generalized to any $[m\times m]$ matrix in ref. [6], sect. 6.4 Determinants *(click here)*.

The $\ln\det M = \mathrm{tr}\,\ln M$ relation, valid for any square matrix M (even the infinite dimensional 'trace class' operators M, as long as all $|\mathrm{tr}\, M^k|$ are bounded) offers a powerful alternative, universally used, for evaluating determinants.

First, observe that both the determinant and the trace are invariant under similarity transformations $\hat{M} = S^{-1}M\,S$, $\det S \neq 0$:

$$\begin{aligned} \det \hat{M} &= \det(S^{-1}M\,S) = (\det S^{-1})\,(\det M)\,(\det S) = \det M \\ \mathrm{tr}\, \hat{M}^k &= \mathrm{tr}\, S^{-1}M\,S \cdots S^{-1}M\,S = \mathrm{tr}\, M\,S \cdots S^{-1}M\,S\,S^{-1} = \mathrm{tr}\, M^k\,, \end{aligned} \tag{1.38}$$

so any quantity, in particular the eigenvalues of M, expressed in terms of its traces and its determinant is also invariant under all linear coordinate changes.

Next, consider the characteristic polynomial (1.15) of $[m \times m]$ matrix T, and change the variable to $z = 1/\lambda$ in $\det(T - \lambda\mathbf{1})$. The zeros $z_j = 1/\lambda_j$ of

$$\det(\mathbf{1} - zT) = 0 \tag{1.39}$$

now yield the non-zero eigenvalues λ_j of T. That $\lambda_j = 0$ eigenvalues are gone is a blessing; nobody liked them anyway. By the determinant–trace relation $\ln\det M = \mathrm{tr}\ln M$, the determinant of $M = 1 - zT$ is always *expressible as* exercise 1.1

$$\det(1 - zT) = \exp(\mathrm{tr}\,\ln(1 - zT)) = e^{-\sum_{n=1}\frac{z^n}{n}\mathrm{tr}\,T^n}\,. \tag{1.40}$$

We evaluate such formulas in two steps. First, expand $\exp(f(z))$ as Taylor series in $f(z)$

$$\det(1 - zT) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\sum_{n=1}\frac{z^n}{n}\mathrm{tr}\,T^n\right)^k$$

Then expand $(\cdots)^k$ as series in z^n and combine terms of order z^n. The result is central to much statistical physics and field theory, where it is known as the cumulant expansion*:*

$$\begin{aligned}
\det(1 - zT) &= 1 - z\,\mathrm{tr}\,T - \frac{z^2}{2}\left((\mathrm{tr}\,T)^2 - \mathrm{tr}\,T^2\right) \\
&\quad -\frac{z^3}{3!}\left((\mathrm{tr}\,T)^3 - 3\,(\mathrm{tr}\,T)\,\mathrm{tr}\,T^2 + 2\,\mathrm{tr}\,T^3\right) \qquad (1.41)\\
&\quad -\frac{z^4}{4!}\left((\mathrm{tr}\,T)^4 - 3\,(2\,(\mathrm{tr}\,T)^2 - \mathrm{tr}\,T^2)\,\mathrm{tr}\,T^2 + 8\,\mathrm{tr}\,T\,\mathrm{tr}\,T^3 - 6\,\mathrm{tr}\,T^4\right) - \ldots
\end{aligned}$$

If T is an $[m \times m]$ matrix, the characteristic polynomial is at most of order m, so the infinity of coefficients of z^n must vanish exactly *for $n > m$! For example, for a $[2 \times 2]$ matrix, the z^2 coefficient in (1.41) is a traces expansion for the determinant (1.35),*

$$\det(T) = \frac{1}{2}\left((\mathrm{tr}\,T)^2 - \mathrm{tr}\,T^2\right)\,, \tag{1.42}$$

and for a $[3 \times 3]$ matrix, the z^3 coefficient in (1.41) is a traces expansion for the determinant (1.36)

$$\begin{aligned}
\det(T) &= \frac{1}{3!}\left((\mathrm{tr}\,T)^3 - 3\,(\mathrm{tr}\,T)\,\mathrm{tr}\,T^2 + 2\,\mathrm{tr}\,T^3\right) \\
&= M_{11}M_{22}M_{33} - M_{11}M_{23}M_{32} - M_{12}M_{21}M_{33} + M_{12}M_{23}M_{31} \\
&\quad + M_{13}M_{21}M_{32} - M_{13}M_{22}M_{31}\,, \qquad (1.43)
\end{aligned}$$

as you can verify by hand, if you do not believe me (you should never believe anything anyone over 30 says). If you still do not believe me, verify that the z^4 coefficient vanishes

$$0 = \frac{1}{4!}\left(-6(\mathrm{tr}\,T)^2\mathrm{tr}\,T^2 + 8(\mathrm{tr}\,T)\mathrm{tr}\,T^3 + 3(\mathrm{tr}\,T^2)^2 - 6\mathrm{tr}\,T^4 + (\mathrm{tr}\,T)^4\right)$$

for $m = 1, 2, 3$, but is a traces expansion for the determinant of a $[4 \times 4]$ matrix. If you need to know more, these relations were noted by Albert Girard (1629), so they are called Newton's (1666) identities.

Note also that derivative of (1.40) relates the determinant to the resolvent*,*

$$\begin{aligned}
-z\frac{d}{dz}\ln\det(1 - zT) &= -\mathrm{tr}\left(z\frac{d}{dz}\ln(1 - zT)\right) \\
&= \mathrm{tr}\,\frac{zT}{1 - zT} = \sum_{k=1}^{\infty} z^n\mathrm{tr}\,(T^n)\,, \qquad (1.44)
\end{aligned}$$

a simple but very useful relation expressing a determinant in terms of traces.

What are all these relationships? Have a fresh look at the Hamilton-Cayley theorem (1.18) that states that the matrix $\mathbf{M}$ satisfies its characteristic equation, and to be specific, look at the $m = 3$ case. The Hamilton-Cayley characteristic equation *expanded in terms of traces is*

$$0 \quad = \quad T^3 - (\mathrm{tr}\, T)\, T^2 + \frac{1}{2}\left((\mathrm{tr}\, T)^2 - \mathrm{tr}\, T^2\right) T - (\det T)\, \mathbf{1}\,. \qquad (1.45)$$

This is the first 3 terms of the cumulant expansion *(1.41), with λ restored by $z \to 1/\lambda$, i.e., the characteristic equation for A $[3 \times 3]$ matrix, and the λ replaced by T. The Hamilton-Cayley formula says that whenever you see $[m \times m]$ matrix T^m you can express it in terms of T^{m-1}, T^{m-1}, $\cdots$, T.*

To be very specific and pedestrian, consider the $[3 \times 3]$ matrix

$$\begin{aligned} T \quad &= \quad \begin{pmatrix} 2 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 2 \end{pmatrix}, \qquad T^2 = \begin{pmatrix} 12 & 20 & 12 \\ 20 & 44 & 20 \\ 12 & 20 & 12 \end{pmatrix} \\ \mathrm{tr}\, T \quad &= \quad 10\,, \qquad \mathrm{tr}\, T^2 = 68\,. \end{aligned} \qquad (1.46)$$

From the shape of T clearly $\det T = 0$, so the characteristic equation is

$$\begin{aligned} 0 \quad &= \quad \left(T^2 - (\mathrm{tr}\, T)\, T + \frac{1}{2}\left((\mathrm{tr}\, T)^2 - \mathrm{tr}\, T^2\right) \mathbf{1}\right) T\,. \\ &= \quad \left(T^2 - 10\, T + \frac{1}{2}\, (100 - 68)\, \mathbf{1}\right) T \\ &= \quad \left(T^2 - 10\, T + 16\, \mathbf{1}\right) T = (T - 8\, \mathbf{1})(T - 2\, \mathbf{1})\, T\,, \end{aligned} \qquad (1.47)$$

with eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\} = \{8, 2, 0\}$.

For the associated projection operators, see (2.30).

Commentary

Remark 1.1. Projection operators. The construction of projection operators given in sect. 1.4 is taken from refs. [4, 5]. Sylvester [14] wrote down the spectral decomposition (1.24) in 1883 in the form we use, but lineage certainly goes all the way back to 1795 Lagrange polynomials [12], and Euler 1783. Often projection operators get drowned in sea of algebraic details. Halmos [8] is a good early reference - but we like Harter's exposition [9–11] best, for its multitude of specific examples and physical illustrations. In particular, by the time we get to (1.21) 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] J. D. Aplevich, *The Essentials of Linear State-space Systems* (Wiley, 2000).

[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] P. Cvitanović, "Group theory for Feynman diagrams in non-Abelian gauge theories", Phys. Rev. D **14**, 1536–1553 (1976).

[5] P. Cvitanović, Classical and exceptional Lie algebras as invariance algebras, Oxford Univ. preprint 40/77, unpublished., 1977.

[6] P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Groups* (Princeton Univ. Press, Princeton NJ, 2008).

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. (J. Hopkins Univ. Press, Baltimore, MD, 2013).

[8] P. R. Halmos, *Finite-Dimensional Vector Spaces* (Princeton Univ. Press, Princeton NJ, 1948).

[9] W. G. Harter, "Algebraic theory of ray representations of finite groups", J. Math. Phys. **10**, 739–752 (1969).

[10] W. G. Harter, *Principles of Symmetry, Dynamics, and Spectroscopy* (Wiley, New York, 1993).

[11] 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).

[12] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed. (Prentice-Hall, Englewood Cliffs NJ, 1971).

[13] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

[14] J. J. Sylvester, "XXXIX. On the equation to the secular inequalities in the planetary theory", Philos. Mag. **16**, 267–269 (1883).

Exercises

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

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

for an arbitrary finite dimensional square matrix M, $\det M \neq 0$. (If you are not getting it, see AWH(3.171).)

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

$$J^t = e^{tA} = \mathbf{U}^{-1} e^{tA_D} \mathbf{U}\,, \quad A_D = \mathbf{U} \mathbf{A} \mathbf{U}^{-1}\,. \tag{1.48}$$

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 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 [2×2] 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. **Exponential of a matrix of Jordan form.** A matrix B with all eigenvalues degenerate that cannot be diagonalized can always be brought to upper triangular Jordan form $B = \lambda\mathbf{1} + E$, where E is its strictly upper bidiagonal part. As an example, consider [4×4] matrix B, with

$$E = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.49}$$

a) Write down E, E^2, E^2, E^3, ...

b) Write down explicitly the exponential [4×4] matrix function $\exp(tE)$.

c) Bonus points, some assembly required: Work out the kth term in the Taylor expansion of a $[d \times d]$ matrix function $f(B)$, $B = \lambda\mathbf{1} + E$ a $[d \times d]$ matrix,

$$f(B) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (B - x_0\mathbf{1})^k\,. \tag{1.50}$$

.

A side remark to the masters of QM: E is a 'raising operator'.

Chapter 1 solutions: Linear algebra

Solution 1.1 - Trace-log of a matrix.

1) Consider $M = \exp A$.

$$\det M = \lim_{n\to\infty} \det \left(\mathbf{1} + \frac{1}{n}A\right)^n = \lim_{n\to\infty} (\mathbf{1} + \frac{1}{n}\mathrm{tr}\, A + \ldots)^n = \exp(\mathrm{tr}\,(\ln M))$$

2) A rephrasing of the solution 1): evaluate $\frac{d}{dt}\det\left(e^{t\ln M}\right)$ by definition of derivative in terms of infinitesimals. (Kasper Juel Eriksen)

3) Here is an example of wrong/incomplete answer, hiding behind fancier notation: This identity makes sense for a matrix $M \in \mathbb{C}^{n\times n}$, if $|\prod_{i=1}^{n}\lambda_i| < \infty$ and $\{|\lambda_i| > 0, \forall i\}$, where $\{\lambda_i\}$ is a set of eigenvalues of M. Under these conditions there exist a nonsingular $O : M = ODO^{-1}$, $D = \mathrm{diag}[\{\lambda_i, i = 1, \ldots, n\}]$. If $f(M)$ is a matrix valued function defined in terms of power series then $f(M) = Of(D)O^{-1}$, and $f(D) = \mathrm{diag}[\{f(\lambda_i)\}]$. Using these properties and cyclic property of the trace we obtain

$$\exp(\mathrm{tr}\,(\ln M)) = \exp\left(\sum_i \ln\lambda_i\right) = \prod_i \lambda_i = \det\,(M)$$

What's wrong about it? If a matrix with degenerate eigenvalues, $\lambda_i = \lambda_j$ is of Jordan type, it cannot be diagonalized, so a bit more of discussion is needed to show that the identity is satisfied by upper-triangular matrices.

4) First check that this is true for any Hermitian matrix M. Then write an arbitrary complex matrix as sum $M = A + zB$, A, B Hermitian, Taylor expand in z and prove by analytic continuation that the identity applies to arbitrary M. (David Mermin)

5) Suppose $\ln M = A$, then $e^A = M$. Let $A = C^{-1}JC$, where J is the Jordan canonical form of A, then we have $e^J = CMC^{-1}$. J can be written as $D+N$, where D is diagonal and N is nilpotent with diagonal elements be zero. So e^J has the same diagonal elements with e^D, we have that the eigenvalues of J is the logarithm of eigenvalues of CMC^{-1}. J has the same eigenvalues with A, CMC^{-1} has the same eigenvalues with M. So the eigenvalues of A is the logarithm of eigenvalues of M. Determinant of a matrix is product of its all eigenvalues. Combine all these, we have $\det M = e^{tr\ln M}$. (Lei Zhang)

Solution 1.2 - Stability, diagonal case.

$$\begin{aligned}
J^t &= e^{t\mathbf{A}} = \sum_{k=0}^{\infty}\frac{1}{k!}(t\mathbf{A})^k \\
&= \sum_{k=0}^{\infty}\frac{1}{k!}t^k(\mathbf{U}^{-1}A_D\mathbf{U})^k = \mathbf{U}^{-1}\sum_{k=0}^{\infty}\frac{1}{k!}t^k(A_D)^k\mathbf{U} \\
&= \mathbf{U}^{-1}e^{tA_D}\mathbf{U}\,.
\end{aligned}$$

(Han Liang)

Solution 1.2 - Stability, diagonal case. *The relation (1.48) can be verified by noting* *that the defining Euler product can be rewritten as*

$$\begin{aligned} e^{t\mathbf{A}} &= \left(\mathbf{U}\mathbf{U}^{-1} + \frac{t\mathbf{U}\mathbf{A}_D\mathbf{U}^{-1}}{m}\right)\left(\mathbf{U}\mathbf{U}^{-1} + \frac{t\mathbf{U}\mathbf{A}_D\mathbf{U}^{-1}}{m}\right)\cdots \\ &= \mathbf{U}\left(I + \frac{t\mathbf{A}_D}{m}\right)\mathbf{U}^{-1}\mathbf{U}\left(I + \frac{t\mathbf{A}_D}{m}\right)\mathbf{U}^{-1}\cdots = \mathbf{U}e^{t\mathbf{A}_D}\mathbf{U}^{-1}. \end{aligned} \tag{1.51}$$

Solution 1.3 - The matrix square root. .
a) It is easy to check that

$$A = \begin{bmatrix} 4 & 10 \\ 0 & 9 \end{bmatrix} = \left(A_{ij}^{1/2}\right)^2$$

for the matrices

$$\begin{aligned} A_{++}^{1/2} &= \begin{bmatrix} 2 & 2 \\ 0 & 3 \end{bmatrix}, \qquad A_{+-}^{1/2} = \begin{bmatrix} -2 & 10 \\ 0 & 3 \end{bmatrix} \\ A_{--}^{1/2} &= \begin{bmatrix} -2 & -2 \\ 0 & -3 \end{bmatrix}, \quad A_{-+}^{1/2} = \begin{bmatrix} 2 & -10 \\ 0 & -3 \end{bmatrix} \end{aligned} \tag{1.52}$$

Being upper-triangular, the eigenvalues of the four matrices can be read off their diagonals: there are four square root $\pm$ eigenvalue combinations {3,2}, {-3,2}, {3,-2}, and {-3,-2}.

Associated with each set $\lambda_i \in \{\lambda_1, \lambda_2\}$ is the projection operator *(1.30)*

$$P_{ij}^{(1)} = \frac{1}{\lambda_1 - \lambda_2}(A_{ij}^{1/2} - \lambda_2 \mathbf{1}) = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \tag{1.53}$$

$$P_{ij}^{(2)} = \frac{1}{\lambda_2 - \lambda_1}(A_{ij}^{1/2} - \lambda_1 \mathbf{1}) = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}. \tag{1.54}$$

This calculation reveals that all 'square root' matrices have the same projection operators / eigenvectors as the matrix A itself. We know $\{\lambda_1, \lambda_2\}$ and $P^{(\alpha)}$ for A, and the four 'square root' eigenvalues are clearly $\{\pm\lambda_1^{1/2}, \pm\lambda_2^{1/2}\}$. That suggest finding the 'square root' matrices by reverse-engineering (1.53), (1.54):

$$A_{ij}^{1/2} = (\lambda_1 - \lambda_2)P_{ij}^{(1)} + \lambda_2 \mathbf{1}.$$

For example,

$$A_{+-}^{1/2} = (+3 - (-2))\begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} + (-2)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

b) If the eigenvalues of a [$d \times d$] matrix are all distinct, the matrix is diagonalizable, so the number of square root $\pm$ combinations is 2^d. However, for general matrices things can get crazy - there can be no, or some, or ∞ *of 'square root' matrices.*

Solution 1.4 - Exponential of a matrix of Jordan form. *A matrix B with all eigenvalues degenerate of upper triangular Jordan form $B = \lambda\mathbf{1} + E$. Consider [4×4] matrix B, with*

a) Write down E, E^2, E^2, E^3, ...:

$$E^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad E^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad , \quad E^4 = 0. \tag{1.55}$$

b) *The Taylor expansion of the exponential matrix function $f(E) = e^{tE}$ stops at the 4th term*

$$e^{tE} = \mathbf{1} + tE + \frac{t^2}{2}E^2 + \frac{t^3}{6}E^3 = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{6} \\ 0 & 1 & t & \frac{t^2}{2} \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

c) *Bonus points, some assembly required: Write down the formula for the exponential matrix function $f(B)$, $B = \lambda\mathbf{1} + E$ with a single eigenvalue, but of arbitrary dimensions.*

$$e^{tB} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \,. \tag{1.56}$$

mathematical methods - week 2

Eigenvalue problems

Georgia Tech PHYS-6124

Homework HW #2 due Tuesday, September 1, 2020

== 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 exerWeek2.tex

Exercise 2.1 *Three masses on a loop*	8 points
Exercise 2.2 *Examples of singular value decomposition*	2 points + 3 bonus points

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

edited May 15, 2021

Week 2 syllabus

Tuesday, August 25, 2020

> If I had had more time, I would have written less
> — Blaise Pascal, a remark made to a correspondent

Tuesday's lecture is related to AWH Chapter 6 *Eigenvalue problems* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run

the course playlist

- Please do not get intimated by the length of this week's notes - they are here more for me than for you, as notes on these topics for future reference. If you understand the online lectures and can solve the exercises, you are good. The notes you can quickly skim over...
  - Sect. 2.2 *Using symmetries*
  - Sect. 2.3 *Normal modes*: The free vibrations of systems, for undamped systems with total energy conserved for which the frequencies of oscillation are real.

  Normal modes
  - Example 2.1 *Vibrations of a classical CO_2 molecule*

  A Hamiltonian with a symmetry (4:46 min)

  CO_2 molecule (4:07 min)

  Projection operators (5:33 min)

  (Anti)symmetric subspaces (3:04 min)

  Zero mode (5:19 min)

  AWH Example 6.2.3 *Degenerate eigenproblem*

  AWH Example 6.5.2 *Normal modes*

- Matrix decompositions in data science
  - Sect. 2.4 *Singular Value Decomposition*

  Matrices: physics vs data science

  Singular value decomposition (SVD)

  SVD sample calculation

2.1 Other sources

Normal modes are important in aeronautical and mechanical engineering (optional reading for week 2, not required for this course):

- MIT 16-07-dynamics is a typical mathematical methods in engineering course. Normal modes are discussed here.

- Example 2.2 pen & paper derivation of normal-modes of the ring of N asymmetric pairs of oscillators (from Gutkin lecture notes example 5.1 *C_n symmetry*).

- Srdjan Ostojić @ostojic_srdjan writes: The singular value decomposition (SVD) course by @eigensteve is great: "These lectures go into depth on the singular value decomposition (SVD), one of the most widely used algorithms for data processing, reduced-order modeling, and high-dimensional statistics, following Chapter 1 of *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* by Brunton and Kutz [2], with databookuw website and chapters."

- We like the discussion of norms, least square problems, and differences between singular value and eigenvalue decompositions in Trefethen and Bau [4], cited in sect. 2.4.1.

- Andrew: In *Understanding SVD* Reza Bagheri develops SVD step-by-step, starting with the concept of eigenvalues through eigenvalue decomposition and then to SVD. I found it good to review some of the linear algebra I had forgotten. It is long, but it takes time to develop each concept which is a style I find very helpful.

- If you later need SVD in your research, Cline and Dhillon [3] *Computation of the singular value decomposition* seems to be a handy cookbook.

- *Eigen Grandito* - `u/cactus`'s *Principal Components Analysis of the Taco Bell menu*, an NumPy SVD exploration of the *Onion* (1998) classic on Taco Bell's revolutionary Grandito

- ChaosBook sect. 6.1 explains the geometrical intuition behind matrix decompositions.

- In ChaosBook remark 6.1. *Lyapunov exponents are uncool* Predrag claims that SVD is the wrong thing in dynamics.

- If instead, bedside crocheting is your thing, click here.

2.2 Using symmetries

The big idea #1 of this is week is *symmetry*.

If our physical problem is defined by a (perhaps complicated) Hamiltonian $\mathbf{H}$, another matrix $\mathbf{M}$ (hopefully a very simple matrix) is a symmetry if it commutes with the Hamiltonian

$$[\mathbf{M}, \mathbf{H}] = 0\,. \tag{2.1}$$

Than we can use the spectral decomposition (1.24) of $\mathbf{M}$ to block-diagonalize $\mathbf{H}$ into a sum of lower-dimensional sub-matrices,

$$\mathbf{H} = \sum_i \mathbf{H}_i\,, \qquad \mathbf{H}_i = P_i \mathbf{H} P_i\,, \tag{2.2}$$

and thus significantly simplify the computation of eigenvalues and eigenvectors of $\mathbf{H}$, the matrix of physical interest.

2.3 Normal modes

> The big idea #2 of this is week is : *many body systems* (molecules, neuronal networks, ...) are ruled by *collective modes*, not individual particles (atoms, neurons, ...).

In the linear, harmonic oscillator approximation, the classical dynamics of a molecule is governed by the Hamiltonian

$$H = \sum_{i=1}^{N} \frac{m_i}{2} \dot{x}_i^2 + \frac{1}{2} \sum_{i,j=1}^{N} x_i^\top V_{ij} x_j \,,$$

where $\{x_i\}$ are small deviations from the equilibrium, resting points of the molecules labelled i. V_{ij} is a symmetric matrix, so it can be brought to a diagonal form by an orthogonal transformation, to a set of N uncoupled harmonic oscillators or *normal modes* of frequencies $\{\omega_i\}$.

$$x \to y = Ux, \qquad H = \sum_{i=1}^{N} \frac{m_i}{2} \left(\dot{y}_i^2 + \omega_i^2 y_i^2 \right) \,. \tag{2.3}$$

2.4 Singular Value Decomposition

> Everybody knows that the SVD is the best matrix decomposition !!!
>
> — @Daniela_Witten, 21 July 2020

Daniela's Twitter lecture (tweaked by Predrag): If you are in statistics or data science, SVD is the #1 matrix decomposition, and likely the only one you will ever need. And believe me: you are going to need it.

In data science, one often deals with a very large data set X that can be laid out as an rectangular array, vertically arbitrarily high (n time measurements $x_{1j}, x_{2j}, \cdots, x_{nj}$; n faces), and horizontally relatively short (m neuronal voltages $x_{k1}, x_{k2}, \cdots, x_{km}$; m facial features).

What does the SVD do? You give me an $[n \times m]$, $n \geq m$ rectangular matrix X, and I'll give you back 3 matrices, an $[n \times m]$ rectangular matrix U, a diagonal $[m \times m]$ matrix Σ, and unitary $[m \times m]$ matrix V that together "decompose" the matrix X:

$$X = U \Sigma V^T \,. \tag{2.4}$$

U and V are orthogonal matrices (if X is complex, unitary matrices),

$$U^T U = V^T V = V V^T = I_{[m \times m]} \tag{2.5}$$

Σ is diagonal with nonnegative and decreasing elements:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0\,. \tag{2.6}$$

Some terminology: the diagonal elements of Σ are the *singular values*, and the columns of U and V are the *left* and *right singular vectors* u_k, v_j. Multiply (2.4) from right by V. This implies that if I have the singular value σ_j and the short $[m \times 1]$ singular vector v_j, I can multiply it with my data array X to compute the tall, $[n \times 1]$ singular vector u_j to

$$X\, v_j = \sigma_j u_j\,. \tag{2.7}$$

How do I compute the singular eigenvalues? From (2.4) it follows that V is a rotation that diagonalizes the symmetric $X^T X = \{\sum_j x_{kj} x_{j\ell}\}$ correlation matrix

$$X^T X = V \Sigma^2 V^T\,. \tag{2.8}$$

So this gives us v_k and σ_k (one always picks the positive root of σ_k^2) which we label by the decreasing eigenvalues convention (2.6), and evaluate u_j using (2.7). Why σ_k^2? Rectangular matrix X is dimensionally a strange beast; it relates bricks to oranges, and it's transpose returns oranges to bricks. The result is an (hyper)ellipsoid, with singular vectors as semiaxes, and singular values as lengths along the semiaxes.

Simple as that. What makes this decomposition special (and unique) is the particular set of properties of U, Σ, and V.

Don't be fooled tho: $UU^T \neq I_{n \times n}$!!!!!!! In layperson's terms, the columns of U and V are special: the squared elements of each column of U and V sums to 1, and also the inner product (dot product) between each pair of columns in U equals 0. And the inner product between each pair of columns of V equals 0.

First of all, let's marvel that this decomposition is not only possible, but easily computable, and even unique (up to sign flips of columns of U and V). Like, why on earth should *every matrix* X be decomposable in this way?

Magic, that's why. OK, so, its existence is magic. But, is it also useful? Well, YES.

Suppose you want to approximate X with a pair of vectors: that is, a rank-1 approx. Well, the world's best rank-1 approximation to X, in terms of residual sum of squares, is given by the first columns of U and V:

$$X \approx \sigma_1 u_1 v_1^T \tag{2.9}$$

is literally the best you can do!!

OK, but what if you want to approximate X using two pairs of vectors (a rank-2 approximation)? Just calculate

$$X \approx \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T\,, \tag{2.10}$$

and call it a day.

Want an even better approximation, using rank-k? You *literally can't beat this one*

$$X \approx \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots + \sigma_k u_k v_k^T \tag{2.11}$$

so please don't bother trying.

OK, so, the SVD gives me the best possible way to approximate any matrix. What is this good for??!!

Ever heard of *principal components analysis* (PCA)? This is just the SVD (after centering columns of X to have mean 0). Columns of V are PC loading vectors. Columns of U (up to scaling) are PC score vectors.

Bam!!!

How about if you want to impute missing values in your data matrix X? (In finance 'impute' means "to assign (a value) to something by inference from the value of the products to which it contributes.) Assuming that the elements of X are randomly missing (rather than, for instance, larger elements more likely to be missing), the SVD gives you an effective and easy way to impute those values!!

First, fill in missing elements using (say) the column mean. Then, compute SVD to get rank-k approx for (e.g.) $k = 3$. Replace the missing elements with the elements of rank-k approx. Rinse and repeat until your answer stops changing.

Voila!! I am not making this up!! The SVD is like a matrix X-ray. For instance:

$$X^T X = V\Sigma^2 V^T\,, \quad (X^T X)^{-1} = V\Sigma^{-2} V^T \tag{2.12}$$

$$X(X^T X)^{-1} X^T = UU^T\,. \tag{2.13}$$

Take a minute to breathe this in. Formulas (2.12), (2.13) have no U's, and the 3rd (hat matrix from least squares) has no V's or Σ's!

Wowzers!

Now, you may ask "well what about the eigen-decomposition?" Well I can dispense with that concern in 1 tweet. A symmetric matrix A (the only type worth eigen-decomposing, IMO - internetese for "in my opinion") is just $A = X^T X$ for some X. And $X^T X = V\Sigma^2 V^T$ by (2.12). SVD $\gg$ eigen-decomposition. QED

Not convinced? Need me to spell it out for u? Singular vectors are the eigenvectors of A, and singular values are the square roots of the eigenvalues!!! So, the SVD gives you the eigen-decomposition for free!!! Eigen-decomposition *just got owned by the* SVD. That's how it's done!!!!!

The SVD is a great 1-stop shop for data analysis.

Need to know if X is multi-collinear, before fitting least squares? Check out the singular values. If σ_1/σ_m is huge then least squares is a bad idea.

If $n > m$, or if $\sigma_m = 0$ then bad news bears, X isn't even invertible!! To know if the matrix $X^T X$ is invertible, you just have to check whether the smallest singular value is non-zero. Want to know the rank of X? It's just # of non-zero singular values!

Want the Moore-Penrose pseudo-inverse (though please be careful — there are better ways to approximate a matrix inverse)? That's basically the rank-k approx from earlier, but with $1/\sigma_k$ instead of σ_k!

And please don't troll me with your comments about how you prefer the QR or LU decompositions. I'm a working mom with 3 kids at home in the midst of a pandemic, I know you don't mean it, and I literally don't have time for this. (@SusCrockford, March 12, 2020 concurs: "The next academic dude who posts about how much work Isaac Newton or whoever got done at Cambridge during the plague I'm coming over to their house with my snotty 4y and staying there to develop my genius while he deals with the kid and then we'll see who discovers laws of nature.")

And if you prefer a specialty matrix decomposition, like NMF, then I've got news for you: you got fooled, because that's just a souped up SVD. Honest to god. If you remain convinced that the NMF or any other decomposition discovered in the past 80 years can hold a candle to the SVD, then I can get you a great price on a Tesla-branded vegan unicorn made out of CRISPR. I'll send it to you as soon as you give me all your Bitcoin.

The SVD is super magical and there's so much I've left unsaid. While you can compute it using a single line of code in R or any other halfway decent programming language, it's fun, easy, and safe to DIY with matrix multiplies!!!!

I hope that this thread has helped to grow your appreciation for this magical decomposition. The more you learn about the SVD, the more you will love it. It will take you far on your statistics and data science journey. Godspeed.

The rest is on YouTube.

2.4.1 Eigen values vs. singular values

> It's now more important to learn boring algebra than to practice fun rock throwing. So you take your choice. If you choose happiness over survival too consistently–well, then you die happy. Or else, you thrive grumpily. It's the tragedy of the human condition. About time we changed it, in my humble opinion.
>
> — Hans Moravec

Trefethen and Bau [4] *Numerical Linear Algebra*:

> $\cdots$ not all matrices (even square ones) have an eigenvalue decomposition [here week 1 and sect. 2.3], but all matrices (even rectangular ones) have a singular value decomposition [here sect. 2.4]. In applications, eigenvalues tend to be relevant to problems involving the behavior of iterated forms of A, such as matrix powers A^k or exponentials $\exp A$, whereas singular vectors tend to be relevant to problems involving the magnitudes of elements of A , or its inverse.

2.4.2 SVD in dynamical systems

Tosif Ahamed @_mlechha : There are lots of applications in recovering dynamics from data. I'd also like to plug our own eigen-worms, using SVD to go from experimental observations of moving worms to their periodic orbits (and more) arXiv:1911.10559.

On 2020-08-20 graduate student Daniel Dylewsky, U Washington, gave a good presentation *Koopman Approximations for Multiscale Nonlinear Physics using Dynamic Mode Decomposition*, based on Daniel Dylewsky, Eurika Kaiser, Steven L. Brunton and J. Nathan Kutz *Principal Component Trajectories (PCT): Nonlinear dynamics as a superposition of time-delayed periodic orbits* arXiv:2005.14321.

2.4.3 SVD in rocket science

Steven L. Brunton *et al. Data-Driven Aerospace Engineering: Reframing the Industry with Machine Learning* arXiv:2008.10740, a review: "$\cdots$ The aerospace industry is poised to capitalize on big data and machine learning, which excels at solving the types of multi-objective, constrained optimization problems that arise in aircraft design and manufacturing. Indeed, emerging methods in machine learning may be thought of as data-driven optimization techniques that are ideal for high-dimensional, non-convex, and constrained, multi-objective optimization problems, and that improve with increasing volumes of data."

2.4.4 SVD in theoretical neuroscience

Srdjan Ostojić again: During my physics education, I have never heard of singular value decomposition.

Almost all matrices in physics are symmetric, and in that case SVD reduces to eigenvalue decomposition.

But for non-symmetric, or especially non-square matrices, SVD is the fundamental tool. So over the recent years, part of the theoretical neuroscience community has been rediscovering how useful SVD is.

For instance, basic results on perceptrons can be understood in a simple way using SVD. Dynamics of learning in deep networks can be understood based on SVD: arXiv:1312.6120, on Pnas, arXiv:1809.10374. Non-linear dynamics in recurrent neural networks can be analyzed by starting from the SVD of the connectivity matrix, and keeping dominant terms: arXiv:1711.09672, bioRxiv:350801v3, arXiv:2007.02062, arXiv:1909.04358, bioRxiv:2020.07.03.185942v1. Non-normal transient dynamics in recurrent networks: on sciencedirect, arXiv:1811.07592.

References

[1] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. (Academic, New York, 2013).

[2] S. L. Brunton and J. N. Kutz, *Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* (Cambridge Univ. Press, Cambridge UK, 2019).

[3] A. K. Cline and I. S. Dhillon, "Computation of the singular value decomposition", in *Handbook of Linear Algebra* (CRC Press, 2006), pp. 45-1–45-13.

[4] L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (SIAM, 1997).

2.5 Examples

Example 2.1. ***Vibrations of a classical CO_2 molecule:*** *Consider one carbon and two oxygens constrained to the x-axis [1] and joined by springs of stiffness k, as shown*

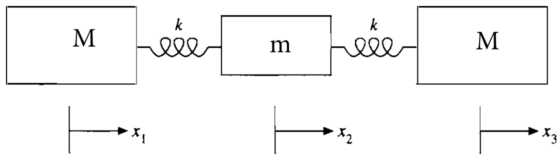


Figure 2.1: A classical colinear CO_2 molecule [1].

in figure 2.1. Newton's second law says

$$\begin{aligned}\ddot{x}_1 &= -\frac{k}{M}(x_1 - x_2) \\ \ddot{x}_2 &= -\frac{k}{m}(x_2 - x_3) - \frac{k}{m}(x_2 - x_1) \\ \ddot{x}_3 &= -\frac{k}{M}(x_3 - x_2)\,. \end{aligned} \tag{2.14}$$

The normal modes, with time dependence $x_j(t) = x_j \exp(it\omega)$, *are the common frequency* ω *vibrations that satisfy (2.14),*

$$\mathbf{H}\mathbf{x} = \begin{pmatrix} A & -A & 0 \\ -a & 2\,a & -a \\ 0 & -A & A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \omega^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{2.15}$$

where $a = k/m$, $A = k/M$. *Secular determinant* $\det(\mathbf{H} - \omega^2 \mathbf{1}) = 0$ *now yields a cubic equation for* ω^2.

You might be tempted to stick this $[3\times 3]$ *matrix into* `Mathematica` *or whatever, but please do that in some other course. What would understood by staring at the output? In this course we think.*

First thing to always *ask yourself is: does the system have a* symmetry*? Yes! Note that the* CO_2 *molecule (2.14) of figure 2.1 is invariant under* $x_1 \leftrightarrow x_3$ *interchange, i.e., coordinate relabeling by matrix* σ *that commutes with our law of motion* $\mathbf{H}$*,*

$$\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \sigma\mathbf{H} = \mathbf{H}\sigma = \begin{pmatrix} 0 & -A & A \\ -a & 2\,a & -a \\ A & -A & 0 \end{pmatrix}. \tag{2.16}$$

We can now use the symmetry operator σ *to simplify the calculation. As* $\sigma^2 = \mathbf{1}$*, its eigenvalues are* ± 1*, and the corresponding symmetrization, anti-symmetrization* projection operators *(1.30) are*

$$P_+ = \frac{1}{2}(\mathbf{1} + \sigma)\,, \qquad P_- = \frac{1}{2}(\mathbf{1} - \sigma)\,. \tag{2.17}$$

The dimensions $d_i = \mathrm{tr}\, P_i$ *of the two subspaces are*

$$d_+ = 2\,, \qquad d_- = 1\,. \tag{2.18}$$

As σ *and* $\mathbf{H}$ *commute, we can now use spectral decomposition (1.24) to block-diagonalize* $\mathbf{H}$ *to a 1-dimensional and a 2-dimensional matrix.*

On the 1-dimensional antisymmetric subspace, the trace of a $[1\times 1]$ *matrix equals its sole matrix element equals it eigenvalue*

$$\lambda_- = \mathbf{H}P_- = \frac{1}{2}(\mathrm{tr}\,\mathbf{H} - \mathrm{tr}\,\mathbf{H}\sigma) = (a + A) - a = \frac{k}{M}\,,$$

so the corresponding eigenfrequency is $\omega_-^2 = k/M$*. To understand its physical meaning, write out the antisymmetric subspace projection operator (2.18) explicitly. Its non-vanishing columns are proportional to the sole eigenvector*

$$P_- = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \Rightarrow \mathbf{e}^{(-)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \tag{2.19}$$

In this subspace the outer oxygens are moving in opposite directions, with the carbon stationary.

On the 2-dimensional symmetric subspace, the trace yields the sum of the remaining two eigenvalues

$$\lambda_+ + \lambda_0 = \mathrm{tr}\,\mathbf{H}P_+ = \frac{1}{2}(\mathrm{tr}\,\mathbf{H} + \mathrm{tr}\,\mathbf{H}\sigma) = (a + A) + a = \frac{k}{M} + 2\,\frac{k}{m}\,.$$

We could disentangle the two eigenfrequencies by evaluating $\mathrm{tr}\,\mathbf{H}^2 P_+$*, for example, but thinking helps again.*

There is still another, translational symmetry, so obvious that we forgot it; if we change the origin of the x*-axis, the three coordinates* $x_j \to x_j - \delta x$ *change, for any continuous translation* δx*, but the equations of motion (2.14) do not change their form,*

$$\mathbf{H}\,\mathbf{x} = \mathbf{H}\,\mathbf{x} + \mathbf{H}\,\delta\mathbf{x} = \omega^2\mathbf{x} \quad \Rightarrow \quad \mathbf{H}\,\delta\mathbf{x} = 0\,. \tag{2.20}$$

So any translation $\mathbf{e}^{(0)} = \delta\mathbf{x} = (\delta x, \delta x, \delta x)$ *is a nul, 'zero mode' eigenvector of* $\mathbf{H}$ *in (2.16), with eigenvalue* $\lambda_0 = \omega_0^2 = 0$*, and thus the remaining eigenfrequency is* $\omega_+^2 = k/M + 2\,k/m$*. As we can add any nul eigenvector* $\mathbf{e}^{(0)}$ *to the corresponding* $\mathbf{e}^{(+)}$ *eigenvector, there is some freedom in choosing* $\mathbf{e}^{(+)}$*. One visualization of the corresponding eigenvector is the carbon moving opposite to the two oxygens, with total momentum set to zero.*

(Taken from AWH Example 6.2.3 Degenerate eigenproblem*, but done here using symmetries.)*

Example 2.2. ***Vibrational spectra of molecules:*** *Consider the ring of pair-wise interactions of two kinds of molecules sketched in figure 2.2 (a), given by the potential*

$$V(z) = \frac{1}{2}\sum_{i=1}^{N}\left(k_1(x_i - y_i)^2 + k_2(x_{i+1} - y_i)^2\right), \qquad z_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \tag{2.21}$$

whose $[2N \times 2N]$ *matrix form is (aside to the cognoscenti: this is a Toeplitz matrix):*

$$V_{ij} = \frac{1}{2}\begin{pmatrix} k_1+k_2 & -k_1 & 0 & 0 & 0 & \dots & 0 & 0 & -k_2 \\ -k_1 & k_1+k_2 & -k_2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -k_2 & k_1+k_2 & -k_1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & -k_1 & k_1+k_2 & -k_2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -k_2 & k_1+k_2 & -k_1 \\ -k_2 & 0 & 0 & 0 & 0 & \dots & 0 & -k_1 & k_1+k_2 \end{pmatrix}$$

This potential matrix is a holy mess. How do we find an orthogonal transformation (2.3) that diagonalizes it? Look at figure 2.2 (a). Molecules lie on a circle, so that suggests we should use a Fourier representation. As the $i = 1$ *labelling of the starting molecule*

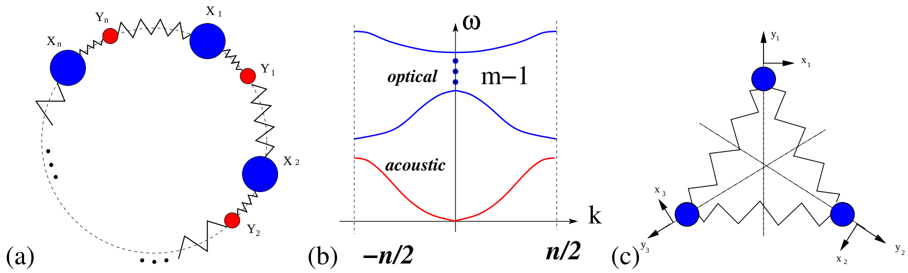


Figure 2.2: (a) Chain with circular symmetry. (b) Dependance of frequency on the representation wavenumber k. (c) Molecule with D_3 symmetry. (B. Gutkin)

on a ring is arbitrary, we are free to relabel them, for example use the next molecule pair as the starting one. This relabelling is accomplished by the $[2N\times 2N]$ *permutation matrix (or 'one-step shift', 'stepping' or 'translation' matrix)* M *of form*

$$\underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 & I \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{pmatrix}}_{M} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} z_n \\ z_1 \\ z_2 \\ \vdots \\ z_{n-1} \end{pmatrix}, \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad z_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{2.22}$$

Projection operators corresponding to M *are worked out in example 10.1. They are* N *distinct* $[2N\times 2N]$ *matrices,*

$$P_k = \begin{pmatrix} I & \bar\lambda I & \bar\lambda^2 I & \dots & \bar\lambda^{N-2} I & \bar\lambda^{N-1} I \\ \lambda I & I & \bar\lambda I & \dots & \bar\lambda^{N-3} I & \bar\lambda^{N-2} I \\ \lambda^2 I & \lambda I & I & \dots & \bar\lambda^{N-4} I & \bar\lambda^{N-3} I \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda^{N-2} I & \lambda^{N-3} I & \lambda^{N-4} I & \dots & I & \bar\lambda I \\ \lambda^{N-1} I & \lambda^{N-2} I & \lambda^{N-2} I & \dots & \lambda I & I \end{pmatrix}, \qquad \lambda = \exp\left(\frac{2\pi i}{N} k\right) \tag{2.23}$$

which decompose the $2N$*-dimensional configuration space of the molecule ring into a direct sum of* N *2-dimensional spaces, one for each discrete Fourier mode* $k = 0, 1, 2, \cdots, N-1$.

The system (2.21) is clearly invariant under the cyclic permutation relabelling M*,* $[V, M] = 0$ *(though checking this by explicit matrix multiplications might be a bit tedious), so the* P_k *decompose the interaction potential* V *as well, and reduce its action to the* k*th 2-dimensional subspace. Thus the* $[2N\times 2N]$ *diagonalization (2.3) is now reduced to a* $[2\times 2]$ *diagonalization which one can do by hand. The resulting* k*th space is spanned* *by two $2N$-dimensional vectors, which we guess to be of form:*

$$\eta_1 = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ 0 \\ \lambda \\ 0 \\ \vdots \\ \lambda^{n-1} \\ 0 \end{pmatrix} , \qquad \eta_2 = \frac{1}{\sqrt{n}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \lambda \\ \vdots \\ 0 \\ \lambda^{n-1} \end{pmatrix} .$$

In order to find eigenfrequences we have to consider action of V on these two vectors:

$$V\eta_1 = (k_1 + k_2)\eta_1 - (k_1 + k_2\lambda)\eta_2 \,, \qquad V\eta_2 = (k_1 + k_2)\eta_2 - (k_1 + k_2\bar{\lambda})\eta_1 \,.$$

The corresponding eigenfrequencies are determined by the equation:

$$\begin{aligned} 0 &= \det \left(\begin{pmatrix} k_1 + k_2 & -(k_1 + k_2\lambda) \\ -(k_1 + k_2\bar{\lambda}) & k_1 + k_2 \end{pmatrix} - \frac{\omega^2}{2} I \right) \qquad \Longrightarrow \\ \frac{1}{2}\omega^2_{\pm}(k) &= k_1 + k_2 \pm |k_1 + k_2\lambda^k| \,, \end{aligned} \tag{2.24}$$

one acoustic ($\omega(0) = 0$), one optical, see figure 2.2 (b) and the acoustic and optical phonons wiki. *(B. Gutkin)*

Example 2.3. ***An SVD hand calculation.*** *Given a rectangular $[n \times m] = [4 \times 3]$ "data matrix"*

$$X = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.25}$$

in this example we implement by hand its singular value decomposition

$$X = U\Sigma V^T \,. \tag{2.26}$$

A side remark: by inspection, the 1st and 3rd rows of X are not independent from the 2nd, the rank of the data matrix X is 2, so expect one zero eigenvalue.

The 'right', $[m \times m] = [3 \times 3]$ correlation matrix (see (2.7) and (2.8)) is

$$C_r = X^T X = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 2 \end{pmatrix} . \tag{2.27}$$

The zeroes of its characteristic polynomial

$$\det(C_r - \lambda 1) = (-8 + \lambda)(-2 + \lambda)\lambda = 0 \tag{2.28}$$

yield eigenvalues

$$\{\lambda_1, \lambda_2, \lambda_3\} = \{8, 2, 0\} \tag{2.29}$$

You are free to find the corresponding eigenvectors any way you like. If you use projection operators, you will also need the matrix squared:

$$C_r^2 = \begin{pmatrix} 12 & 20 & 12 \\ 20 & 44 & 20 \\ 12 & 20 & 12 \end{pmatrix}$$

The associated projection operators are:

$$
\begin{aligned}
P_1 &= \frac{(C_r - 2\cdot 1)(C_r - 0\cdot 1)}{(8-2)(8-0)} = \frac{1}{6\cdot 8}(C_r^2 - 2C_r) = \frac{1}{6}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \\
P_2 &= \frac{(C_r - 8\cdot 1)(C_r - 0\cdot 1)}{(2-8)(2-0)} = \frac{1}{6\cdot 2}(-C_r^2 + 8C_r) = \frac{1}{3}\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & - & 1 \end{pmatrix} \\
P_3 &= \frac{(C_r - 8\cdot 1)(C_r - 2\cdot 1)}{(0-8)(0-2)} = \frac{1}{8\cdot 2}(C_r^2 - 10C_r + (8\cdot 2)\,1) = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}
\end{aligned}
\tag{2.30}
$$

Each column of a projection operator is the same right eigenvector, with a different prefactor, and its rows are likewise proportional to the same left eigenvector. SVD, however, demands that the eigenvectors be normalized to unit length, for example

$$
v_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}
\tag{2.31}
$$

The three normalized right singular vectors, taken as the columns, form the rotation matrix

$$
V = \left(\; v_1 \;\middle|\; v_2 \;\middle|\; v_3 \; \right) = \left(\begin{array}{c|c|c} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & -\frac{2}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{array} \right)
\tag{2.32}
$$

The "left", "big" correlation matrix

$$
C_l = X\,X^T = \begin{pmatrix} 2 & 2\sqrt{2} & 0 & 0 \\ 2\sqrt{2} & 6 & 2 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\tag{2.33}
$$

characteristic polynomial has the same non-zero eigenvalues

$$
\det(C_l - \lambda 1) = (-8+\lambda)(-2+\lambda)\lambda^2 = 0\,,
\tag{2.34}
$$

but an extra zero eigenvalue. Going through the same algebra as for C_r, we find that C_l (unnormalized) eigenvectors can be presented as columns of matrix

$$
\hat{U} = \begin{pmatrix} \sqrt{2} & 3 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \sqrt{2} & -1 & 1 & 0 \end{pmatrix}
\tag{2.35}
$$

After normalization to unit length we refer to them as the left singular vectors.

The singular values are, by definition, the positive square roots of C_r or C_l eigenvalues

$$
\{\sigma_1, \sigma_2, \sigma_3\} = \{2\sqrt{2}, \sqrt{2}, 0\}
\tag{2.36}
$$

so the diagonal singular values matrix is given by

$$
\Sigma = \begin{pmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\tag{2.37}
$$

SVG decomposition (2.26) is of form $X = U\Sigma V^T$. Now that we have the right eigenvectors matrix V, and the diagonal singular values matrix Σ, we could compute the left eigenvectors matrix $U = XV/\Sigma$, as can be done in some of the examples in exercise 2.2. But zero singular values make this a bit tricky, so here we compute instead also U from the C_l eigenvalue equation, and verify that we indeed get the rectangular data matrix exercise 2.25 back:

$$U\Sigma V^T = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

A sanity check: Mathematica does all this in one line:
$\{U, \Sigma, V\} =$ ***SingularValueDecomposition***$[X] \qquad \Rightarrow$

$$U = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{\sqrt{3}}{2} & 0 & 0 & -\frac{1}{2} \\ \frac{1}{2\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.38}$$

$$V = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{2.39}$$

verifying (2.32), (2.37), and (2.35).

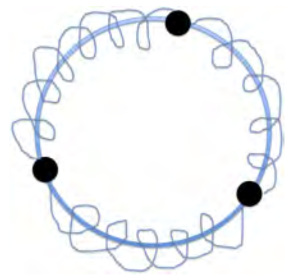


Figure 2.3: Three identical masses are constrained to move on a hoop, connected by three identical springs such that the system wraps completely around the hoop. Find the normal modes.

Exercises

2.1. **Three masses on a loop.** Three identical masses, connected by three identical springs, are constrained to move on a circle hoop as shown in figure 2.3. Find the normal modes. Hint: write down coupled harmonic oscillator equations, guess the form of oscillatory solutions. Then use basic matrix methods, i.e., find zeros of a characteristic determinant, find the eigenvectors, etc.. (Kimberly Y. Short)

2.2. **Examples of singular value decomposition.** Bring, by hand calculation, the following matrices into SVD form:

$$\begin{aligned} A &= \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}, \; B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\ D &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \; E = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \end{aligned} \tag{2.40}$$

The goal is to verify that any matrix, including these, has the unique SVD decomposition, and (2.8), (2.6) and (2.7) should suffice for the job.

Chapter 2 solutions: Eigenvalue problems

Solution 2.1 - Three masses on a loop. *As the masses and springs are identical, the equilibrium positions, x_1, x_2, and x_3, of the masses are equally spaced on the hoop, i.e., separated by 120° or $2\pi/3$ rads. The equations of motion are*

$$\begin{aligned} m\ddot{x}_1 + k(x_1 - x_2) + k(x_1 - x_3) &= 0 \\ m\ddot{x}_2 + k(x_2 - x_3) + k(x_2 - x_1) &= 0 \\ m\ddot{x}_3 + k(x_3 - x_1) + k(x_3 - x_2) &= 0 \end{aligned} \tag{2.41}$$

Anticipating oscillatory solutions, we introduce trial solutions of the form

$$x_1 = A_1 e^{i\alpha t}, \quad x_2 = A_2 e^{i\alpha t}, \quad x_3 = A_3 e^{i\alpha t}$$

Plugging these expressions into our system of equations yields

$$\begin{pmatrix} -\alpha^2 + 2\omega^2 & -\omega^2 & -\omega^2 \\ -\omega^2 & -\alpha^2 + 2\omega^2 & -\omega^2 \\ -\omega^2 & -\omega^2 & -\alpha^2 + 2\omega^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.42}$$

Taking the determinant yields a cubic equation for α^2:

$$-(\alpha^2)^3 + 6(\alpha^2)^2\omega^2 - 9(\alpha^2)\omega^4 = 0 \tag{2.43}$$

The three solutions are

$$\alpha^2 = 0, \qquad \alpha^2 = 3\omega^2(\text{multiplicity} = 2)$$

Consider first the solution corresponding to $\alpha^2 = 0$ with eigenvector $(1, 1, 1)^\top/\sqrt{3}$. In this case, the normal mode is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (At + B) \tag{2.44}$$

where have absorbed the $1/\sqrt{3}$ factor into the coefficients A and B. This normal mode has zero frequency and corresponds to the masses sliding around the hoop, equally spaced, at constant speed. The normal coordinates are simply $x_1 + x_2 + x_3$.

The remaining two roots corresponding to $\alpha^2 = 3\omega^2$ describe oscillations. Any vector of the form (a, b, c) satisfying $a + b + c = 0$ is a valid normal mode with frequency $\sqrt{3}\omega$. Here we have chosen vectors $(0, 1, -1)^\top/\sqrt{2}$ and $(1, 0, -1)^\top/\sqrt{2}$ as the basis for the two-dimensional subspace of normal modes. Consequently, the normal modes may be written as linear combinations of the vectors

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = C_1 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \cos(\sqrt{3}\omega t + \phi_1) \tag{2.45}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = C_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \cos(\sqrt{3}\omega t + \phi_2) \tag{2.46}$$

with normal coordinates $x_1 - 2x_2 + x_3$ and $-2x_1 + x_2 + x_3$, respectively.

(Kimberly Y. Short)

Solution 2.2 - Examples of singular value decomposition. *The five matrices in SVD form (LaTex generated by Mathematica `SingularValueDecomposition[A]`):*

$$
\begin{aligned}
A &= \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \left(\begin{array}{cc} 3 & 0 \\ 0 & 2 \end{array} \right) \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \\
B &= \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) \left(\begin{array}{cc} 3 & 0 \\ 0 & 2 \end{array} \right) \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) \\
C &= \begin{pmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right) \left(\begin{array}{cc} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{array} \right) \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) \\
D &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \left(\begin{array}{cc} \sqrt{2} & 0 \\ 0 & 0 \end{array} \right) \left(\begin{array}{cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array} \right) \\
E &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \left(\begin{array}{cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array} \right) \left(\begin{array}{cc} 2 & 0 \\ 0 & 0 \end{array} \right) \left(\begin{array}{cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array} \right) .
\end{aligned}
\tag{2.47}
$$

Matrix C is rank 1, so only the first u_1 column matters, u_2 and u_3 are any 2 unit length singular vectors that span the remaining 2 dimensions. The same for rank 1 matrices D and E. The "data" has nothing to say about singular value $\sigma_j = 0$ singular vectors.

mathematical methods - week 3

Go with the flow

Georgia Tech PHYS-6124

Homework HW #3 due Tuesday, September 8, 2020

== 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 exerWeek3.tex

Exercise 3.1 *Rotations in a plane*	4 points
Exercise 3.2 *Visualizing 2-dimensional linear flows*	6 points

Bonus points

Exercise 3.3 *Visualizing Duffing flow*	3 points
Exercise 3.4 *Visualizing Lorenz flow*	2 points
Exercise 3.5 *A limit cycle with analytic Floquet exponent*	6 points

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

edited September 13, 2020

Week 3 syllabus

Tuesday, September 1, 2020

Typical ordinary differential equations course spends most of time teaching you how to solve linear equations, and for those our spectral decompositions are very instructive. Nonlinear differential equations (as well as the differential geometry) are much harder, but still (as we already discussed in sect. 1.3), linearizations of flows are a very powerful tool.

This week's lectures are related to AWH Chapter 7 *Ordinary Differential Equations* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- *Go with the linear flow : full Tuesday lecture*
  - Sect. 3.2 *Linear flows*
  - AWH Section 7.2 *First-order equations*
  - Sect. 3.3 *Stability of linear flows*
  - *Local stability: Stability matrix (already covered in the above full lecture*
- *Go nonlinear : full Thursday lecture* (includes the next 7 videos, but in BlueJeans mangled resolution; still the lecture is more than the sum of the 7 clips)
  - *Lorenz flow*
  - *Strange attractors*
  - *Strange attractors - Lorenz again*
  - *Lorenz again (apologies)*
  - *Roessler flow*
  - *Computing is like hygiene, personal*
  - *Dynamical systems : a summary*
- *Mixed phase space; Jacques Laskar rant*
- *Computing hygiene; the obligatory Gibson rant, take #2*

3.1 Other sources

As every week, feel free to ignore extra reading and videos for this week. What I cover in the online lecture is all that I hope you take home with you.

- Just as you had learned everything about linear ODEs, this tweet comes along :(

- *Dynamical systems*
  GaTech College of Unprofessional Education insisted on making me a talking head in a GaTech branded video. I hated it. I fired them. They fired me. I do not even know who "they" were, but I got to teach the rest of the course on a blackboard. Until COVID-19 that reduced us all to talking heads.

- *Trajectories*
- *Equilibria*
- *Orbits are time-invariant*
  I'm so happy. I'm divorced of Unprofessional Education, and free of their moronic PowerPoints! But if you must,
- *Do the course in the Power Point format*
- *Long live Bologna!*
- *John F Gibson solves the Navier-Stokes, take #1*
- *Life in extreme dimensions: Fluttering flame front*
- *Life in extreme dimensions: Constructing state spaces*
- *Life in extreme dimensions: As visualized by dummies*
- *What do these equations do?*

- MIT 16-90 Computational methods is a typical mathematical methods in engineering course. ODEs are discussed here.
- There are no doubt many online courses vastly better presented than this one - here is a glimpse into our competition:
  MIT 18.085 Computational Science and Engineering I .
- Optional reading: Sect. 3.4 *Nonlinear flows*
- Optional reading: AWH Section 7.8 *Nonlinear differential equations*
- Sect. 3.5 *Optional listening*

3.2 Linear flows

Linear is good, nonlinear is bad.
—Jean Bellissard

(Notes based on ChaosBook *Chapter 2 : Go with the flow*).

A *dynamical system* is defined by specifying a state space $\mathcal{M}$, and a law of motion, typically an ordinary differential equation (ODE), first order in time,

$$\dot{x} = v(x)\,. \tag{3.1}$$

The vector field $v(x)$ can be any nonlinear function of x, so it pays to start with a simple example. *Linear* dynamical system is the simplest example, 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 state space point x and a constant A as

$$\dot{x} = A\,x\,. \tag{3.2}$$

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.

Solution of (3.2) is given by the exponential of a constant matrix

$$x(t) \;=\; J^t\, x_0\,, \tag{3.3}$$

usually defined by its series expansion (1.10)

$$J^t = e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k\,, \qquad A^0 = \mathbf{1}\,, \tag{3.4}$$

and that is why we started the course by defining functions of matrices, and in particular the matrix exponential. As we discuss next, that means that depending on the eigenvalues of the matrix A, solutions of linear ordinary differential equations are either growing or shrinking exponentially (over-damped oscillators; $\cosh$'s, $\sinh$'s), or oscillating (under-damped oscillators; $\cos$'s, $\sin$'s).

3.3 Stability of linear flows

The system of linear *equations of variations* for the displacement of the infinitesimally close neighbor $x + \delta x$ follows from the flow equations (3.2) by Taylor expanding to linear order

$$\dot{x}_i + \dot{\delta x}_i = v_i(x + \delta x) \approx v_i(x) + \sum_j \frac{\partial v_i}{\partial x_j} \delta x_j\,.$$

The infinitesimal deviation vector δx is thus transported along the trajectory $x(x_0, t)$, with time variation given by

$$\frac{d}{dt}\delta x_i(x_0, t) = \sum_j \frac{\partial v_i}{\partial x_j}(x)\bigg|_{x=x(x_0,t)} \delta x_j(x_0, t)\,. \tag{3.5}$$

As both the displacement and the trajectory depend on the initial point x_0 and the time t, we shall often abbreviate the notation to $x(x_0, t) \to x(t) \to x$, $\delta x_i(x_0, t) \to \delta x_i(t) \to \delta x$ in what follows. Taken together, the set of equations

$$\dot{x}_i = v_i(x)\,, \quad \dot{\delta x}_i = \sum_j A_{ij}(x)\delta x_j \tag{3.6}$$

governs the dynamics in the tangent bundle $(x, \delta x) \in \mathbf{T}\mathcal{M}$ obtained by adjoining the d-dimensional tangent space $\delta x \in T\mathcal{M}_x$ to every point $x \in \mathcal{M}$ in the d-dimensional state space $\mathcal{M} \subset \mathbb{R}^d$. The *stability matrix* or *velocity gradients matrix*

$$A_{ij}(x) = \frac{\partial}{\partial x_j} v_i(x) \tag{3.7}$$

describes the instantaneous rate of shearing of the infinitesimal neighborhood of $x(t)$ by the flow. In case at hand, the linear flow (3.2), with $v(x) = A\,x$, the stability matrix

$$A_{ij}(x) = \frac{\partial}{\partial x_j} v_i(x) = A_{ij} \tag{3.8}$$

is a space- and time-independent constant matrix.

Consider an infinitesimal perturbation of the initial state, $x_0 + \delta x$. The perturbation $\delta x(x_0, t)$ evolves as $x(t)$ itself, so

$$\delta x(t) \;=\; J^t\,\delta x(0)\,. \tag{3.9}$$

The equations are linear, so we can integrate them. In general, the Jacobian matrix J^t is computed by integrating the *equations of variations*

$$\dot{x}_i = v_i(x)\,, \quad \dot{\delta x}_i = \sum_j A_{ij}(x)\delta x_j\,, \tag{3.10}$$

but for linear ODEs everything is known once eigenvalues and eigenvectors of A are known.

Example 3.1. ***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 3.2.*

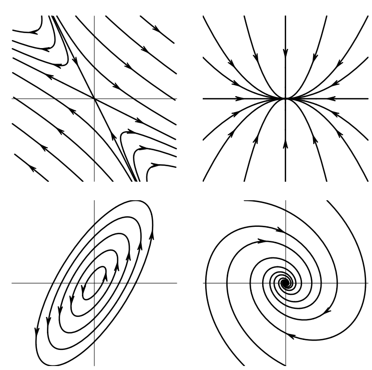


Figure 3.1: Streamlines for several typical 2-dimensional flows: saddle (hyperbolic), in node (attracting), center (elliptic), in spiral.

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 3.1, as are*

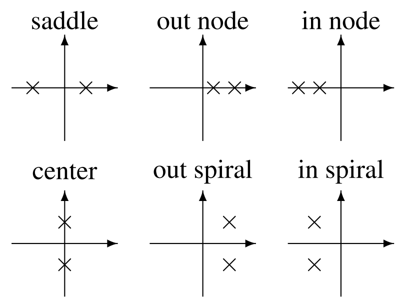


Figure 3.2: Qualitatively distinct types of exponents $\{\lambda_1, \lambda_2\}$ of a $[2\times 2]$ Jacobian matrix.

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 1.1. We distinguish two cases: (a) A can be brought to diagonal form and (b) A can be brought to Jordan 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 3.1 and 3.2.

Example 3.2. *Complex eigenvalues: in-out spirals.* *As $\mathbf{M}$ 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 $\mathbf{M}' \subset \mathbf{M}$ 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, $\{\mathbf{e}^{(k)}, \mathbf{e}^{(k+1)}\} \to \{\mathrm{Re}\,\mathbf{e}^{(k)}, \mathrm{Im}\,\mathbf{e}^{(k)}\}$. In this 2-dimensional real representation, $\mathbf{M}' \to A$, the block A is a sum of the rescaling$\times$identity and the generator of rotations in the $\{\mathrm{Re}\,\mathbf{e}^{(1)}, \mathrm{Im}\,\mathbf{e}^{(1)}\}$ plane.

$$A = \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix} = \mu \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \omega \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} . \tag{3.11}$$

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

$$J^t = e^{tA} = e^{t\mu} \begin{bmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{bmatrix} , \tag{3.12}$$

spiral in/out around $(x, y) = (0, 0)$, see figure 3.1, with the rotation period T and the radial expansion /contraction multiplier along the $\mathbf{e}^{(j)}$ eigen-direction per a turn of the spiral:

exercise 3.1

$$T = 2\pi/\omega\,, \qquad \Lambda_{radial} = e^{T\mu}\,. \tag{3.13}$$

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, $1000\,T$, or $10^{-2}T$).

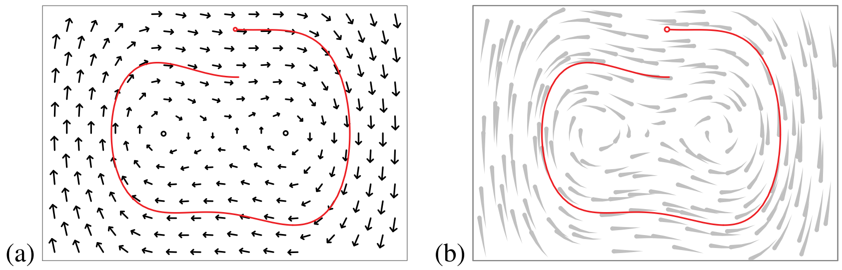


Figure 3.3: (a) The 2-dimensional vector field for the Duffing system (3.14), together with a short trajectory segment. (b) The flow lines. Each 'comet' represents the same time interval of a trajectory, starting at the tail and ending at the head. The longer the comet, the faster the flow in that region. (From ChaosBook [1])

3.4 Nonlinear flows

While linear flows are prettily analyzed in terms of defining matrices and their eigenmodes, understanding nonlinear flows requires many tricks and insights. These days, we start by integrating them, by any numerical code you feel comfortable with: Matlab, Python, Mathematica, Julia, c++, whatever.

Duffing flow of example 3.3 is a typical 2-dimensional flow, with a 'nonlinear oscialltor' limit cycle. Real fun only starts in 3 dimensions, with example 3.4 *Lorenz strange attractor*.

For purposes of this course, it would be good if you coded the next two examples, and just played with their visualizations, without further analysis (that would take us into altogether different ChaosBook.org/course1).

Example 3.3. ***A 2-dimensional vector field*** $v(x)$***.*** *A simple example of a flow is afforded by the unforced Duffing system*

$$\begin{aligned} \dot{x}(t) &= y(t) \\ \dot{y}(t) &= -0.15\, y(t) + x(t) - x(t)^3 \end{aligned} \tag{3.14}$$

plotted in figure 3.3. The 2-dimensional velocity vectors $v(x) = (\dot{x}, \dot{y})$ *are drawn superimposed over the configuration coordinates* (x, y) *of state space* $\mathcal{M}$*.*

Example 3.4. ***Lorenz strange attractor.*** *Lorenz equation*

$$\dot{x} = v(x) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \sigma(y-x) \\ \rho x - y - xz \\ xy - bz \end{bmatrix} \tag{3.15}$$

has played a key role in the history of 'deterministic chaos' for many reasons that you can read about elsewhere [1]. All computations that follow will be performed for the Lorenz parameter choice $\sigma = 10, b = 8/3, \rho = 28$*. For these parameter values the long-time dynamics is confined to the strange attractor depicted in figure 3.4.*

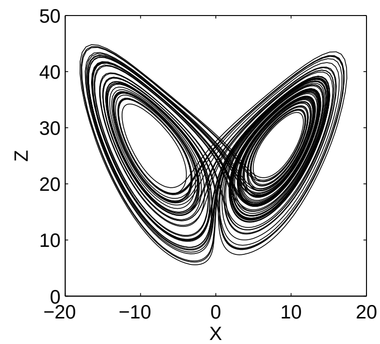


Figure 3.4: Lorenz “butterfly” strange attractor. (From ChaosBook [1])

3.5 Optional listening

If you do not know Emmy Noether, one of the great mathematicians of the 20th century, the time to make up for that is now. All symmetries we will use in this course are for kindergartners: flips, slides and turns. Noether, however, found a profound connections between these and invariants of our world - masses, charges, elementary particles. Then the powerful plutocrats of Germany made a clown the Chancellor of German Reich, because they could easily control him. They were wrong, and that’s why you are not getting this lecture in German. Noether lost interest in physics and went on to shape much of what is today called pure mathematics.

References

[1] R. Mainieri, P. Cvitanović, and E. A. Spiegel, “Go with the flow”, in *Chaos: Classical and Quantum*, edited by P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay (Niels Bohr Inst., Copenhagen, 2020).

Exercises

3.1. **Rotations in a plane:** In order to understand the role complex eigenvalues in example 3.2 play, it is helpful to show by exponentiation $J^t = \exp(tA) = \sum_{k=0}^{\infty} t^k A^k/k!$ with pure imaginary A in (3.11), that

$$A = \omega \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

generates a rotation in the $\{\mathrm{Re}\,\mathbf{e}^{(1)}, \mathrm{Im}\,\mathbf{e}^{(1)}\}$ plane,

$$\begin{aligned} J^t &= e^{At} = \cos\omega t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin\omega t \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix}. \end{aligned} \tag{3.16}$$

3.2. **Visualizing 2-dimensional linear flows.** Either sketch by hand, or use any integration routine to integrate numerically and plot, or plot the analytic solution of the linear flow (3.2) for all examples of qualitatively different eigenvalue pairs of figure 3.2. As noted in (1.29), the eigenvalues

$$\lambda_{1,2} = \frac{1}{2}\mathrm{tr}\,A \pm \frac{1}{2}\sqrt{(\mathrm{tr}\,A)^2 - 4\det A}$$

depend only on $\mathrm{tr}\,A$ and $\det A$, so you can get two examples by choosing any A such that $\mathrm{tr}\,A = 0$ (symplectic or Hamiltonian flow), vary $\det A$. For other examples choose A such that $\det A = 1$, vary $\mathrm{tr}\,A$. Do your plots capture the qualitative features of the examples of figure 3.1?

3.3. **Visualizing Duffing flow.** Use any integration routine to integrate numerically the Duffing flow (3.14). Take a grid of initial points, integrate each for some short time δt. Does your result look like the vector field of figure 3.3? What does a generic long-time trajectory look like?

3.4. **Visualizing Lorenz flow.** Use any integration routine to integrate numerically the Lorenz flow (3.15). Does your result look like the 'strange attractor' of figure 3.4?

3.5. **A limit cycle with analytic Floquet exponent.** There are only two examples of nonlinear flows for which the Floquet multipliers can be evaluated analytically. Both are cheats. One example is the 2-dimensional flow

$$\begin{aligned} \dot{q} &= p + q(1 - q^2 - p^2) \\ \dot{p} &= -q + p(1 - q^2 - p^2). \end{aligned}$$

Determine all periodic solutions of this flow, and determine analytically their Floquet exponents. Hint: go to polar coordinates $(q, p) = (r\cos\theta, r\sin\theta)$. G. Bard Ermentrout

Chapter 3 solutions: Go with the flow

Solution 3.1 - $\mathrm{SO}(2)$ rotations in a plane. *To compute $g(\theta) = \exp(\theta T)$, expand it in a Taylor series, noting that T is a real 2-dimensional representation of the imaginary unit i, so the powers of Lie algebra element T satisfy*

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad T^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbf{1},$$

so the terms in the expansion simplify to either T or $\mathbf{1}$: $T^3 = T^2 T = -\mathbf{1}T = -T$, $T^4 = (T)^2 = \mathbf{1}$, $T^5 = T^4 T = T$, etc. Hence,

$$\begin{aligned} e^{\theta T} &= \mathbf{1} + \theta T + \frac{1}{2!}(\theta T)^2 + \frac{1}{3!}(\theta T)^3 + \frac{1}{4!}(\theta T)^4 + \frac{1}{5!}(\theta T)^5 + \ldots \\ &= \mathbf{1} + \theta T - \frac{1}{2!}\theta^2 \mathbf{1} - \frac{1}{3!}\theta^3 T + \frac{1}{4!}\theta^4 \mathbf{1} + \frac{1}{5!}\theta^5 T + \ldots \\ &= \left(\mathbf{1} - \frac{1}{2!}\theta^2 \mathbf{1} + \frac{1}{4!}\theta^4 \mathbf{1} - \ldots\right) + \left(\theta T - \frac{1}{3!}\theta^3 T + \frac{1}{5!}\theta^5 T - \ldots\right) \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots\right)\mathbf{1} + \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots\right) T \\ &= \cos\theta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin\theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{aligned}$$

is a 2-dimensional rotation by angle θ.

Solution 3.4 - Visualizing Lorenz flow. *You will probably want the matlab function* `ode45` *to do this. There are several others which perform better in different situations (for example* `ode23` *for stiff ODEs), but* `ode45` *seems to be the best for general use.*

To use `ode45` *you must create a function, say '*`Lorenz`*', which will take in a time and a vector of* `[x,y,z]` *and return* `[xdot, ydot, zdot]`*. Then the command would be something like*

```
ode45([tmin, tmax], [x0 y0 z0], @Lorenz)
```

(Jonathan Halcrow)

Solution 3.5 - A limit cycle with analytic Floquet exponent. *The 2-dimensional flow is cooked up so that $x(t) = (q(t), p(t))$ is separable (check!) in polar coordinates $q = r\cos\phi$, $\quad p = r\sin\phi$:*

$$\dot{r} = r(1 - r^2), \qquad \dot{\phi} = 1. \tag{3.17}$$

In the (r, ϕ) coordinates the flow starting at any $r > 0$ is attracted to the $r = 1$ limit cycle, with the angular coordinate ϕ wrapping around with a constant angular velocity $\omega = 1$. The non–wandering set of this flow consists of the $r = 0$ equilibrium and the $r = 1$ limit cycle.

Equilibrium stability: *As the change of coordinates is defined everywhere except at the equilibrium point $(r = 0,$ any $\phi)$, the equilibrium stability matrix has to be computed in the original (q, p) coordinates,*

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \tag{3.18}$$

The eigenvalues are $\lambda = \mu \pm i\omega = 1 \pm i$, indicating that the origin is linearly unstable, with nearby trajectories spiralling out with the constant angular velocity $\omega = 1$. The Poincaré section ($p = 0$, for example) return map is in this case also a stroboscopic

map, strobed at the period (Poincaré section return time) $T = 2\pi/\omega = 2\pi$. *The radial Floquet multiplier per one Poincaré return is* $|\Lambda| = e^{\mu T} = e^{2\pi}$.
Limit cycle stability: *From (3.17) the stability matrix is diagonal in the* (r, ϕ) *coordinates,*

$$A = \begin{bmatrix} 1 - 3r^2 & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.19}$$

The vanishing of the angular $\lambda^{(\theta)} = 0$ *eigenvalue is due to the rotational invariance of the equations of motion along* ϕ *direction. The expanding* $\lambda^{(r)} = 1$ *radial eigenvalue of the equilibrium* $r = 0$ *confirms the above equilibrium stability calculation. The contracting* $\lambda^{(r)} = -2$ *eigenvalue at* $r = 1$ *decreases the radial deviations from* $r = 1$ *with the radial Floquet multiplier* $\Lambda_r = e^{\mu T} = e^{-4\pi}$ *per one Poincaré return. This limit cycle is* very *attracting.*
Stability of a trajectory segment: *Multiply (3.17) by* r *to obtain* $\frac{1}{2}\dot{r^2} = r^2 - r^4$, *set* $r^2 = 1/u$, *separate variables* $du/(1-u) = 2\,dt$, *and integrate:* $\ln(1-u) - \ln(1-u_0) = -2t$. *Hence the* $r(r_0, t)$ *trajectory is*

$$r(t)^{-2} = 1 + (r_0^{-2} - 1)e^{-2t}\,. \tag{3.20}$$

The $[1\times1]$ *Jacobian matrix*

$$J(r_0, t) = \left.\frac{\partial r(t)}{\partial r_0}\right|_{r_0 = r(0)}\,. \tag{3.21}$$

satisfies

$$\frac{d}{dt}J(r,t) = A(r)\,J(r,t) = (1 - 3r(t)^2)\,J(r,t)\,, \qquad J(r_0, 0) = 1\,.$$

This too can be solved by separating variables $d(\ln J(r,t)) = dt - 3r(t)^2 dt$, *substituting (3.20) and integrating. The stability of any finite trajectory segment is:*

$$J(r_0, t) = (r_0^2 + (1 - r_0^2)e^{-2t})^{-3/2}e^{-2t}\,. \tag{3.22}$$

On the $r = 1$ *limit cycle this agrees with the limit cycle multiplier* $\Lambda_r(1, t) = e^{-2t}$, *and with the radial part of the equilibrium instability* $\Lambda_r(r_0, t) = e^t$ *for* $r_0 \ll 1$.
(P. Cvitanović)

Solution 3.2 - Visualizing 2-dimensional linear flows.

Solution 3.3 - Visualizing Duffing flow.

Solution 3.4 - Visualizing Lorenz flow. *My code is ugly, and I see that many of you have code which is much clearer and prettier than mine... But just in case you don't, this is my code for these three problems.* *(Han Liang)*

I found the easiest way to visualize the velocity field of the flows is using the function "StreamPlot" in Mathematica. So for 3.2:

```
A = {{1, -1}, {1, 1}};
StreamPlot[A.{x, y}, {x, -10, 10}, {y, -10, 10}]
```

And for problem 3.3:

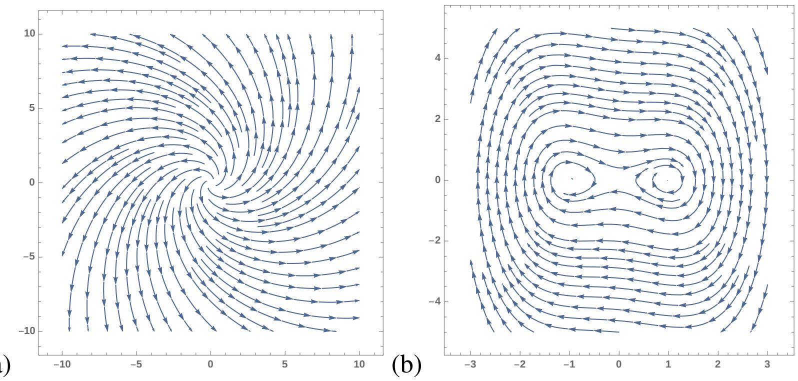


Figure 3.5: (a) A 2-dimensional vector field for the linear flows. (b) The 2-dimensional vector field for the Duffing flow

```
StreamPlot[{y, -0.15 y + x - x^3}, {x, -3, 3}, {y, -5, 5}]
```

The plots figure 3.5 have the desired feature.

For problem 3.4, to integrate the Lorenz flow, my code in MATLAB is:

```
clear
x(1)=1;
y(1)=1;
z(1)=1;
for t = 0:0.01:50
   dx=10*(y(end)-x(end));
   dy=28*x(end)-y(end)-x(end)*z(end);
   dz=x(end)*y(end)-8/3*z(end);
   x(end+1)=x(end)+0.01*dx;
   y(end+1)=y(end)+0.01*dy;
   z(end+1)=z(end)+0.01*dz;
end
plot3(x,y,z)
```

The result I get is figure 3.6.

(Han Liang)

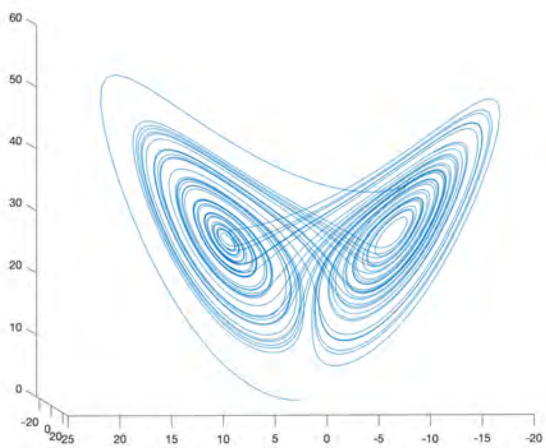


Figure 3.6: Lorenz strange attractor.

mathematical methods - week 4

Complex differentiation

Georgia Tech PHYS-6124

Homework HW #4 due Thursday, September 17, 2020

== 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 exerWeek4.tex

Exercise 4.2 *Complex arithmetic* 10 (+3 bonus) points
Exercise 4.5 *Circles and lines with complex numbers* 3 points

Bonus points
Exercise 4.1 *Complex arithmetic – principles* 6 points

Total of 13 points = 100 % score. Extra points accumulate, can help you later if you miss a few problems.

edited September 13, 2020

Week 4 syllabus

Tuesday, September 8, 2020

This week's lectures are related to Arfken, Weber & Harris [2] Chapter 11 *Complex variable theory* (click here). For the 3 next weeks of lectures on complex analysis, I am very much indebted to Paul Goldbart and his 1993 math methods lecture notes. The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

Complex variables; History; algebraic and geometric insights; De Moivre's formula; roots of unity; functions of complex variables as mappings (32 min; the last 7 minutes were not in the live lecture)

AWH 11.1 *Complex variables and functions*

Cauchy-Riemann - differentiation of complex functions; Cauchy-Riemann conditions; holomorphic (analytic) functions; conformal mappings (1h 9min, includes a conformal mapping clip that was not in the live lecture)

AWH 11.2 *Cauchy-Riemann conditions*

Everything is allowed in love and war (how to do problem sets)

Optional reading

- Grigoriev notes pages 2.1 - 2.3 (clean and concise)
- I personally am a big fan of Stone and Goldbart [4] (click here); our lectures on complex numbers follow Paul Goldbart's lectures.

  SG 17.1 *Cauchy-Riemann equations*

  SG 17.1.2 *Conformal mapping*

  SG 17.3.3 *Blasius and Kutta-Joukowski theorems* (for the rocket scientists among us)

  SG 17.6.1 *The point at infinity* (Riemann sphere)

- Ahlfors [1] (click here)
- Needham [3] (click here)

Alex Kontorovich, Rutgers MAT 640:503 Complex Analysis. A wonderful lecturer, here he diverges into the story of Cardano and cubics. They are *cube*-ic for a reason. Did you know people learned to use $\sqrt{-1}$ before they understood that a number can be negative, like -1? Listen to his first lecture. Oh no! He just made me solve the cubic, something I had avoided my entire life. So far. You'll love it.

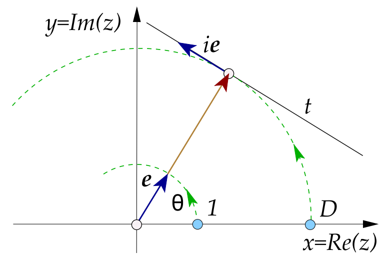


Figure 4.1: A unit vector $\mathbf{e}$ multiplied by a real number D traces out a circle of points in the complex plane. Multiplication by the imaginary unit i rotates a complex vector by 90^0, so $D\mathbf{e} + it\mathbf{e}$ is a tangent to this circle, a line parametrized by a real number t.

Question 4.1. Henriette Roux asks

Q You made us do exercise 4.5, but you did not cover this in class? I left it blank!

A Mhm. I told you that complex numbers can be understood as vectors in the complex plane, vectors that can be added and multiplied by scalars. I told you that the multiplication by the imaginary unit i rotates a complex vector by 90^0. I told you that in the polar representation, complex numbers define circle parametrized by their argument (phase). For example, a line is defined by its orientation $\mathbf{e}$, and its shortest distance to the origin is along the vector $D\mathbf{e}$, of length D, see figure 4.1.

The point of the exercise is that if you use your high school $\sin$'s and $\cos$'s, this simple formula (and the other that have to do with circles) is a mess.

References

[1] L. V. Ahlfors, *Complex Analysis*, 3rd ed. (Mc Graw Hill, 1979).

[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] T. Needham, *Visual Complex Analysis* (Oxford Univ. Press, Oxford UK, 1997).

[4] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

Exercises

4.1. **Complex arithmetic – principles:** (Ahlfors [1], pp. 1-3, 6-8)

(a) (bonus) Show that $\frac{A+iB}{C+iD}$ is a complex number provided that $C^2 + D^2 \neq 0$. Show that an efficient way to compute a quotient is to multiply numerator and denominator by the conjugate of the denominator. Apply this scheme to compute the quotient $\frac{A+iB}{C+iD}$.

(b) (bonus) By considering the equation $(x + iy)^2 = (A + iB)$ for real x, y, A and B, compute the square root of $A + iB$ explicitly for the case $B \neq 0$. Repeat for the case $B = 0$. (To avoid confusion it is useful to adopt he convention that square roots of positive numbers have real signs.) Observe that the square root of any complex number exists and has two (in general complex) opposite values.

(c) (bonus) Show that $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ and that $\overline{z_1\, z_2} = \bar{z}_1\, \bar{z}_2$. Hence show that $\overline{z_1/z_2} = \bar{z}_1/\bar{z}_2$. Note the more general result that for any rational operation R applied to the set of complex numbers $z_1, z_2, \ldots$ we have $\overline{R(z_1, z_2, \ldots)} = R(\bar{z}_1, \bar{z}_2, \ldots)$. Hence, show that if ζ solves $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 = 0$ then $\bar{\zeta}$ solves $\bar{a}_n z^n + \bar{a}_{n-1} z^{n-1} + \cdots + \bar{a}_0 = 0$.

(d) (bonus) Show that $|z_1\, z_2| = |z_1|\, |z_2|$. Note that this extends to arbitrary finite products $|z_1\, z_2\, \ldots| = |z_1|\, |z_2|\, \ldots$. Hence show that $|z_1/z_2| = |z_1|/|z_2|$. Show that $|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2\mathrm{Re}\, z_1\, \bar{z}_2$ and that $|z_1 - z_2|^2 = |z_1|^2 + |z_2|^2 - 2\mathrm{Re}\, z_1\, \bar{z}_2$.

4.2. **Complex arithmetic.** (Ahlfors [1], pp. 2-4, 6, 8, 9, 11)

(a) Find the values of

$$\big(1 + 2i\big)^3, \qquad \frac{5}{-3 + 4i}, \qquad \left(\frac{2 + i}{3 - 2i}\right),$$

$$\big(1 + i\big)^N + \big(1 - i\big)^N \quad \text{for} \quad N = 1, 2, 3, \ldots.$$

(b) If $z = x + iy$ (with x and y real), find the real and imaginary parts of

$$z^4, \qquad \frac{1}{z}, \qquad \frac{z - 1}{z + 1}, \qquad \frac{1}{z^2}.$$

(c) Show that, for all combinations of signs,

$$\left(\frac{-1 \pm i\sqrt{3}}{2}\right)^3 = 1, \qquad \left(\frac{\pm 1 \pm i\sqrt{3}}{2}\right)^6 = 1.$$

(d) By using their Cartesian representations, compute $\sqrt{i}$, $\sqrt{-i}$, $\sqrt{1 + i}$ and $\sqrt{\frac{1 - i\sqrt{3}}{2}}$.

(e) By using the Cartesian representation, find the four values of $\sqrt[4]{-1}$.

(f) By using their Cartesian representations, compute $\sqrt[4]{i}$ and $\sqrt[4]{-i}$.

(g) Solve the following quadratic equation (with real A, B, C and D) for complex z:

$$z^2 + (A + iB)z + C + iD = 0.$$

(h) Show that the system of all matrices of the form

$$\begin{bmatrix} A & B \\ -B & A \end{bmatrix}$$

(with real A and B), when combined by matrix addition and matrix multiplication, is isomorphic to the field of complex numbers.

(i) Verify by calculation that the values of $z/\big(z^2+1\big)$ for $z = x+iy$ and $z = x-iy$ are conjugate.

(j) Find the absolute values of

$$-2i\big(3+i\big)\big(2+4i\big)\big(1+i\big), \qquad \frac{\big(3+4i\big)\big(-1+2i\big)}{\big(-1-i\big)\big(3-i\big)}.$$

(k) Prove that, for complex a and b, if either $|a| = 1$ or $|b| = 1$ then

$$\left|\frac{a-b}{1-\bar{a}b}\right| = 1.$$

What exception must be made if $|a| = |b| = 1$?

(l) Show that there are complex numbers z satisfying $|z-a| + |z+a| = 2|c|$ if and only if $|a| \le |c|$. If this condition is fulfilled, what are the smallest and largest values of $|z|$?

(m) Prove the complex form of Lagrange's identity, *viz.*, for complex $\{a_j, b_j\}$

$$\Big|\sum_{j=1}^{n} a_j\, b_j\Big|^2 = \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2 - \sum_{1\le j<k\le n} \big|a_j\,\bar{b}_k - a_k\,\bar{b}_j\big|^2.$$

4.3. **Complex inequalities – principles:** (Ahlfors [1], pp. 9-11)

(a) (bonus) Show that $-|z| \le \operatorname{Re} z \le |z|$ and that $-|z| \le \operatorname{Im} z \le |z|$. When do the equalities $\operatorname{Re} z = |z|$ or $\operatorname{Im} z = |z|$ hold?

(b) (bonus) Derive the so-called triangle inequality $|z_1 + z_2| \le |z_1| + |z_2|$. Note that it extends to arbitrary sums: $|z_1 + z_2 + \cdots| \le |z_1| + |z_2| + \cdots$. Under what circumstances does the equality hold? Show that $|z_1 - z_2| \ge \big||z_1| - |z_2|\big|$.

(c) (bonus) Derive Cauchy's inequality, i.e., show that

$$\Big|\sum_{j=1}^{n} w_j\, z_j\Big|^2 \le \Big|\sum_{j=1}^{n} \big|w_j\big|^2\,\Big|\sum_{j=1}^{n} \big|z_j\big|^2.$$

4.4. **Complex inequalities:** (Ahlfors [1], p. 11)

(a) (bonus) Prove that, for complex a and b such that $|a| < 1$ and $|b| < 1$, we have $|(a-b)/(1-\bar{a}b)| < 1$.

(b) (bonus) Let $\{a_j\}_{j=1}^n$ be a set of n complex variables and let $\{\lambda_j\}_{j=1}^n$ be a set of n real variables.
If $|a_j| < 1$, $\lambda_j \ge 0$ and $\sum_{j=1}^n \lambda_j = 1$, show that $\big|\sum_{j=1}^n \lambda_j\, a_j\big| < 1$.

4.5. **Circles and lines with complex numbers:** (Needham [3] p. 46)

(a) If c is a fixed complex number and R is a fixed real number, explain with a picture why $|z - c| = R$ is the equation of a circle. Given that z satisfies the equation $|z + 3 - 4i| = 2$, find the minimum and maximum values of $|z|$ and the corresponding positions of z.

(b) Consider the two straight lines in the complex plane that make an angle $(\pi/2) + \phi$ with the real axis and lie a distance D from the origin. Show that points z on the lines satisfy one or other of $\mathrm{Re}\,\big(\cos\phi - i\sin\phi\big)z = \pm D$.

(c) Consider the circle of points obeying $\big|z - (D + R)\big(\cos\phi + i\sin\phi\big)\big| = R$. Give the centre of this circle and its radius. Determine what happens to this circle in the $R \to \infty$ limit. (Note: In the *extended* complex plane the properties of circles and lines are unified. For this reason they are sometimes referred to as *circlines*.)

4.6. **Plane geometry with complex numbers:** (Ahlfors [1], p. 15)

(a) Prove that if the points a_1, a_2 and a_3 are the vertices of an equilateral triangle then $a_1\,a_1 + a_2\,a_2 + a_3\,a_3 = a_1\,a_2 + a_2\,a_3 + a_3\,a_1$.

(b) Suppose that a and b are two vertices of a square in the complex plane. Find the two other vertices in all possible cases.

(c) (bonus) Find the center and the radius of the circle that circumscribes the triangle having vertices a_1, a_2 and a_3. Express the result in symmetric form.

(d) (bonus) Find the symmetric points of the complex number z with respect to each of the lines that bisect the coordinate axes.

4.7. **More plane geometry with complex numbers:** (Needham [3] p. 16)
Consider the quadrilateral having *sides* given by the complex numbers $2a_1$, $2a_2$, $2a_3$ and $2a_4$, and construct the squares on these sides. Now consider the two line-segments joining the centres of squares on opposite sides of the quadrilateral. Show that these line-segments are perpendicular and of equal length.

4.8. **More plane geometry with complex numbers:** (Ahlfors [1], p. 9, 17)

(a) Find the conditions under which the equation $az + b\bar{z} + c = 0$ (with complex a, b and c) in one complex unknown z has exactly one solution, and compute that solution. When does the equation represent a line?

(b) (bonus) Write the equation of an ellipse, hyperbola and parabola in complex form.

(c) (bonus) Show, using complex numbers, that the diagonals of a parallelogram bisect each other.

(d) (bonus) Show, using complex numbers, that the diagonals of a rhombus are orthogonal.

(e) (bonus) Show that the midpoints of parallel chords to a circle lie on a diameter perpendicular to the chords.

(f) (bonus) Show that all circles that pass through a and $1/a$ intersect the circle $|z| = 1$ at right angles.

4.9. **Number theory with complex numbers:** (Needham [3] p. 45)
Here is a basic fact that has many uses in number theory: *If two integers can be expressed as the sum of two squares then so can their product*. Prove this result by considering $\big|(A + iB)(C + iD)\big|^2$ for integers A, B, C and D.

4.10. **Trigonometry with complex numbers:** (Ahlfors [1], pp. 16-17)

(a) Express $\cos 3\phi$, $\cos 4\phi$ and $\sin 5\phi$ in terms of $\cos\phi$ and $\sin\phi$.

(b) Simplify $1 + \cos\phi + \cos 2\phi + \cdots + \cos N\phi$ and $\sin\phi + \sin 2\phi + \sin 3\phi + \cdots + \sin N\phi$.

(c) Express the fifth and tenth roots of unity in algebraic form.

(d) (bonus) If ω is given by $\omega = \cos\big(2\pi/N\big) + i\sin\big(2\pi/N\big)$ (for $N = 0, 1, 2, \ldots$), show that, for any integer H that is not a multiple of N, $1 + \omega^H + \omega^{2H} + \cdots + \omega^{(N-1)H} = 0$. What is the value of $1 - \omega^H + \omega^{2H} - \cdots + (-1)^{N-1}\omega^{(N-1)H}$?

Chapter 4 solutions: Complex differentiation

Solution 4.2 - Complex arithmetic.

Uploaded to Canvas as HW04.pdf.

mathematical methods - week 5

Complex integration

Georgia Tech PHYS-6124

Homework HW #5 due Sunday, September 27, 2020

== 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 5.1 *More holomorphic mappings* 10 (+6 bonus) points

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

edited September 24, 2020

Week 5 syllabus

September 15, 2020

Arfken, Weber & Harris [1] (click here) Chapter 11 *Complex variable theory*. The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- Complex integration : Tuesday lecture segments
  - AWH Sect. 11.3 *Cauchy's integral theorem*
  - *Holomorphic functions*
  - *Complex exponential, logarithm*
  - *Complex logarithm, take #2*
  - *Integration in the complex plane, part I*
- Cauchy contour integral : Thursday lecture
  - *Cauchy integration theorem*
  - AWH Sect. 11.4 *Cauchy's integral formula*

Optional reading

- Grigoriev pages 3.1 - 3.3 (Cauchy's contour integral)

Riemann sphere by waving hands; Quaternions, octonions; bifurcations

Is conformal a 2D version of symplectic? David Finkelstein discovers Higgs before Higgs; Why cannot colloquia speakers stop blabbing?

- Discussion continued, **Andrew Wu**:
  - Conservative field in complex plane:
    Cauchy's theorem and conservative vector fields
  - Cauchy-Riemann equations wikipedia sect. *Harmonic vector field*.
    Predrag: Also the next sect. *Preservation of complex structure* is interesting, points out that Cauchy–Riemann equation have symplectic structure, the defining property of Hamilton's equations. There is no clear path forward to mechanics with more degrees of freedom, though.

Tuesday morning chit-chat

Thursday morning chit-chat : baseball, and getting humiliated by the 10 year old Peter Serene

- Stone and Goldbart [2] (click here)
  - SG 17.2-3 *Complex integration: Cauchy and Stokes*
  - SG 17.2.2 *Cauchy's theorem*
  - SG 17.2.3 *The residue theorem*

References

[1] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. (Academic, New York, 2013).

[2] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

Exercises

5.1. **More holomorphic mappings.** **Needham**, pp. 211-213

(a) **(bonus)** Use the Cauchy-Riemann conditions to verify that the mapping $z \mapsto \bar{z}$ is not holomorphic.

(b) The mapping $z \mapsto z^3$ acts on an infinitesimal shape and the image is examined. It is found that the shape has been rotated by π, and its linear dimensions expanded by 12. Determine the possibilities for the original location of the shape, i.e., find all values of the complex number z for which an infinitesimal shape at z is rotated by π, and its linear dimensions expanded by 12. Hint: write z in polar form, first find the appropriate $r = |z|$, then find all values of the phase of z such that $\arg(z^3) = \pi$.

(c) Consider the map $z \mapsto \bar{z}^2/z$. Determine the geometric effect of this mapping. By considering the effect of the mapping on two small arrows emanating from a typical point z, one arrow parallel and one perpendicular to z, show that the map fails to produce an *amplitwist*.

(d) The interior of a simple closed curve $\mathcal{C}$ is mapped by a holomorphic mapping into the exterior of the image of $\mathcal{C}$. If z travels around the curve counterclockwise, which way does the image of z travel around the image of $\mathcal{C}$?

(e) Consider the mapping produced by the function $f(x+iy) = \big(x^2+y^2\big) + i\big(y/x\big)$.

  (i) Find and sketch the curves that are mapped by f into horizontal and vertical lines. Notice that f appears to be conformal.

  (ii) Now show that f is *not* in fact a conformal mapping by considering the images of a pair of lines (e.g. , one vertical and one horizontal).

  (iii) By using the Cauchy-Riemann conditions confirm that f is not conformal.

  (iv) Show that no choice of $v(x,y)$ makes $f(x+iy) = \big(x^2+y^2\big) + iv(x,y)$ holomorphic.

(f) **(bonus)** Show that if f is holomorphic on some connected region then each of the following conditions forces f to reduce to a constant:
(i) $\mathrm{Re}\, f(z) = 0$; (ii) $|f(z)| = \mathrm{const.}$; (iii) $\bar{f}(z)$ is holomorphic too.

(g) **(bonus)** Suppose that the holomorphic mapping $z \mapsto f(z)$ is expressed in terms of the modulus R and argument Φ of f, i.e.,
$f(z) = R(x,y) \exp i\Phi(x,y)$.
Determine the form of the Cauchy-Riemann conditions in terms of R and Φ.

(h) (i) By sketching the image of an infinitesimal rectangle under a holomorphic mapping, determine the the local magnification factor for the area and compare it with that for a infinitesimal line. Re-derive this result by examining the Jacobian determinant for the transformation.

  (ii) Verify that the mapping $z \mapsto \exp z$ satisfies the Cauchy-Riemann conditions, and compute $\big(\exp z\big)'$.

  (iii) **(bonus)** Let S be the square region given by $A - B \le \mathrm{Re}\, z \le A + B$ and $-B \le \mathrm{Im}\, z \le B$ with A and B positive. Sketch a typical S for which $B < A$ and sketch the image $\tilde{S}$ of S under the mapping $z \mapsto \exp z$.

  (iv) **(bonus)** Deduce the ratio $(\text{area of } \tilde{S})/(\text{area of } S)$, and compute its limit as $B \to 0^+$.

  (v) **(bonus)** Compare this limit with the one you would expect from part (i).

Chapter 5 solutions: Complex integration

Solution 5.1 - More holomorphic mappings.

Uploaded to Canvas as `HW05.pdf`

mathematical methods - week 6

Cauchy theorem at work

Georgia Tech PHYS-6124

Homework HW #6 due Thursday, October 1, 2020

== 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 6.1 *Complex integration* (a) 4; (b) 2; (c) 2; and (d) 3 points
Exercise 6.2 *Fresnel integral* 7 points

Bonus points
Exercise 6.4 *Cauchy's theorem via Green's theorem in the plane* 6 points

Total of 16 points = 100 % score. Extra points accumulate, can help you later if you miss a few problems.

edited September 26, 2020

Week 6 syllabus

September 22, 2020

Mephistopheles knocks at Faust's door and says, "Du mußt es dreimal sagen!"

"You have to say it three times"

— Johann Wolfgang von Goethe
Faust I - Studierzimmer 2. Teil

Arfken, Weber & Harris [1] (click here) Chapter 11 *Complex variable theory*. The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

The essence of complex; Taylor, Laurent series; residue calculus part 1

- AWH 11.5 *Laurent expansion*
- AWH 11.6 *Singularities*
- AWH 11.7 *Calculus of residues*

Calculus of residues. A few integrals, evaluated by Cauchy contours

- AWH 11.8 *Evaluation of definite integrals*

- Grigoriev examples worked out in the lecture:
  - *Meromorphic in upper half-plane*
  - *Singularity on the contour*
  - *Pole in upper half-plane*
  - *Singularity on the contour*

Optional reading

Stone and Goldbart [2] (click here) Chapter 17

- SG 17.4 *Applications of Cauchy's theorem*
- SG 17.4.2 *Taylor and Laurent series*
- SG 17.4.3 *Zeros and singularities*
- SG 17.4.4 *Analytic continuation*

Wolfram rant: the wunderkid vs. Gradshteyn and Ryzhik; opinions of blackest reactionary professor on graduate educations (the kids are OK). Click on this at your own risk - 30 minutes! Absolutely no science.

The meaning of the things complex rant: The power of visual thinking; Data and dimension reduction; AI, hype and morality. Click on this at your own risk.

References

[1] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. (Academic, New York, 2013).

[2] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

Exercises

6.1. **Complex integration.**

(a) Write down the values of $\oint_C (1/z)\,dz$ for each of the following choices of C:
(i) $|z| = 1$, (ii) $|z-2| = 1$, (iii) $|z-1| = 2$.
Then confirm the answers the hard way, using parametric evaluation.

(b) Evaluate parametrically the integral of $1/z$ around the square with vertices $\pm 1 \pm i$.

(c) Confirm by parametric evaluation that the integral of z^m around an origin centered circle vanishes, except when the integer $m = -1$.

(d) Evaluate $\int_{1+i}^{3-2i} dz\ \sin z$ in two ways: (i) via the fundamental theorem of (complex) calculus, and (ii) (bonus) by choosing any path between the end-points and using real integrals.

6.2. **Fresnel integral.**

We wish to evaluate the $I = \int_0^\infty \exp\big(ix^2\big)\,dx$. To do this, consider the contour integral $I_R = \int_{C(R)} \exp\big(iz^2\big)\,dz$, where $C(R)$ is the closed circular sector in the upper half-plane with boundary points 0, R and $R\exp(i\pi/4)$. Show that $I_R = 0$ and that $\lim_{R\to\infty} \int_{C_1(R)} \exp\big(iz^2\big)\,dz = 0$, where $C_1(R)$ is the contour integral along the circular sector from R to $R\exp(i\pi/4)$. [Hint: use $\sin x \geq (2x/\pi)$ on $0 \leq x \leq \pi/2$.] Then, by breaking up the contour $C(R)$ into three components, deduce that

$$\lim_{R\to\infty} \left(\int_0^R \exp\big(ix^2\big)\,dx - \mathrm{e}^{i\pi/4} \int_0^R \exp\big(-r^2\big)\,dr \right) = 0$$

and, from the well-known result of real integration $\int_0^\infty \exp\big(-x^2\big)\,dx = \sqrt{\pi}/2$, deduce that $I = \mathrm{e}^{i\pi/4}\sqrt{\pi}/2$.

6.3. **Fresnel integral.**

(a) Derive the Fresnel integral

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, e^{-\frac{x^2}{2ia}} = \sqrt{ia} = |a|^{1/2} e^{i\frac{\pi}{4}\frac{a}{|a|}}\,.$$

Consider the contour integral $I_R = \int_{C(R)} \exp\big(iz^2\big)\,dz$, where $C(R)$ is the closed circular sector in the upper half-plane with boundary points 0, R and $R\,\exp(i\pi/4)$. Show that $I_R = 0$ and that $\lim_{R\to\infty} \int_{C_1(R)} \exp\big(iz^2\big)\,dz = 0$, where $C_1(R)$ is the contour integral along the circular sector from R to $R\exp(i\pi/4)$. [Hint: use $\sin x \geq (2x/\pi)$ on $0 \leq x \leq \pi/2$.] Then, by breaking up the contour $C(R)$ into three components, deduce that

$$\lim_{R\to\infty} \left(\int_0^R \exp\big(ix^2\big)\,dx - \mathrm{e}^{i\pi/4} \int_0^R \exp\big(-r^2\big)\,dr \right)$$

vanishes, and, from the real integration $\int_0^\infty \exp\big(-x^2\big)\,dx = \sqrt{\pi}/2$, deduce that

$$\int_0^\infty \exp\big(ix^2\big)\,dx = \mathrm{e}^{i\pi/4}\sqrt{\pi}/2\,.$$

Now rescale x by real number $a \neq 0$, and complete the derivation of the Fresnel integral.

(b) In exercise 9.2 the exponent in the d-dimensional Gaussian integrals is real, so the real symmetric matrix M in the exponent has to be strictly positive definite. However, in quantum physics one often has to evaluate the d-dimenional Fresnel integral

$$\frac{1}{(2\pi)^{d/2}} \int d^d\phi e^{-\frac{1}{2i}\phi^\top \cdot M^{-1} \cdot \phi + i\,\phi \cdot J} ,$$

with a Hermitian matrix M. Evaluate it. What are conditions on its spectrum in order that the integral be well defined?

6.4. **Cauchy's theorem via Green's theorem in the plane.** Express the integral $\oint_C dz\, f(z)$ of the analytic function $f = u + iv$ around the simple contour C in parametric form, apply the two-dimensional version of Gauss' theorem (a.k.a. Green's theorem in the plane), and invoke the Cauchy-Riemann conditions. Hence establish Cauchy's theorem $\oint_C dz\, f(z) = 0$.

Chapter 6 solutions: Cauchy theorem at work

Solution 6.1 - Complex integration. *Uploaded to Canvas as* `HW06-1.pdf`

Solution 6.1 - Complex integration.

(a) *Using Cauchy's residue theorem, the curves (i) and (iii) enclose the origin but (ii) does not:*

(i) $2\pi i$, *(ii)* 0, *(iii)* $2\pi i$. *With parametric integration:*

(i) If we write z *as* $re^{i\theta}$, $dz = ire^{i\theta}\, d\theta$ *(*$r = 1$*). Thus,* $\int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}}\, d\theta = 2\pi i$.

(ii) z *is now* $re^{i\theta} + 2$*;* dz *is still* $ire^{i\theta} d\theta$ *and* $r = 1$*. Our integral is now*

$$\begin{aligned} \oint_C \frac{1}{z}\, dz &= \int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}+2}\, d\theta \\ &= \int_{\theta=0}^{2\pi} \frac{1}{u}\, du \quad (u \equiv e^{i\theta}+2) \\ &= \Big[\ln(e^{i\theta}+2)\Big]_0^{2\pi} = 0\,. \end{aligned} \tag{6.1}$$

(iii) Similar to part (ii), $z = 2e^{i\theta} + 1$*;* dz *is now* $i2e^{i\theta} d\theta$*, so:*

$$\begin{aligned} \oint_C \frac{1}{z}\, dz &= \int_0^{2\pi} \frac{2ie^{i\theta}}{2e^{i\theta}+1}\, d\theta \\ &= \int_0^{2\pi} \frac{2ie^{i\theta'}-i}{2e^{i\theta'}}\, d\theta' \quad (2e^{i\theta'} \equiv 2e^{i\theta}+1) \\ &= \int_0^{2\pi} \Big(\frac{2ie^{i\theta'}}{2e^{i\theta'}} - \frac{i}{2e^{i\theta'}}\Big)\, d\theta' \\ &= 2\pi i - \left[\frac{-1}{2} e^{-i\theta'}\right]_0^{2\pi} = 2\pi i \end{aligned} \tag{6.2}$$

(b) *We split the integral into 4 parts (the four parts of the square), with the four paths:* $z_1 = x + i, z_2 = -1 + iy, z_3 = x - i, z_4 = 1 + iy$*. Then,*

$$\begin{aligned} \oint_1 \frac{1}{z}\, dz = \int_1^{-1} \frac{1}{x+i}\, dx = \int_1^{-1} \frac{x-i}{1+x^2}\, dx &= \left[\ln(1+x^2) - i\tan^{-1}(x)\right]_1^{-1} = i\pi/2 \\ \oint_2 \frac{1}{z}\, dz = \int_1^{-1} \frac{i}{-1+iy}\, dy = \int_1^{-1} \frac{-i+y}{1+y^2}\, dy &= \left[\ln(1+y^2) - i\tan^{-1}(y)\right]_1^{-1} = i\pi/2 \\ \oint_3 \frac{1}{z}\, dz = \int_{-1}^{1} \frac{1}{x-i}\, dx = \int_{-1}^{1} \frac{x+i}{1+x^2}\, dx &= \left[\ln(1+x^2) + i\tan^{-1}(x)\right]_{-1}^{1} = i\pi/2 \\ \oint_2 \frac{1}{z}\, dz = \int_{-1}^{1} \frac{i}{1+iy}\, dy = \int_{-1}^{1} \frac{i+y}{1+y^2}\, dy &= \left[\ln(1+y^2) + i\tan^{-1}(y)\right]_{-1}^{1} = i\pi/2 \end{aligned}$$

Add the four integrals. All the messy terms cancel, leaving four $i\pi/2$*. Thus,* $\oint_C \frac{1}{z}\, dz = 2\pi i$.

(c) *Write* z *as* $r^{ie\theta}$ *and use* $dz = ire^{i\theta}\, d\theta$.

$$\oint_C z^m\, dz = \int_0^{2\pi} ir^{m+1} e^{i(m+1)\theta}\, d\theta = \left[\frac{ir^{m+1}}{i(m+1)} e^{i(m+1)\theta}\right]_0^{2\pi} \tag{6.3}$$

$e^{i(m+1)\theta}$ *is periodic by* 2π*, thus if this term remains,* $\oint_C z^m\,dz = 0$*. To kill this term we set* $m = -1$ *and* $\oint_C z^m\,dz$ *turns into a singularity. We reevaluate the integral for* $m = -1$ *and get* $\oint_C z^{-1}\,dz = \int_0^{2\pi} i\,d\theta = 2\pi i$*.* □

(d)

(i)

$$\int \sin(z)\,dz = \int \frac{i}{2}(e^{-iz} - e^{iz})\,dz = \frac{i}{2}(ie^{-iz} + ie^{iz}) = -\cos(z) \quad (6.4)$$

Using the fundamental theorem of calculus,

$$\begin{aligned}
\int_{1+i}^{3-2i} \sin(z)\,dz &= [-\cos(z)]_{1+i}^{3-2i} \\
&= -\cos(3-2i) + \cos(1+i) \\
&= \frac{-1}{2}(e^{i(3-2i)} + e^{-i(3-2i)}) + \frac{1}{2}(e^{i(1+i)} + e^{-i(1+i)}) \\
&= \frac{1}{2}\left[e^{-1}e^{i} + e^{1}e^{-i} - e^{2}e^{i3} - e^{-2}e^{-i3}\right] \\
&= \cosh(1)\cos(1) - i\sinh(1)\sin(1) - \cosh(2)\cos(3) - i\sinh(2)\sin(3)
\end{aligned}$$

(ii) ***(bonus)***

$$\begin{aligned}
\sin(z) &= \frac{i}{2}(e^{-iz} - e^{iz}) = \frac{i}{2}(e^{-i(x+iy)} - e^{i(x+iy)}) \\
&= \frac{i}{2}(e^{y}e^{-ix} - e^{-y}e^{ix}) \\
&= \frac{1}{2}(ie^{y}(\cos(x) - i\sin(x)) - ie^{-y}(\cos(x) + i\sin(x))) \\
&= \frac{1}{2}(e^{y}\sin(x) + e^{-y}\sin(x) + ie^{y}\cos(x) - e^{-y}cos(x)) \\
&= \sin(x)\cosh(y) + i\cos(x)\sinh(y) \qquad (6.5)
\end{aligned}$$

We follow the two paths $z = x + i$ *and* $z = 3 + iy$*.*

$$\begin{aligned}
\int_{1+i}^{3-2i} \sin(z)\,dz &= \int_1^3 \sin(x)\cosh(1) + i\cos(x)\sinh(1)\,dx \\
&\quad + \int_1^{-2} \sin(3)\cosh(y) + i\cos(3)\sinh(y)i\,dy \\
&= \cosh(1)(-\cos(3) + \cos(1)) + i\sinh(1)(\sin(3) - \sin(1)) \\
&\quad +i\sin(3)(\sinh(-2) - \sinh(1)) - \cos(3)(\cosh(-2) - \cosh(1)) \\
&= \cosh(1)cos(1) - i\sinh(1)\sin(1) - i\sinh(2)\sin(3) - \cosh(2)\cos(3) \\
&= \text{what we previously got (rearranged)} \qquad (6.6)
\end{aligned}$$

Arthur Lin

Solution 6.2 - Fresnel integral. *Uploaded to Canvas as* `HW06-2.pdf`

Solution 6.2 - Fresnel integral. *We break the contour integral into three parts,*

$z_1(x) = x, z_2(\theta) = R\,e^{i\theta}, z_3(r) = r\,e^{i\pi/4}$:

$$\begin{aligned}
\oint_C e^{iz^2}\,dz &= \int_0^R e^{ix^2}\,dx + \int_0^{\pi/4} e^{i(Re^{i\theta})^2} Rie^{i\theta}\,d\theta + \int_R^0 e^{i(re^{i\pi/4})^2} e^{i\pi/4} dr \\
&= \int_0^R e^{ix^2}\,dx + iR\int_0^{\pi/4} e^{i(R^2e^{i2\theta})} e^{i\theta}\,d\theta + \int_R^0 e^{ir^2(e^{i\pi/2})} e^{i\pi/4} dr \\
&= \int_0^R e^{ix^2}\,dx + iR\int_0^{\pi/4} e^{i(R^2e^{i2\theta})} e^{i\theta}\,d\theta - e^{i\pi/4}\int_0^R e^{-r^2} dr \qquad (6.7) \\
&= 0 \quad \text{(there are no residues in this region).}
\end{aligned}$$

We now take a closer look at the integral $\int_0^{\pi/4} e^{i(R^2e^{i2\theta})}e^{i\theta}\,d\theta$.

$$\int_0^{\pi/4} e^{i(R^2e^{i2\theta})} e^{i\theta}\,d\theta = \int_0^{\pi/4} e^{i(R^2\cos(2\theta))} e^{-(R^2\sin(2\theta))} e^{i\theta}\,d\theta$$

Both $e^{i(R^2\cos(2\theta))}$ and $e^{i\theta}$ are always oscillatory terms while $e^{-(R^2\sin(2\theta))}$ is decaying with R. Noting that $e^{-(R^2\sin(2\theta))} = 0$ when $R\to\infty$, we assume that this integral also goes to 0 when $R\to\infty$. Substituting these back into (6.7) and taking $R\to\infty$,

$$\lim_{R\to\infty}\left(\int_0^R e^{ix^2}\,dx - e^{i\pi/4}\int_0^R e^{-r^2} dr\right) = 0\,.$$

Taking the limit and rearranging we have:

$$\int_0^\infty e^{ix^2}\,dx = e^{i\pi/4}\int_0^\infty e^{-r^2}\,dr = e^{i\pi/4}\sqrt{\pi}/2\,.$$

Arthur Lin

Solution 6.3 - Fresnel integral.

1. *Predrag 2019-09-29 Uploaded to T-square:* `HW04-2.pdf` Exercise 5.2 *taken out of HW12.pdf.*
2. *No solution available.*

Solution 6.4 - Cauchy's theorem via Green's theorem in the plane.

Uploaded to Canvas as `HW06-3.pdf`

Solution 6.4 - Cauchy's theorem via Green's theorem in the plane.

$$\oint_C f(z)\,dz = \oint_C (u+iv)\,d(x+iy) = \oint_C (u\,dx - v\,dy) + i\oint_C (v\,dx + u\,dy) \qquad (6.8)$$

Using Gauss's Law, the integral becomes:

$$\int_S \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) dx\,dy + i\int_S \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy \qquad (6.9)$$

If the function is analytic, the Cauchy-Riemann conditions require

$$\begin{aligned}
\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} &= 0 \qquad (6.10) \\
\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} &= 0. \qquad (6.11)
\end{aligned}$$

With both integrands equal to zero, the whole integral is also zero, thus $\oint_C f(z)\,dz = 0$ for analytic functions. □ *Arthur Lin*

mathematical methods - week 7

Method of steepest descent

Georgia Tech PHYS-6124

Homework HW #7 due Thursday, October 8, 2020

== 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 7.1 *In high dimensions any two vectors are (nearly) orthogonal* 16 points

Bonus points
Exercise 7.2 *Airy function for large arguments* 10 points

Total of 16 points = 100 % score. Extra points accumulate, can help you later if you miss a few problems.

edited October 10, 2020

Week 7 syllabus

September 29, 2020

Arfken, Weber & Harris [1] Chapter 12 *Further Topics in Analysis* (click here); Chapter 13 *Gamma function* (click here). saddle point method; Gamma, Airy function estimates; beta function is also often encountered. The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- *Asymptotic evaluation of integrals: perturbation expansions; Laplace, saddle, steepest descent leading term. (1 hour)*
  - AWH 12.6 *Asymptotic series*
  - AWH 12.7 *Method of steepest descents*
  - Grigoriev lecture notes
- *Steepest descent I: Gamma function, Sterling formula. (27 min)*
- *Steepest descent II, for physicists: Zero-dimensional field theory - perturbation expansion is an asymptotic series.*
  - Sect. 7.1 *Saddle-point expansions are asymptotic*
- *Steepest descent III, for data scientists: How tall is my graduate student?*
  - Exercise 7.1 *In high dimensions any two vectors are (nearly) orthogonal*

Optional reading

- Steven Strogatz *Asymptotics and perturbation methods* (2021)
- AWH 11.6 *Singularities*; Branch-cut integrals
- *If they only got the phase in the Fresnel integral right, QM would look different*
- *Got problems? Do them like a journalist*
- *I heard it through the grapevine: how to pick a tolerable adviser?*
- *You think you are stressed? Try finishing your thesis*

7.1 Saddle-point expansions are asymptotic

The first trial ground for testing our hunches about field theory is the *zero-dimensional field theory*, the field theory of a lattice consisting of one point, in case of the "ϕ^4 theory" given a humble 1-dimensional integral

$$Z[J] = \int \frac{d\phi}{\sqrt{2\pi}} e^{-\phi^2/2 - g\phi^4/4 + \phi J} \,. \tag{7.1}$$

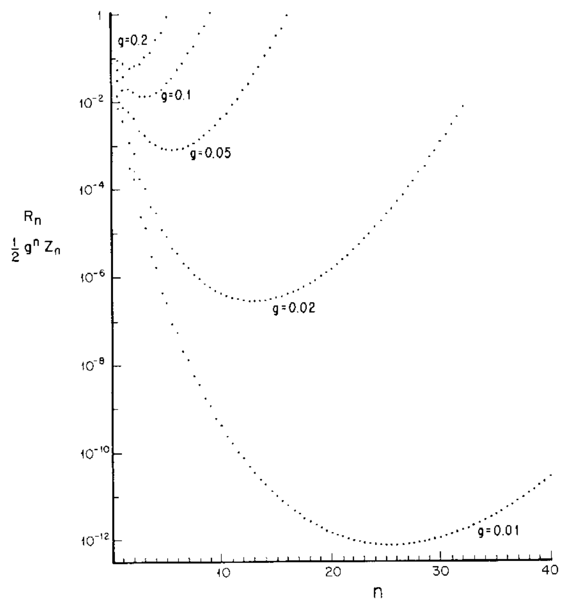


Figure 7.1: Plot of the saddle-point estimate of Z_n vs. the exact result (7.2) for $g = 0.1$, $g = 0.02$, $g = 0.01$.

The idea of the saddle-point expansions is to keep the Gaussian part $\phi^2/2$ ("free field", with a quadratic H_0 "Hamiltonian") as is, and expand the rest (H_I "interacting Hamiltonian") as a power series, and then evaluate the perturbative corrections using the moments formula

$$\int \frac{d\phi}{\sqrt{2\pi}}\, \phi^n e^{-\phi^2/2} = \left(\frac{d}{dJ}\right)^n e^{J^2/2}\Big|_{J=0} = (n-1)!! \quad \text{if } n \text{ even, 0 otherwise}\,.$$

In this zero-dimensional theory the n-point correlation is a number exploding combinatorially, as $(n-1)!!$. And here our troubles start.

To be concrete, let us work out the *exact* zero-dimensional ϕ^4 field theory in the saddle-point expansion to all orders:

$$\begin{aligned} Z[0] &= \sum_n Z_n g^n\,, \\ Z_n &= \frac{(-1)^n}{n!4^n} \int \frac{d\phi}{\sqrt{2\pi}} \phi^{4n} e^{-\phi^2/2} = \frac{(-1)^n}{16^n n!} \frac{(4n)!}{(2n)!}\,. \end{aligned} \tag{7.2}$$

The Stirling formula $n! = \sqrt{2\pi}\, n^{n+1/2} e^{-n}$ yields for large n

$$g^n Z_n \approx \frac{1}{\sqrt{n\pi}} \left(\frac{4g}{e} n\right)^n\,. \tag{7.3}$$

As the coefficients of the parameter g^n are blowing up combinatorially, no matter how small g might be, the perturbation expansion is not convergent! Why? Consider again

(7.1). We have tacitly assumed that $g > 0$, but for $g < 0$, the potential is unbounded for large ϕ, and the integrand explodes. Hence the partition function in not analytic at the $g = 0$ point.

Is the whole enterprise hopeless? As we shall now show, even though divergent, the perturbation series is an *asymptotic* expansion, and an asymptotic expansion can be extremely good [6]. Consider the residual error after inclusion of the first n perturbative corrections:

$$\begin{aligned} R_n &= \left| Z(g) - \sum_{m=0}^{n} g^m Z_m \right| \\ &= \int \frac{d\phi}{\sqrt{2\pi}} e^{-\phi^2/2} \left| e^{-g\phi^4/4} - \sum_{m=0}^{n} \frac{1}{m!} \left(-\frac{g}{4}\right)^m \phi^{4m} \right| \\ &\le \int \frac{d\phi}{\sqrt{2\pi}} e^{-\phi^2/2} \frac{1}{(n+1)!} \left(\frac{g\phi^4}{4}\right)^{n+1} = g^{n+1} \left| Z_{n+1} \right| . \end{aligned} \tag{7.4}$$

The inequality follows from the convexity of exponentials, a generalization of the inequality $e^x \ge 1 + x$. The error decreases as long as $g^n |Z_n|$ decreases. From (7.3) the minimum is reached at $4g\, n_{min} \approx 1$, with the minimum error

$$g^n Z_n|_{min} \approx \sqrt{\frac{4g}{\pi}}\, e^{-1/4g}. \tag{7.5}$$

As illustrated by the figure 7.1, a perturbative expansion can be, for all practical purposes, very accurate. In Quantum ElectroDynamics, or QED, this argument had led Dyson to suggest that the QED perturbation expansions are good to $n_{min} \approx 1/\alpha \approx 137$ terms. Due to the complicated relativistic, spinorial and gauge invariance structure of perturbative QED, there is not a shred of evidence that this is so. The very best calculations performed so far stop at $n \le 5$.

Predrag The truth is, perturbation theory as taught in all Quantum Mechanics textbooks is wrong, as one should also include the exponentially small contributions. That goes by names 'resurgence', 'Borel transforms', 'lateral Borel resummations', 'trans-series', and is not an easy subject. I found the May 5, 2021 École Normale Supŕieure lecture by Marcos Mariño the most accessible overview, but I do not see it online. Perhaps the start of his lecture *Resurgence, BPS counting, and knot invariants* has some of the ideas. His lectures on resurgence in mathematics and physics are more technical, if you want to learn more.

Predrag I find Córdova, Heidenreich, Popolitov and Shakirov [4] *Orbifolds and exact solutions of strongly-coupled matrix models* very surprising. The introduction is worth reading. They compute analytically the matrix model (QFT in zero dimensions) partition function for trace potential

$$S[X] = \mathrm{tr}\,(X^r)\,, \qquad \text{integer } r \ge 2\,. \tag{7.6}$$

Their "non-perturbative ambiguity" in the case of the $N = 1$ cubic matrix model seem to amount to the Stokes phenomenon, i.e., choice of integration contour for the Airy function.

Unlike the weak coupling expansions, the strong coupling expansion of

$$Z = \frac{1}{2\pi} \int dx e^{-\frac{1}{2g^2}x^2 - x^4}, \tag{7.7}$$

is convergent, not an asymptotic series.

There is a negative dimensions type duality $N \to -N$, their eq. (3.27). The loop equations, their eq. (2.10), are also interesting - they seem to essentially be the Dyson-Schwinger equations and Ward identities in my book's [5] formulation of QFT.

7.2 Notes on life in extreme dimensions

You can safely ignore this section, it's "math methods," as much as Predrag's musings about current research...

Exercise 7.1 is something that anyone interested in computational neuroscience [9] and/or machine learning already knows. It is also something that many a contemporary physicist *should* know; a daily problem for all of us, from astrophysics to fluid physics to biologically inspired physics is how to visualize large, extremely large data sets.

Possibly helpful references:

Distribution of dot products between two random unit vectors. They denote $Z = \langle X, Y \rangle = \sum X_i Y_i$. Define

$$f_{Z_i}(z_i) = \int_{-\infty}^{\infty} f_{X_i,Y_i}(x, \frac{z_i}{x}) \frac{1}{|x|} dx$$

then since $Z = \sum Z_i$,

$$f_Z(z) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{Z_1,\ldots,Z_D}(z_1, \ldots, z_d)\, \delta(z - \sum z_i)\, dz_1 \ldots dz_d \,.$$

There is a Georgia Tech paper on this [12]. See also cosine similarity and Mathworld. There is even a python tutorial. `scikit-learn` is supposed to be 'The de facto Machine Learning package for Python'.

Remark 7.1. High-dimensional flows and their visualizations. Dynamicist's vision of turbulence was formulated by Eberhard Hopf in his seminal 1948 paper [11]. Computational neuroscience grapples with closely related visualization and modeling issues [7, 8]. Much about high-dimensional state spaces is counterintuitive. The literature on why the expectation value of the angle between any two high-dimensional vectors picked at random is 90^o is mostly about spikey spheres: see the draft of the Hopcroft and Kannan [3] book and Ravi Kannan's course; lecture notes by Hermann Flaschka on *Some geometry in high-dimensional spaces*; Wegman and Solka [13] visualizations of high-dimensional data; Spruill paper [12]; a lively mathoverflow.org thread on "Intuitive crutches for higher dimensional thinking."

The 'good' coordinates, introduced in ref. [10] are akin in spirit to the low-dimensional projections of the POD modeling [2], in that both methods aim to capture key features and dynamics of the system in just a few dimensions. But the ref. [10] method is very different from POD in a key way: we construct basis sets from *exact solutions of the fully-resolved dynamics* rather than from the empirical eigenfunctions of the POD. Exact solutions and their linear stability modes (a) characterize the spatially-extended states precisely, as opposed to the truncated expansions of the POD, (b) allow for different basis sets and projections for different purposes and different regions of state space, (c) these low-dimensional projections are not meant to suggest low-dimensional ODE models; they are only visualizations, every point in these projections is still a point the full state space, and (d) the method is not limited to Fourier mode bases.

References

[1] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. (Academic, New York, 2013).

[2] N. Aubry, P. Holmes, J. L. Lumley, and E. Stone, "The dynamics of coherent structures in the wall region of turbulent boundary layer", J. Fluid Mech. **192**, 115–173 (1988).

[3] A. Blum, J. Hopcroft, and R. Kannan, *Foundations of Data Science* (Cambridge Univ. Press, Cambridge UK, 2020).

[4] C. Córdova, B. Heidenreich, A. Popolitov, and S. Shakirov, "Orbifolds and exact solutions of strongly-coupled matrix models", Commun Math Phys **361**, 1235–1274 (2018).

[5] P. Cvitanović, *Field Theory*, Notes prepared by E. Gyldenkerne (Nordita, Copenhagen, 1983).

[6] R. B. Dingle, *Asymptotic Expansions: Their Derivation and Interpretation* (Academic Press, London, 1973).

[7] A. Farshchian, J. A. Gallego, J. P. Cohen, Y. Bengio, L. E. Miller, and S. A. Solla, Adversarial domain adaptation for stable brain-machine interfaces, in International Conference on Learning Representations (2019), pp. 1–14.

[8] J. A. Gallego, M. G. Perich, R. H. Chowdhury, S. A. Solla, and L. E. Miller, "Long-term stability of cortical population dynamics underlying consistent behavior", Nature Neuroscience **23**, 260–270 (2020).

[9] J. A. Gallego, M. G. Perich, S. N. Naufel, C. Ethier, S. A. Solla, and L. E. Miller, "Cortical population activity within a preserved neural manifold underlies multiple motor behaviors", Nat. Commun. **9**, 4233 (2018).

[10] J. F. Gibson, J. Halcrow, and P. Cvitanović, "Visualizing the geometry of state-space in plane Couette flow", J. Fluid Mech. **611**, 107–130 (2008).

[11] E. Hopf, "A mathematical example displaying features of turbulence", Commun. Pure Appl. Math. **1**, 303–322 (1948).

[12] M. C. Spruill, "Asymptotic distribution of coordinates on high dimensional spheres", Elect. Comm. in Probab. **12**, 234–247 (2007).

[13] E. J. Wegman and J. L. Solka, "On some mathematics for visualizing high dimensional data", Sankhya: Indian J. Statistics, Ser. A **64**, 429–452 (2002).

Exercises

7.1. **In high dimensions any two vectors are (nearly) orthogonal.** Among humble plumbers laboring with extremely high-dimensional ODE discretizations of fluid and other PDEs, there is an inclination to visualize the ∞-dimensional state space flow by projecting it onto a basis constructed from a few random coordinates, let's say the 2nd Fourier mode along the spatial x direction against the 4th Chebyshev mode along the y direction. It's easy, as these are typically the computational degrees of freedom. As we will now show, it's easy but not smart, with vectors representing the dynamical states of interest being almost orthogonal to any such random basis.

Suppose your state space $\mathcal{M}$ is a real 10 247-dimensional vector space, and you pick from it two vectors $x_1, x_2 \in \mathcal{M}$ at random. What is the angle between them likely to be?

In the literature you might run into this question, formulated as the 'cosine similarity'

$$\cos(\theta_{12}) = \frac{x_1{}^\top \cdot x_2}{|\,x_1\,|\,|\,x_2\,|}\,. \tag{7.8}$$

Two vectors with the same orientation have a cosine similarity of 1, two vectors at 90^o have a similarity of 0, and two vectors diametrically opposed have a similarity of -1. By asking for 'angle between two vectors' we have implicitly assumed that there exist is a dot product

$$x_1{}^\top \cdot x_2 = |\,x_1\,|\,|\,x_2\,| \cos(\theta_{12})\,,$$

so let's make these vectors unit vectors, $|\,x_j\,| = 1$. When you think about it, you would be hard put to say what 'uniform probability' would mean for a vector $x \in \mathcal{M} = \mathbb{R}^{10\,247}$, but for a unit vector it is obvious: probability that x direction lies within a solid angle $d\Omega$ is $d\Omega/$(unit hyper-sphere surface).

So what is the surface of the unit sphere (or, the total solid angle) in d dimensions? One way to compute it is to evaluate the Gaussian integral

$$I_d = \int_{-\infty}^{\infty} dx_1 \cdots dx_d \, e^{-\frac{1}{2}\left(x_1^2+\cdots+x_d^2\right)} \tag{7.9}$$

in cartesian and polar coordinates. Show that

(a) In cartesian coordinates $I_d = (2\pi)^{d/2}$.

(b) Show, by examining the form of the integrand in the polar coordinates, that for an arbitrary, even complex dimension $d \in \mathbb{C}$

$$S_{d-1} = 2\pi^{d/2}/\Gamma(d/2)\,. \tag{7.10}$$

In QFT, or Quantum Field Theory, integrals over 4-momenta are brought to polar form and evaluated as functions of a complex dimension parameter d. This procedure is called the 'dimensional regularization'.

(c) Recast the integrals in polar coordinate form. You know how to compute this integral in 2 and 3 dimensions. Show by induction that the surface S_{d-1} of unit d-ball, or the total solid angle in even and odd dimensions is given by

$$S_{2k} = \frac{2(2\pi)^k}{(2k-1)!!}\,, \qquad S_{2k+1} = \frac{2\pi^{k+1}}{k!}\,. \tag{7.11}$$

However irritating to Data Scientists (these are just the Gamma function (7.10) written out as factorials), the distinction between even and odd dimensions is not silly - in Cartan's classification of all compact Lie groups, special orhtogonal groups SO$(2k)$ and SO$(2k{+}1)$ belong to two distinct infinite families of special orthogonal symmetry groups, with implications for physics in 2, 3 and 4 dimensions. For example, by the hairy ball theorem, there can be no non-vanishing continuous tangent vector field on even-dimensional d-spheres; you cannot smoothly comb hair on a 3-dimensional ball.

(d) Check your formula for $d = 2$ (1-sphere, or the circle) and $d = 3$ (2-sphere, or the sphere).

(e) What limit does S_d does tend to for large d? (Hint: it's not what you think. Try Sterling's formula).

So now that we know the volume of a sphere, what is a the most likely angle between two vectors x_1, x_2 picked at random? We can rotate coordinates so that x_1 is aligned with the 'z-axis' of the hypersphere. An angle θ then defines a meridian around the 'z-axis'.

(f) Show that probability $P(\theta)d\theta$ of finding two vectors at angle θ is given by the area of the meridional strip of width $d\theta$, and derive the formula for it:

$$P(\theta) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(d/2)}{\Gamma((d-1)/2)} \,.$$

(One can write analytic expression for this in terms of beta functions, but it is unnecessary for the problem at hand).

(g) Show that for large d the probability $P(\theta)$ tends to a normal distribution with mean $\theta = \pi/2$ and variance $1/d$.

So, in d-dimensional vector space the two random vectors are nearly orthogonal, within accuracy of $\theta = \pi/2 \pm 1/d$.

Null distribution: For data which can be negative as well as positive, the null distribution for cosine similarity is the distribution of the dot product of two independent random unit vectors. This distribution has a mean of zero and a variance of $1/d$ (where d is the number of dimensions), and although the distribution is bounded between -1 and +1, as d grows large the distribution is increasingly well-approximated by the normal distribution.

In high dimensions any two vectors are (nearly) orthogonal - If I am 2 meters tall, how tall does a graduate student look to me, if grad students are randomly distributed in a million directions?

If you are a humble plumber simulating turbulence, and trying to visualize its state space and the notion of a vector space is some abstract hocus-pocus to you, try thinking this way. Your 2nd Fourier mode basis vector is something that wiggles twice along your computation domain. Your turbulent state is very wiggly. The product of the two functions integrated over the computational domain will average to zero, with a small leftover. We have just estimated that with dumb choices of coordinate bases this leftover will be of order of $1/10\,247$, which is embarrassingly small for displaying a phenomenon of order ≈ 1.

Several intelligent choices of coordinates for state space projections are described in ChaosBook section 2.4, the web tutorial ChaosBook.org/tutorials, and Gibson *et al.* [10].

Sara A. Solla and P. Cvitanović

7.2. **Airy function for large arguments.** Important contributions as stationary phase points may arise from extremal points where the first non-zero term in a Taylor expansion of the phase is of third or higher order. Such situations occur, for example, at bifurcation points or in diffraction effects, such as waves near sharp corners, waves creeping around obstacles, etc.. In such calculations, one meets Airy functions integrals of the form

$$Ai(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy \, e^{i(xy - \frac{y^3}{3})} \, . \tag{7.12}$$

Calculate the Airy function $Ai(x)$ using the stationary phase approximation. What happens when considering the limit $x \to 0$? Estimate for which value of x the stationary phase approximation breaks down.

Chapter 7 solutions: Method of steepest descent

Solution 7.1 - In high dimensions any two vectors are (nearly) orthogonal.

(a) *The d Gaussian integrals are independent and identical; compute one, raise it to power d.*

(b) *There are two ways of writing the volume element dV. One of them is $dx_1....dx_d$, as in (7.9). The other way is to note that as in polar coordinates*

$$I_d = \int d\Omega_d \int_0^\infty dr^{d-1} e^{-\frac{1}{2}r^2}$$

the integrand depends only on r^2, we can construct volume elements for which r is constant, by considering the volume between two surfaces: one being the outer shell of a sphere of radius r, and the next one being the outer shell of a sphere of radius $r + dr$. What is the volume of the space between these two spheres? It is the surface S_{d-1} of the shell that bounds the sphere of radius r times dr (the convention is that the surface of 3-dimensionalunit ball is the 2-dimensional, and denoted S_2). So: $dV = S_{d-1}dr$. What is S_{d-1}? By definition, the solid angle $\Omega_d = S_{d-1}/r^{d-1}$. Thus, $S_{d-1} = \Omega_d r^{d-1}$, and $dV = \Omega_d r^{d-1}dr$. With substitution $x = r^2/2$, the integral then becomes

$$I_d = \Omega_d \int_0^\infty dr\, r^{d-1} e^{-r^2/2} = \Omega_d\, 2^{d/2-1} \int_0^\infty dx\, x^{d/2-1} e^{-x} = \Omega_d\, 2^{d/2-1}\Gamma(d/2)\,.$$

Equating the two expressions for I_d:

$$(2\pi)^{d/2} = \Omega_d 2^{d/2-1}\, \Gamma(d/2)\,.$$

So the surface of the d-dimensional unit sphere, or the total solid angle for arbitrary, perhaps even complex dimension $d \in \mathbb{C}$ is

$$S_{d-1} = 2\pi^{d/2}/\Gamma(d/2)\,.$$

This derivation is much simpler than the double factorials (7.11), and having two expressions, one for even d, one for odd d. The only tricky part, conceptually, is recognizing that $\Omega_d = S_{d-1}/r^{d-1}$. But this is the very definition of a solid angle.

(c) *You can Google for volumes of spheres. The point of this exercise is to illustrate how much more elegant the solution is if you do not think of dimension as an integer, as in part (b).*

(d) *Surface of the unit 1-sphere, i.e., perimeter of the unit circle is $S_1 = 2\pi$. Surface of the unit 2-sphere, i.e., surface of the unit 3-ball is $S_2 = 4\pi$.*

(e) *The largest S_d turns out to be $d = 7$; for d larger than that, S_d tends to zero. For an explanation, see, for example, Haber's notes.*

(f) *As*

$$S_d = \int_0^\pi S_{d-1}(\sin\theta)^{d-1} d\theta\,, \tag{7.13}$$

the probability $P(\theta)\; d\theta$ of finding two vectors at an angle θ is

$$P(\theta)\; d\theta = \frac{S_{d-1}}{S_d}(\sin\theta)^{d-1} d\theta \tag{7.14}$$

Using the result from part (c),

$$
\begin{aligned}
P(\theta)\, d\theta &= \frac{(2\pi)^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})} \frac{\Gamma(\frac{d}{2})}{(2\pi)^{\frac{d}{2}}} (\sin\theta)^{d-2} d\theta \\
&= \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} (\sin\theta)^{d-2} d\theta \qquad (7.15)
\end{aligned}
$$

In high dimensions $\sin\theta \to 1$ *(equivalently,* $\cos\theta \to 0$*) and*

$$P(\theta) \approx \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} \qquad (7.16)$$

when $d \gg 1$.

(g) *Using the result from part (f),*

$$P(\theta) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} (\sin\theta)^{d-2}$$

For simplicity, we make a change of coordinates: $\theta' \equiv \theta - \pi/2$ *so that* $\sin\theta \to \cos\theta'$*. Then,*

$$P(\theta') = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} (\cos\theta')^{n} \qquad (7.17)$$

where $n \equiv d-2$*. Taylor expanding the cosine and considering the large dimension limit* $n \to \infty$*,*

$$(\cos\theta')^n \to \left(1 - \frac{\theta'^2}{2}\right)^n \qquad (7.18)$$

Recalling the binomial theorem,

$$(1+x)^n = \sum_{j=0}^{n} \binom{n}{j} x^j = \sum_{j=0}^{n} \frac{n!}{j!\,(n-j)!} x^j \qquad (7.19)$$

The cosine term can then be written as

$$
\begin{aligned}
(\cos\theta')^n \to \sum_{j=0}^{n} \frac{n!}{j!\,(n-j)!} \left(-\frac{\theta'^2}{2}\right)^j &\approx \sum_{j=0}^{n} \frac{n^j}{j!} \left(-\frac{\theta'^2}{2}\right)^j \\
&= \sum_{j=0}^{n} \frac{1}{j!} \left(-\frac{n}{2}\,\theta'^2\right)^j = \exp\left(-\frac{n}{2}\,\theta'^2\right) \qquad (7.20)
\end{aligned}
$$

in the limit of large n*.Thus, the probability is nearly Gaussian with mean* $\mu = \pi/2$ *and the variance* $\sigma^2 = 1/n = 1/(d-2) \approx 1/d$*, standard deviation* $\approx \sqrt{d}$*.*

Sara A. Solla, Kimberly Y. Short, and P. Cvitanović

Solution 7.1 - In high dimensions any two vectors are (nearly) orthogonal.

(a) *The exponential* $\exp(-\frac{1}{2}(x_1^2 + x_2^2 + \cdots + x_d^2))$ *is a product of the exponentials of independent coordinates:* $\prod_{n=1}^{d} \exp(-x_n^2/2)$*. Gaussian integral we know how to do:*

$$\prod_{n=1}^{d} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x_n^2}\, dx_n = \prod_{n=1}^{d} \sqrt{2\pi} = (2\pi)^{\frac{d}{2}}$$

(c) In polar coordinates, $r^2 = x_1^2 + \cdots x_d^2$. The integral is now $\int_0^\infty \int_{min}^{max} e^{-\frac{1}{2}r^2} r^{d-1}\, d\Omega_{d-1}\, dr$, where Ω is the solid angle. The integral of $d\Omega_{d-1}$ yields the surface area = S_{d-1}, so

$$\int_{min}^{max} d\Omega_{d-1} \int_0^\infty e^{-\frac{1}{2}r^2} r^{d-1}\, dr = S_{d-1} \int_0^\infty e^{-\frac{1}{2}r^2} r^{d-1}\, dr = (2\pi)^{\frac{d}{2}}\,. \tag{7.21}$$

For the familiar 2- and 3-dimensional sphere, the surface area is 2π and 4π, respectively:

$$\begin{aligned} S_1 \int_0^\infty e^{-\frac{1}{2}r^2} r\, dr = 2\pi \quad &\Rightarrow \quad \int_0^\infty e^{-\frac{1}{2}r^2} r\, dr = 1 \\ S_2 \int_0^\infty e^{-\frac{1}{2}r^2} r^2\, dr = (2\pi)^{\frac{3}{2}} \quad &\Rightarrow \quad \int_0^\infty e^{-\frac{1}{2}r^2} r^2\, dr = \sqrt{2/\pi}\,. \end{aligned} \tag{7.22}$$

In higher dimensions

$$\begin{aligned} S_{d=even} &= 2(\pi)^{\frac{d}{2}}/(d/2-1)! \\ S_{d=odd} &= 2^{\frac{d}{2}+\frac{1}{2}}(\pi)^{\frac{d}{2}-\frac{1}{2}}/(d-2)!!. \end{aligned} \tag{7.23}$$

Substituting $d = 2k$ for even and $d = 2k+1$ for odd:

$$\begin{aligned} S_{2k-1} &= 2\pi^k/(k-1)! \\ S_{2k} &= 2(2\pi)^k/(2k-1)!!. \end{aligned} \tag{7.24}$$

(b) From (7.21):

$$S_{d-1} = \frac{(2\pi)^{\frac{d}{2}}}{\int_0^\infty e^{-r^2/2} r^{d-1}\, dr}\,. \tag{7.25}$$

Noting that $\int_0^\infty e^{-\frac{1}{2}r^2} r^{d-1}\, dr$ can be rewritten as a gamma function by substituting $\frac{1}{2}r^2$ for t:

$$\begin{aligned} \int_0^\infty e^{-\frac{1}{2}r^2} r^{d-1}\, dr &= \int_0^\infty e^{-t} \sqrt{2t}^{d}\, dt = 2^{\frac{d}{2}} \int_0^\infty e^{-t} t^{\frac{d}{2}}\, dt \\ &= 2^{\frac{d}{2}-1} \int_0^\infty e^{-t} t^{\frac{d}{2}-1}\, dt = 2^{\frac{d}{2}-1}\Gamma(d/2)\,, \end{aligned}$$

and substituting back into (7.25) we get:

$$S_{d-1} = 2\pi^{\frac{d}{2}}/\Gamma(d/2)\,. \tag{7.26}$$

(d) For $d = 2$: $S_1 = 2\pi/1 = 2\pi$, and for $d = 3$: $S_2 = 2\pi^{\frac{2}{2}}/\sqrt{\pi}/2 = 4\pi$.

*(e) **(bonus)** The Sterling's formula says that $\Gamma(x) \approx \sqrt{2\pi}\, x^{x-1/2} e^{-x}$. Applied to (7.26), it yields:*

$$\begin{aligned} S_{d\to\infty} &= \frac{2\pi^{\frac{d}{2}}}{\sqrt{2\pi}\frac{d}{2}^{\frac{d}{2}-\frac{1}{2}} e^{-\frac{d}{2}}} = \frac{\sqrt{2}\pi^{\frac{d}{2}-\frac{1}{2}}}{e^{-\frac{d}{2}}\frac{d}{2}^{\frac{d}{2}-\frac{1}{2}}} \\ &= \sqrt{2} e^{\frac{d}{2}} \left(\frac{2\pi}{d}\right)^{\frac{d}{2}-\frac{1}{2}} ????? \end{aligned} \tag{7.27}$$

See the other solution.

(f) See the other solution.

(g) See the other solution.

Arthur Lin

Solution 7.2 - Airy function for large arguments. *We start with the integral form of the Airy function:*

$$Ai(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(xy+y^3/3)}\,dy\,.$$

Rewrite the Airy function as a contour integral

$$Ai(x) = \frac{1}{2\pi i}\int_C e^{(xt-t^3/3)}\,dt,$$

where $t = iy$ is our new parameter, thus our contour goes from $-i\infty$ to $i\infty$. Apply the steepest decent method by first finding the derivatives of $f(t)$:

$$\begin{aligned} f(t) &= tx - t^3/3 \\ f'(t) &= x - t^2 \\ f''(t) &= -2t \end{aligned} \tag{7.28}$$

From the first derivative we find that the saddle points t_s are at $\pm\sqrt{x}$, which means that we have to treat two cases separately. If $x > 0$, t_s is real and if $x < 0$, t_s is purely imaginary.

(i) *If $x > 0$: our contour would only pass through one saddle point, which we choose to be $t_s = -\sqrt{x}$, since going from $-i\infty$ to $i\infty$ naturally passes through $-\sqrt{x}$ at steepest decent. Then, our values of $f(t_s)$ and its derivatives at $f(t_s)$ are:*

$$\begin{aligned} f(t_s) &= -x^{\frac{3}{2}} + x^{\frac{3}{2}}/3 = -\frac{2}{3}x^{\frac{3}{2}} \\ f'(t_s) &= 0 \\ f''(t_s) &= 2x^{\frac{1}{2}} \\ f'''(t_s) &= -2 \end{aligned} \tag{7.29}$$

After a change of variable ($t = x^{-\frac{1}{2}} + ix^{-\frac{1}{4}}\tau$), the Taylor expansion of $f(t)$ is $-\frac{2}{3}x^{\frac{3}{2}} - \tau^2 - \frac{i}{3}x^{-\frac{3}{4}}\tau^3$. We then separate terms of the exponent of $f(t)$ and Taylor expand the last term (under the condition that $x^{-\frac{3}{4}}$ is small or x is large):

$$\begin{aligned} Ai(x) &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} e^{(-\frac{2}{3}x^{\frac{3}{2}})}e^{-\tau^2}e^{(-\frac{i}{3}x^{-\frac{3}{4}}\tau^3)}ix^{-\frac{1}{4}}\,d\tau \\ &= \frac{e^{(-\frac{2}{3}x^{\frac{3}{2}})}}{2\pi x^{\frac{1}{4}}}\int_{-\infty}^{\infty} e^{-\tau^2}(1 - \frac{i}{3}x^{-\frac{3}{4}}\tau^3 - \frac{1}{18}x^{-\frac{3}{2}}\tau^6 + \cdots)\,d\tau \end{aligned}$$

All the odd terms in the Taylor expansion equal zero, and the integral becomes:

$$Ai(x>0) = \frac{e^{(-\frac{2}{3}x^{\frac{3}{2}})}}{2\pi^{\frac{1}{2}}x^{\frac{1}{4}}}(1 - \frac{5}{48}x^{-\frac{3}{2}} + \cdots)$$

(ii) If $x < 0$: $t_s = \pm\sqrt{x} = \pm i\sqrt{-x}$. The values of $f(t_s)$ and its derivatives at $f(t_s)$ are:

$$\begin{aligned} f(t_s) &= \pm i\frac{2}{3}(-x)^{\frac{3}{2}} \\ f'(t_s) &= 0 \\ f''(t_s) &= \mp 2i(-x)^{\frac{1}{2}} \\ f'''(t_s) &= -2 \end{aligned} \tag{7.30}$$

The Taylor expansion is now

$$f(x) = \pm i\frac{2}{3}(-x)^{\frac{3}{2}} \mp 2i(-x)^{\frac{1}{2}}(t \mp i\sqrt{-x})^2 - 2(t \mp i\sqrt{-x})^3 \,.$$

Following the same steps as the case $x > 0$, we arrive at the first order result:

$$Ai(x<0)_{t_{s_\pm}} = \frac{e^{\pm i\frac{2}{3}(-x)^{\frac{3}{2}}}}{2\pi\sqrt{\pm i\sqrt{-x}}}\int_C e^{-t^2}\,dt = \frac{e^{\pm i\frac{2}{3}(-x)^{\frac{3}{2}}}}{2\sqrt{\pi}\sqrt{\pm i\sqrt{-x}}}\,,$$

for each saddle point. The final value of the integral is the sum of the two saddle points:

$$Ai(x<0) = \frac{e^{i\frac{2}{3}(-x)^{\frac{3}{2}}}}{2\sqrt{\pi}\sqrt{i\sqrt{-x}}} + \frac{e^{-i\frac{2}{3}(-x)^{\frac{3}{2}}}}{2\sqrt{\pi}\sqrt{-i\sqrt{-x}}} = \frac{\sin(\frac{2}{3}(-x)^{\frac{3}{2}} + \frac{\pi}{4})}{\sqrt{\pi}(-x)^{\frac{1}{4}}}\,.$$

Conclusion: As we can see from the derivation, our approximation depends on $|x| \to \infty$, and if $x = 0$ we get the value of ∞ for the integrals.

Arthur Lin

mathematical methods - week 8

Discrete Fourier representation

Georgia Tech PHYS-6124

Homework HW #8 due Thursday, October 15, 2020

== 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 8.1 *Laplacian is a non-local operator* 4 points
Exercise 8.2 *Lattice Laplacian diagonalized* 8 points

Total of 12 points = 100 % score.

edited October 13, 2020

Week 8 syllabus

Tuesday, October 5, 2020

Discretization of continuum, lattices, discrete derivatives, discrete Fourier representations.

The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- *Symmetry is your friend - overview. The power of thinking.*
- Applied math version: how to discretize derivatives:
  ChaosBook Appendix A24 Deterministic diffusion
  Sects. A24.1 to A24.1.1 *Lattice Laplacian*.
- *Lattice discretization, lattice state*
- *Lattice derivative*
- *Shift operator: the generator of discrete translations*
  - *Discussion: Shift matrix must have the periodic b.c.; Derivative being non-local is easiest to grasp on discrete lattice. It's so easy to make errors in the continuum formulation.*
- *Derivative is a linear operator*
- *Lattice Laplacian*
- *Derivative is a non-local operator*
  - *Discussion: Lattice discretization; What if geometry is not flat in all directions, but spherical? What about General Relativity? Life's persistent questions, skated around.*
  - *Discussion: What is a derivative? Hypercubic lattice is a graph, with nodes connected by links. Every graph has a notion of derivative associated with it; in particular a Laplacian. I was not allowed to say "Laplacian" here, as I have not gotten to defining it in my lecture at that point...*
- A periodic lattice as the simplest example of the theory of finite groups:
  ChaosBook Sects. A24.1.2 to A24.3.1.
  ChaosBook Example A24.2 *Projection operators for discrete Fourier representation.*
  ChaosBook Example A24.3 *'Configuration-momentum' Fourier space duality.*
- *Have symmetry? Use it!*
  - *Rant: Symmetrize you must. Karl Schwarzschild found his exact solution in 1915, a month after the publication of Einstein's theory of general relativity, while serving on a World War I front.*

- *Have symmetry? Go to "eigen"subspace! Fourier decomposition of a 2-sites periodic lattice.*
- *Periodic lattices*
- *Fourier eigenvalues*
- *Discrete Fourier representation*
- *Laplacian in Fourier representation*
- *Propagator in Fourier representation*
- *A meta truth; We live in The Matrix; Fourier transformation is just a matrix*

Optional reading

- A theoretical physicist's version of the above notes: *Quantum Field Theory - a cyclist tour*, Chapter 1 *Lattice field theory* motivates discrete Fourier representations by computing a free propagator on a lattice.
- *Quantum Mechanics in a box: Sometimes it is simplest to impose the periodic b.c. on a localized solution, than relax it towards the correct (infinite extent) continuum solution.*
- *Rocket science needs complex numbers; Why Fourier? Digital image processing!*

Exercises

8.1. **Laplacian is a non-local operator.**
While the Laplacian is a simple tri-diagonal difference operator, its inverse (the "free" propagator of statistical mechanics and quantum field theory) is a messier object. A way to compute is to start expanding propagator as a power series in the Laplacian

$$\frac{1}{m^2\mathbf{1}-\Delta} = \frac{1}{m^2}\sum_{n=0}^{\infty}\frac{1}{m^{2n}}\Delta^n \,. \tag{8.1}$$

As Δ is a finite matrix, the expansion is convergent for sufficiently large m^2. To get a feeling for what is involved in evaluating such series, show that Δ^2 is:

$$\Delta^2 = \frac{1}{a^4}\begin{bmatrix} 6 & -4 & 1 & & & 1 & -4 \\ -4 & 6 & -4 & 1 & & & \\ 1 & -4 & 6 & -4 & 1 & & \\ & 1 & -4 & \ddots & & & \\ & & & & & 6 & -4 \\ -4 & 1 & & & 1 & -4 & 6 \end{bmatrix} . \tag{8.2}$$

What Δ^3, Δ^4, $\cdots$ contributions look like is now clear; as we include higher and higher powers of the Laplacian, the propagator matrix fills up; while the *inverse* propagator is differential operator connecting only the nearest neighbors, the propagator is integral operator, connecting every lattice site to any other lattice site.

This matrix can be evaluated as is, on the lattice, and sometime it is evaluated this way, but in case at hand a wonderful simplification follows from the observation that the lattice action is translationally invariant, exercise 8.2.

8.2. **Lattice Laplacian diagonalized.** Insert the identity $\sum P^{(k)} = \mathbf{1}$ wherever you profitably can, and use the shift matrix eigenvalue equation to convert shift σ matrices into scalars. If $\mathbf{M}$ *commutes* with σ, then $(\varphi_k^\dagger \cdot \mathbf{M} \cdot \varphi_{k'}) = \tilde{M}^{(k)}\delta_{kk'}$, and the matrix $\mathbf{M}$ acts as a multiplication by the scalar $\tilde{M}^{(k)}$ on the kth subspace. Show that for the 1-dimensional lattice, the projection on the kth subspace is

$$(\varphi_k^\dagger \cdot \Delta \cdot \varphi_{k'}) = \frac{2}{a^2}\left(\cos\left(\frac{2\pi}{N}k\right) - 1\right)\delta_{kk'} \,. \tag{8.3}$$

In the kth subspace the propagator is simply a number, and, in contrast to the mess generated by (8.1), there is nothing to evaluating it:

$$\varphi_k^\dagger \cdot \frac{1}{m^2\mathbf{1}-\Delta} \cdot \varphi_{k'} = \frac{\delta_{kk'}}{m^2 - \frac{2}{(ma)^2}\left(\cos 2\pi k/N - 1\right)} \,, \tag{8.4}$$

where k is a site in the N-dimensional dual lattice, and $a = L/N$ is the lattice spacing.

Chapter 8 solutions: Discrete Fourier representation

Solution 8.1 - Laplacian is a non-local operator.
See Chapter 1 of `http://ChaosBook.org/FieldTheory/QMlectures/lectQM.pdf`

Solution 8.2 - Lattice Laplacian diagonalized.
See Chapter 1 of `http://ChaosBook.org/FieldTheory/QMlectures/lectQM.pdf`

mathematical methods - week 9

Fourier transform

Georgia Tech PHYS-6124

Homework HW #9 due Thursday, October 22, 2020

== 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 9.2 *d-dimensional Gaussian integrals* 5 points
Exercise 9.3 *Convolution of Gaussians* 5 points

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

edited November 2, 2020

Week 9 syllabus

Tuesday, October 13, 2020

> There is only one thing which interests me vitally now, and that is the recording of all that which is omitted in books. Nobody, as far as I can see, is making use of those elements in the air which give direction and motivation to our lives.
>
> — Henry Miller, *Tropic of Cancer*

This week's lectures are related to AWH Chapter 19 *Fourier Series* (click here), but I prefer Stone and Goldbart [3] (click here) Appendix B exposition, which I follow closely in the online recorded lectures.

The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- ChaosBook Sect. A24.4 *Continuum field theory*: Fourier transform as the limit of a discrete Fourier transform.
  - *Propagator in continuum limit*
- Stone and Goldbart (click here) Appendix B.1 *Fourier Series*
  - *Fourier representation, circular Kronecker delta, take #2*
  - *Fourier series*
  - *Circular Dirac delta function*
- Stone and Goldbart (click here) Appendix B.2 *Fourier integral transforms*
  - *Fourier integral transform*
  - *Persival identity*
  - *Fourier transform of a Gaussian*
    - Exercise 9.3 *Convolution of Gaussians*
  - *Convolution of Gaussians*
  - *Covariance evolution*
    - Cigar is sometimes just a cigar
      - sect. 9.2 *A bit of noise.*
    - Noise : seminars and papers
  - *The fearful power of symmetry - translational invariance*
    - example 9.1 *Circulant matrices.*
    - example 9.2 *Convolution theorem for matrices.*

Optional reading

- *Rant: Introductory physics is the greatest con ever. Or,* *Brush up your Gaussians, And they'll all kow-tow*

Discussion: Verbotten! We will not prove Reimann Hypothesis, nor will we explain Wiles proof of Fermat Conjecture in this course. No other course offers intuition. You do not know how lucky you are, boy. You could be back in US back in US back in USSR. As a rule, I do not approve of abuse of children, but Prof. Z is for your own good. Learning from your mistakes is the only way to learn. Countable infinity of professorial opinions. Getting a beating from a class in uprising.

Farazmand notes on Fourier transforms.

Grigoriev notes
4. Integral transforms, 4.3-4.4 square wave, Gibbs phenomenon;
5. Fourier transform: 5.1-5.6 inverse, Parseval's identity, ..., examples

Roger Penrose [2] (click here) chapter on Fourier transforms is sophisticated, but too pretty to pass up.

Alex Kontorovich, on the history of Fourier series: As often happens in mathematics, Fourier was trying to do something completely unrelated when he stumbled on Fourier series. What was it? He was studying the propagation of heat in a uniform medium.

Bernard Maurey, *Fourier, One Man, Several Lives* (2019).

Question 9.1. Henriette Roux asks
Q You usually explain operations by finite-matrix examples, but in exercise 9.3 you asked us to show that the Fourier transform of the convolution corresponds to the product of the Fourier transforms only for continuum integrals. The exercise gives me no intuition for what a convolution *is*.
A "Convolution" is a matrix multiplication for translationally invariant matrix operators. For what that is for discrete Fourier transforms, and what is a "convolution theorem" for matrices, see example 9.2 and *The fearful power of symmetry - translational invariance.*

9.1 Examples

Example 9.1. ***Circulant matrices.*** *An $[L \times L]$ circulant matrix*

$$C = \begin{bmatrix} c_0 & c_{L-1} & \dots & c_2 & c_1 \\ c_1 & c_0 & c_{L-1} & & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{L-2} & & \ddots & \ddots & c_{L-1} \\ c_{L-1} & c_{L-2} & \dots & c_1 & c_0 \end{bmatrix}, \tag{9.1}$$

has eigenvectors (discrete Fourier modes) and eigenvalues $Cv_k = \lambda_k v_k$

$$\begin{aligned} v_k &= \frac{1}{\sqrt{L}}(1, \omega^k, \omega^{2k}, \dots, \omega^{k(L-1)})^{\mathrm{T}}, \qquad k = 0, 1, \dots, L-1 \\ \lambda_k &= c_0 + c_{L-1}\omega^k + c_{L-2}\omega^{2k} + \dots + c_1\omega^{k(L-1)}, \end{aligned} \tag{9.2}$$

where

$$\omega = e^{2\pi \mathrm{i}/L} \tag{9.3}$$

is a root of unity. The familiar examples are the one-lattice site shift matrix ($c_1 = 1$, all other $c_k = 0$), and the lattice Laplacian $\Box$.

Example 9.2. ***Convolution theorem for matrices.*** *Translation-invariant matrices can only depend on differences of lattice positions,*

$$C_{ij} = C_{i-j,0} \tag{9.4}$$

All content of a translation-invariant matrix is thus in its first row C_{n0}, all other rows are its cyclic translations, so translation-invariant matrices are always of the circulant form (9.1). A product of two translation-invariant matrices can be written as

$$A_{im} = \sum_j B_{ij} C_{jm} \quad \Rightarrow \quad A_{i-m,0} = \sum_j B_{i-j,0}\, C_{j-m,0}\,,$$

i.e., in the "convolution" form

$$A_{n0} = (BC)_{n0} = \sum_\ell B_{n-\ell,0}\, C_{\ell 0} \tag{9.5}$$

which only uses a single row of each matrix; N operations, rather than the matrix multiplication N^2 operations for each of the N components A_{n0}.

A circulant matrix is constructed from powers of the shift matrix, so it is diagonalized by the discrete Fourier transform, i.e., unitary matrix U. In the Fourier representation, the convolution is thus simply a product of kth Fourier components (no sum over k):

$$UAU^\dagger = UBU^\dagger UCU^\dagger \quad \to \quad \tilde{A}_{kk} = \tilde{B}_{kk} \tilde{C}_{kk}\,. \tag{9.6}$$

That requires only 1 multiplication for each of the N components A_{n0}.

9.2 A bit of noise

Fourier invented Fourier transforms to describe the diffusion of heat. How does that come about?

Consider a noisy *discrete time* trajectory

$$x_{n+1} = x_n + \xi_n\,, \qquad x_0 = 0\,, \tag{9.7}$$

where x_n is a d-dimensional state vector at time n, $x_{n,j}$ is its jth component, and ξ_n is a noisy kick at time n, with the prescribed probability distribution of zero mean and the covariance matrix (diffusion tensor) Δ,

$$\langle \xi_{n,j} \rangle = 0\,, \qquad \langle \xi_{n,i}\, \xi^T_{m,j} \rangle = \Delta_{ij}\, \delta_{nm}\,, \tag{9.8}$$

where $\langle \cdots \rangle$ stands for average over many realizations of the noise. Each 'Langevin' trajectory $(x_0, x_1, x_2, \cdots)$ is an example of a Brownian motion, or diffusion.

In the Fokker-Planck description individual noisy trajectories (9.7) are replaced by the evolution of a density of noisy trajectories, with the action of discrete one-time step *Fokker-Planck operator* on the density distribution ρ at time n,

$$\rho_{n+1}(y) = [\mathcal{L}\rho_n](y) \;=\; \int dx\, \mathcal{L}(y,x)\, \rho_n(x)\,, \tag{9.9}$$

given by a normalized Gaussian (work through exercise 9.2)

$$\mathcal{L}(y,x) = \frac{1}{N}\, e^{-\frac{1}{2}(y-x)^T \frac{1}{\Delta}(y-x)}\,, \quad N = (2\pi)^{d/2}\sqrt{\det(\Delta)}\,, \tag{9.10}$$

which smears out the initial density ρ_n diffusively by noise of covariance (9.8). The covariance Δ is a symmetric $[d\times d]$ matrix which can be diagonalized by an orthogonal transformation, and rotated into an ellipsoid with d orthogonal axes, of different widths (covariances) along each axis. You can visualise the Fokker-Planck operator (9.9) as taking a δ-function concentrated initial distribution centered on $x = 0$, and smearing it into a cigar shaped noise cloud.

As $\mathcal{L}(y,x) = \mathcal{L}(y-x)$, the Fokker-Planck operator acts on the initial distribution as a *convolution*,

$$[\mathcal{L}\rho_n](y) \;=\; [\mathcal{L} * \rho_n](y) \;=\; \int dx\, \mathcal{L}(y-x)\, \rho_n(x)$$

Consider the action of the Fokker-Planck operator on a normalized, cigar-shaped Gaussian density distribution

$$\rho_n(x) = \frac{1}{N_n} e^{-\frac{1}{2}x^T \frac{1}{\Delta_n}\, x}\,, \qquad N_n = (2\pi)^{d/2}\sqrt{\det(\Delta_n)}\,. \tag{9.11}$$

That is also a cigar, but in general of a different shape and orientation than the Fokker-Planck operator (9.10). As you can check by working out exercise 9.3, a convolution of a Gaussian with a Gaussian is again a Gaussian, so the Fokker-Planck operator maps the Gaussian $\rho_n(x_n)$ into the Gaussian

$$\rho_{n+1}(x) = \frac{1}{N_{n+1}} e^{-\frac{1}{2}x^T \frac{1}{\Delta_n+\Delta}\, x}\,, \qquad N_{n+1} = (2\pi)^{d/2}\sqrt{\det(\Delta_n+\Delta)} \tag{9.12}$$

one time step later.

In other words, covariances Δ_n add up. This is the d-dimensional statement of the familiar fact that cumulative error squared is the sum of squares of individual errors. When individual errors are small, and you are adding up a sequence of them in time, you get Brownian motion. If the individual errors are small and added independently to a solution of deterministic equations (so-called 'drift'), you get the Langevin and the Fokker-Planck equations.

References

[1] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals* (Dover, New York, 1986).

[2] R. Penrose, *The Road to Reality - A Complete Guide to the Laws of the Universe* (A. A. Knopf, New York, 2005).

[3] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

Exercises

9.1. **Who ordered $\sqrt{\pi}$?** Derive the Gaussian integral

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, e^{-\frac{x^2}{2a}} = \sqrt{a}\,, \qquad a>0\,.$$

assuming only that you know to integrate the exponential function e^{-x}. Hint, hint: x^2 is a radius-squared of something. π is related to the area or circumference of something.

9.2. **d-dimensional Gaussian integrals.** Show that the Gaussian integral in d-dimensions is given by

$$\begin{aligned} Z[J] &= \int d^d x\, e^{-\frac{1}{2}x^\top \cdot M^{-1}\cdot x + x^\top\cdot J} \\ &= (2\pi)^{d/2}|\det M|^{\frac{1}{2}}\, e^{\frac{1}{2}J^\top\cdot M\cdot J}\,, \end{aligned} \tag{9.13}$$

where M is a real positive definite $[d\times d]$ matrix, i.e., a matrix with strictly positive eigenvalues, x and J are d-dimensional vectors, and $(\cdots)^\top$ denotes the transpose.

This integral you will see over and over in statistical mechanics and quantum field theory: it's called 'free field theory', 'Gaussian model', 'Wick expansion', etc.. This is the starting, 'propagator' term in any perturbation expansion.

Here we require that the real symmetric matrix M in the exponent is strictly positive definite, otherwise the integral is infinite. Negative eigenvalues can be accommodated by taking a contour in the complex plane [1], see exercise 6.3 *Fresnel integral*. Zero eigenvalues require stationary phase approximations that go beyond the Gaussian saddle point approximation, typically to the Airy-function type stationary points, see exercise 7.2 *Airy function for large arguments*.

9.3. **Convolution of Gaussians.**

(a) Show that the Fourier transform of the convolution

$$[f*g](x) = \int d^d y\, f(x-y)g(y)$$

corresponds to the product of the Fourier transforms

$$[f*g](x) = \frac{1}{(2\pi)^d}\int d^d k\, F(k)G(k)e^{+ik\cdot x}\,, \tag{9.14}$$

where

$$F(k) = \int \frac{d^d x}{(2\pi)^{d/2}}\, f(x)\, e^{-ik\cdot x}\,, \quad G(k) = \int \frac{d^d x}{(2\pi)^{d/2}}\, g(x)\, e^{-ik\cdot x}\,.$$

(b) Consider two normalized Gaussians

$$\begin{aligned} f(x) &= \frac{1}{N_1}\, e^{-\frac{1}{2}x^\top\cdot\frac{1}{\Delta_1}\cdot x}\,, \quad N_1 = \sqrt{\det(2\pi\Delta_1)} \\ g(x) &= \frac{1}{N_2}\, e^{-\frac{1}{2}x^\top\cdot\frac{1}{\Delta_2}\cdot x}\,, \quad N_2 = \sqrt{\det(2\pi\Delta_2)} \\ 1 &= \int d^d k\, f(x) = \int d^d k\, g(x)\,. \end{aligned}$$

Evaluate their Fourier transforms

$$F(k) = \frac{1}{(2\pi)^{d/2}} e^{\frac{1}{2} k^\top \cdot \Delta_1 \cdot k} \,, \qquad G(k) = \frac{1}{(2\pi)^{d/2}} e^{\frac{1}{2} k^\top \cdot \Delta_2 \cdot k} \,.$$

Show that the convolution of two normalized Gaussians is a normalized Gaussian

$$[f * g](x) \quad = \quad \frac{(2\pi)^{-d/2}}{\sqrt{\det\left(\Delta_1 + \Delta_2\right)}} \, e^{-\frac{1}{2} x^\top \cdot \frac{1}{\Delta_1 + \Delta_2} \cdot x} \,.$$

In other words, covariances Δ_j add up. This is the d-dimenional statement of the familiar fact that cumulative error squared is the sum of squares of individual errors. When individual errors are small, and you are adding up a sequence of them in time, you get Brownian motion. If the individual errors are small and added independently to a solution of a deterministic equation, you get Langevin and Fokker-Planck equations.

Chapter 9 solutions: Fourier transform

Solution 9.1 - Who ordered $\sqrt{\pi}$? *No solution available.*

Solution 9.2 - d-dimensional Gaussian integrals. *Make a change of variables $y = Ax$ such that $A^T M^{-1} A = \mathrm{Id}$. Then*

$$I = \frac{1}{(2\pi)^{d/2}} \int dy^d \exp[-\frac{1}{2}\sum_i (y_i^2 - 2(JA)_i y_i)]|\det A|$$

Complete each term under in the sum in the exponent to a full square

$$y_i^2 - 2(JA)_i y_i = (y_i - (JA)_i)^2 - (JA)_i^2$$

and shift the origin of integration to $JA/2$, so that

$$I = \frac{1}{(2\pi)^{d/2}} \exp(\frac{1}{2}J^T AA^T J)|\det A| \int dy^d \exp[-\frac{1}{2}\sum_i y_i^2]\,.$$

Note that $AA^T M^{-1} AA^T = AA^T$, therefore $AA^T = M$ and $|\det A| = \sqrt{\det M}$. The remaining integral is equal to a Gaussian integral raised to the d-th power, i.e., $(2\pi)^{d/2}$. Hence:

$$I = \sqrt{\det M} \exp[\frac{1}{2}J^T M J]$$

(R. Paškauskas)

Solution 9.2 - d-dimensional Gaussian integrals. *We require that the matrix in the exponent is nondegenerate (i.e. has no zero eigenvalues.) The converse may happen when doing stationary phase approximations which requires going beyond the Gaussian saddle point approximation, typically to the Airy-function type stationary points [1]. We also assume that M is positive-definite, otherwise the integral is infinite.*

Make a change of variables $y = Ax$ such that $A^T M^{-1} A = \mathrm{Id}$. Then

$$I = \frac{1}{(2\pi)^{d/2}} \int_{\mathbf{R}^d} \exp[-\frac{1}{2}\sum_i (y_i^2 - 2(JA)_i y_i)]|\det A| dy$$

Complete each term under in the sum in the exponent to a full square

$$y_i^2 - 2(JA)_i y_i = (y_i - (JA)_i)^2 - (JA)_i^2$$

and shift the origin of integration to $JA/2$, so that

$$I = \frac{1}{(2\pi)^{d/2}} \exp(\frac{1}{2}J^T AA^T J)|\det A| \int_{\mathbf{R}^d} \exp[-\frac{1}{2}\sum_i y_i^2] dy$$

Note that $AA^T M^{-1} AA^T = AA^T$, therefore $AA^T = M$ and $|\det A| = \sqrt{\det M}$. The remaining integral is equal to a Poisson integral raised to the d-th power, i.e. $(2\pi)^{d/2}$. Answer:

$$I = \sqrt{\det M} \exp[\frac{1}{2}J^T M J]$$

(R. Paškauskas)

Solution 9.3 - Errors add up as sums of squares. *In one dimension the Fourier transform of f over $x \in \mathbb{R}$ is given by*

$$\mathcal{F}[f](k) = \frac{1}{\sqrt{2\pi}} \int_\Omega dx\, f(x)\, e^{-ikx} \tag{9.15}$$

and the transform of the convolution $f * g$ is

$$\begin{aligned}\mathcal{F}[f*g] &= (2\pi)^{-d/2}\int d^dx \int d^dy f(x-y)g(y)\, e^{-ikx} \\ &= (2\pi)^{-d/2}\int d^dx \int d^dy f(x-y)g(y)\, e^{-ik(x-y)}\, e^{-iky} \\ &= (2\pi)^{-d/2}\int d^dq\, e^{-ikq} f(q) \int d^dy\, e^{-iky} g(y) \\ &= (2\pi)^{-d/2}(2\pi)^{+d/2}F(k)\,(2\pi)^{+d/2}G(k)\end{aligned}$$

$F(k)$ *and* $G(k)$ *follow from* (9.15)*:*

$$F(k) = (2\pi)^{-d/2}\int d^dx\, f(x)\, e^{-ikx} = (2\pi)^{-d/2}\sqrt{\det(\Delta_1)}\, e^{k^\dagger \Delta_1 k/2}$$

$$\Rightarrow F(k) = \sqrt{\det(\Delta_1)}\, e^{k^\dagger \Delta_1 k/2}\,, \qquad \Rightarrow G(k) = \sqrt{\det(\Delta_2)}\, e^{k^\dagger \Delta_2 k/2}$$

Substituting $F(k)$ *and* $G(k)$ *into the inverse transform of* $\mathcal{F}[f*g]$*:*

$$f*g \equiv (2\pi)^{-d/2}\sqrt{\det(\Delta_1)\det(\Delta_2)}\int dk\, e^{k^\dagger\Delta_1 k/2 + k^\dagger\Delta_2 k/2 + ikx}$$

$$f*g = \sqrt{\frac{\det(\Delta_1)\det(\Delta_2)}{\det(\Delta_1+\Delta_2)}}\, e^{-x^\dagger(\Delta_1+\Delta_2)^{-1}x/2}$$

(Chris Marcotte)

mathematical methods - week 10

Discrete symmetries

Georgia Tech PHYS-6124

Homework HW #10 due Thursday, October 29, 2020

== 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 10.1 *1-dimensional representation of anything*	1 point
Exercise 10.2 *2-dimensional representation of* S_3	4 points
Exercise 10.3 D_3*: symmetries of an equilateral triangle*	5 points

Bonus points

Exercise 10.4 (a), (b) and (c) *Permutation of three objects*	2 points
Exercise 10.5 *3-dimensional representations of* D_3	3 points

Total of 10 points = 100 % score.

edited October 28, 2020

Week 10 syllabus

Tuesday, October 20, 2020

Tyger Tyger burning bright,
In the forests of the night:
What immortal hand or eye,
Dare frame thy fearful symmetry?
—William Blake, *The Tyger*

This week's lectures are related to AWH Chapter 17 *Group Theory* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

There is way too much material in this week's notes. Watch the main sequence of video clips, that and recommended reading should suffice to do the problems. The rest is optional. You can glance through sect. 10.1 *Group presentations*, and sect. 10.3 *Literature*, but I do not expect you to understand this material.

Group theory and why I love 808,017,$\cdots$,000 is a great video on group theory from 3Blue1Brown, writes Andrew Wu. I agree: Well worth of your time, more motivational than my lectures. What it actually focuses on - the monster group - is totally useless to us. My focus this week is narrow and technical:

1. theory of finite groups are a natural generalization of discrete Fourier representations
2. it is all about class and character. "Character", in particular, I find very surprising - one complex number suffices to characterize a matrix!

Hang in there! And relax. None of this will be on the test. As a matter of fact, there will be no test.

- It's all about class: Groups, permutations, $D_3 \cong C_{3v} \cong S_3$ symmetries of equilateral triangle, rearrangement theorem, subgroups, cosets, classes.

  Dresselhaus *et al.* [3] Chapter 1 *Basic Mathematical Background: Introduction* (click here). The MIT course 6.734 online version contains much of the same material.

  ChaosBook Chapter 10. Flips, slides and turns

  Clip 1 - discrete symmetry, an example: 3-disk pinball

  Clip 2 - what is a group?

  Clip 2a - discussion : permutations, symmetric group, simple groups, Italian renaissance, French revolution, Galois

  by Socratica*:*
  a delightful introduction to group multiplication (or Cayley) tables.

  Clip 3 - active, passive coordinate transformations

Clip 4 - following Mefisto: symmetry defined three (3) times

Clip 5 - subgroups, classes, group orbits, reduced state space

- Hard work builds character: Irreps, unitary reps, Schur's Lemma.

  Chapter 2 *Representation Theory and Basic Theorems*
  Dresselhaus *et al.* [3], up to and including
  Sect. 2.4 *The Unitarity of Representations* (click here)

  Clip 6 - this requires character

  Clip 7 - hard work builds character

  Clip 8 - the symmetry group of a propeller

  Clip 9 - irreps of C_3

  Clip 10 - rotation in the plane: irreps of D_3

  Clip 10a - Discussion : more symmetries, fewer invariant subspaces

  Clip 10b - Discussion : abelian vs. nonabelian

- "Wonderful Orthogonality Theorem."

  In this course, we learn about full reducibility of finite and compact continuous groups in two parallel ways. On one hand, I personally find the multiplicative *projection operators* (1.19), coupled with the notion of class algebras (Harter [4] (click here) appendix C) most intuitive - a block-diagonalized subspace for each distinct eigenvalue of a given all-commuting matrix. On the other hand, the character weighted sums (here related to the multiplicative projection operators as in ChaosBook Example A24.2 *Projection operators for discrete Fourier transform*) offer a deceptively 'simple' and elegant formulation of full-reducibility theorems, preferred by all standard textbook expositions:

  Dresselhaus *et al.* [3] Sects. 2.5 and 2.6 *Schur's Lemma.*
  a first go at sect. 2.7

  Clip 11 - irreps

  Clip 12 - Frobenius character formula

  Clip 13 - character orthogonality relations

  Clip 14 - the summary: it is all about class and character

  Clip 14a - discussion : class and character

Optional reading

For a deep dive into this material, here is your rabbit hole.

For deeper insights, read Roger Penrose [7] (click here).

- For a typical (but for this course advanced) application see, for example, Stone and Goldbart [10], *Mathematics for Physics: A Guided Tour for Graduate Students*, Section 14.3.2 *Vibrational spectrum of H_2O* (click here).

- Harter's Sect. *3.2 First stage of non-Abelian symmetry analysis*
group multiplication table (3.1.1); class operators; class multiplication table (3.2.1b); all-commuting or central operators;

- Harter's Sect. *3.3 Second stage of non-Abelian symmetry analysis*
projection operators (3.2.15); 1-dimensional irreps (3.3.6); 2-dimensional irrep (3.3.7); Lagrange irreps dimensionality relation (3.3.17)

- *An example: a 1-dimensional system with a symmetry*

- *Fundamental domain*

- *Tiling of state space by a finite group*

- *Make the "fundamental tile" your hood*

- *Symmetry-reduced dynamics*

- *Regular representation of permuting tiles*

- *Group theory voodoo*

- *Tell no Lie to plumbers*

- *Week 12 Clip 3 - Birdtracks (6 min)*

    - Sect. 10.1.1 *Permutations in birdtracks*

- *Discussion 1 - There might be many examples of it, but a 'group' itself is an abstract notion. (3 min)*

- *Discussion 2 - Fourier modes are so simple, that no one calls them irreps. But add more symmetries, and there have to be fewer irreps. (11 min)*

- *Discussion 3 - what are these "characters"? And why is there a* Journal of Linear Algebra*, today? Inconclusive blah blah. (12 min)*

- *Discussion 4 - Homework. (3 min)*

- *It's a matter of no small pride for a card-carrying dirt physics theorist to claim* full and total ignorance of group theory

- ChaosBook Appendix A.6 Gruppenpest

- There is no need to learn all these "Greek" words.

D_3	e	C	C^2	$\sigma^{(1)}$	$\sigma^{(2)}$	$\sigma^{(3)}$
e	e	C	C^2	$\sigma^{(1)}$	$\sigma^{(2)}$	$\sigma^{(3)}$
C	C	C^2	e	$\sigma^{(3)}$	$\sigma^{(1)}$	$\sigma^{(2)}$
C^2	C^2	e	C	$\sigma^{(2)}$	$\sigma^{(3)}$	$\sigma^{(1)}$
$\sigma^{(1)}$	$\sigma^{(1)}$	$\sigma^{(2)}$	$\sigma^{(3)}$	e	C	C^2
$\sigma^{(2)}$	$\sigma^{(2)}$	$\sigma^{(3)}$	$\sigma^{(1)}$	C^2	e	C
$\sigma^{(3)}$	$\sigma^{(3)}$	$\sigma^{(1)}$	$\sigma^{(2)}$	C	C^2	e

Table 10.1: The D_3 group multiplication table.

Question 10.1. Henriette Roux asks
Q What are cosets good for?
A Apologies for glossing over their meaning in the lecture. I try to minimize group-theory jargon, but cosets cannot be ignored.

Dresselhaus *et al.* [3] (click here) Chapter 1 *Basic Mathematical Background: Introduction* needs them to show that the dimension of a subgroup is a divisor of the dimension of the group. For example, C_3 of dimension 3 is a subgroup of D_3 of dimension 6.

In ChaosBook *Chapter 10. Flips, slides and turns* cosets are absolutely essential. The significance of the coset is that if a solution has a symmetry, then the elements in a coset act on the solution the same way, and generate all equivalent copies of this solution. Example 10.7. *Subgroups, cosets of* D_3 should help you understand that.

10.1 Group presentations

Group theory? It is all about class & character.
— Predrag Cvitanović, *One minute elevator pitch*

Group multiplication (or Cayley) tables, such as Table 10.1, *define* each distinct discrete group, but they can be hard to digest. A Cayley graph, with links labeled by generators, and the vertices corresponding to the group elements, has the same information as the group multiplication table, but is often a more insightful presentation of the group.

For example, the Cayley graph figure 10.1 is a clear presentation of the dihedral group D_4 of order 8,

$$D_4 = (e, a, a^2, a^3, b, ba, ba^2, ba^3)\,, \quad \text{generators } a^4 = e\,,\ b^2 = e\,. \tag{10.1}$$

Quaternion group is also of order 8, but with a distinct multiplication table / Cayley graph, see figure 10.2. For more of such, see, for example, mathoverflow Cayley graph discussion.

Example 10.1. ***Projection operators for cyclic group*** C_N**.**
Consider a cyclic group $C_N = \{e, g, g^2, \cdots g^{N-1}\}$*, and let* $M = D(g)$ *be a* $[2N\times 2N]$ *representation of the one-step shift* g*. In the projection operator formulation (1.19), the* N *distinct eigenvalues of* M*, the* N*th roots of unity* $\lambda_n = \lambda^n$*,* $\lambda = \exp(i\,2\pi/N)$*,*

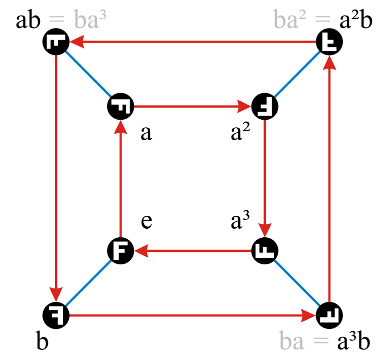


Figure 10.1: A Cayley graph presentation of the dihedral group D_4. The 'root vertex' of the graph, marked e, is here indicated by the letter $\mathbb{F}$, the links are multiplications by two generators: a cyclic rotation by left-multiplication by element a (directed red link), and the flip by b (undirected blue link). The vertices are the 8 possible orientations of the transformed letter $\mathbb{F}$.

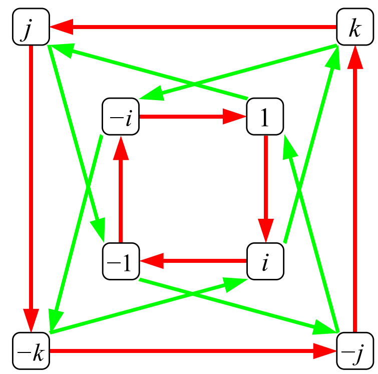


Figure 10.2: A Cayley graph presentation of the quaternion group Q_8. It is also of order 8, but distinct from D_4.

$n = 0, \dots N-1$, split the $2N$-dimensional space into N 2-dimensional subspaces by means of projection operators

$$P_n = \prod_{m \neq n} \frac{M - \lambda_m I}{\lambda_n - \lambda_m} = \prod_{m=1}^{N-1} \frac{\lambda^{-n} M - \lambda^m I}{1 - \lambda^m}\,, \tag{10.2}$$

where we have multiplied all denominators and numerators by λ^{-n}. The numerator is now a matrix polynomial of form $(x-\lambda)(x-\lambda^2)\cdots(x-\lambda^{N-1})$, with the zeroth root $(x-\lambda^0) = (x-1)$ quotiented out from the defining matrix equation $M^N - 1 = 0$. Using

$$\frac{1-x^N}{1-x} = 1 + x + \cdots + x^{N-1} = (x-\lambda)(x-\lambda^2)\cdots(x-\lambda^{N-1})$$

we obtain the projection operator in form of a discrete Fourier sum (rather than the product (1.19)),

$$P_n = \frac{1}{N} \sum_{m=0}^{N-1} e^{i\,\frac{2\pi}{N} nm}\, M^m\,.$$

This form of the projection operator is the simplest example of the key group theory tool, projection operator expressed as a sum over characters,

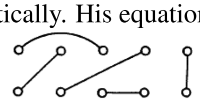
$$P_n = \frac{1}{|G|} \sum_{g \in G} \bar{\chi}(g) D(g)\,.$$

(B. Gutkin and P. Cvitanović)

10.1.1 Permutations in birdtracks

The text that follows is a very condensed extract of chapter 6 *Permutations* from *Group Theory - Birdtracks, Lie's, and Exceptional Groups* [2]. I am usually reluctant to use birdtrack notations in front of graduate students indoctrinated by their professors in the 1890's tensor notation, but I'm emboldened by the very enjoyable article on *The new language of mathematics* by Dan Silver [9]. Your professor's notation is as convenient for actual calculations as -let's say- long division using roman numerals. So leave them wallowing in their early progressive rock of 1968, King Crimsons of their youth. You chill to beats younger than Windows 98, to grime, to trap, to hardvapour, to birdtracks.

In 1937 R. Brauer [1] introduced diagrammatic notation for the Kronecker δ_{ij} operation, in order to represent "Brauer algebra" permutations, index contractions, and matrix multiplication diagrammatically. His equation (39)

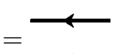


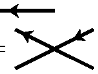
(send index 1 to 2, 2 to 4, contract ingoing (3·4), outgoing (1·3)) is the earliest published diagrammatic notation I know about. While in kindergarten (disclosure: we were too poor to afford kindergarten) I sat out to revolutionize modern group theory [2]. But I suffered a terrible setback; in early 1970's Roger Penrose pre-invented my "birdtracks," or diagrammatic notation, for symmetrization operators [6], Levi-Civita tensors [8], and "strand networks" [5]. Here is a little flavor of how one birdtracks:

We can represent the operation of permuting indices (d "billiard ball labels," tensors with d indices) by a matrix with indices bunched together:

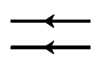
$$\sigma_\alpha^\beta = \sigma_{b_1\ldots b_p}^{a_1 a_2\ldots a_q}{}_{,c_q\ldots c_2 c_1}^{\;d_p\ldots d_1}\,. \tag{10.3}$$

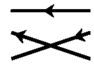
To draw this, Brauer style, it is convenient to turn his drawing on a side. For 2-index tensors, there are two permutations:

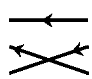
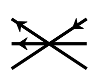
identity: $\mathbf{1}_{ab,}{}^{cd} = \delta_a^d \delta_b^c =$ 

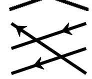
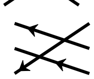
flip: $\sigma_{(12)ab,}{}^{cd} = \delta_a^c \delta_b^d =$  . (10.4)

For 3-index tensors, there are six permutations:

$\mathbf{1}_{a_1a_2a_3,}{}^{b_3b_2b_1} = \delta_{a_1}^{b_1}\delta_{a_2}^{b_2}\delta_{a_3}^{b_3} =$ 

$\sigma_{(12)a_1a_2a_3,}{}^{b_3b_2b_1} = \delta_{a_1}^{b_2}\delta_{a_2}^{b_1}\delta_{a_3}^{b_3} =$ 

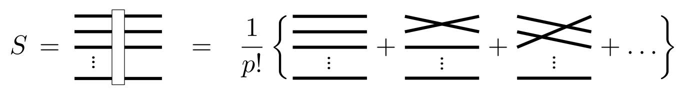
$\sigma_{(23)} =$  , $\sigma_{(13)} =$ 

$\sigma_{(123)} =$  , $\sigma_{(132)} =$  . (10.5)

Here group element labels refer to the standard permutation cycles notation. There is really no need to indicate the "time direction" by arrows, so we omit them from now on.

The symmetric sum of all permutations,

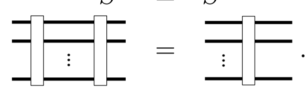
$$S_{a_1 a_2 \ldots a_p,}{}^{b_p \ldots b_2 b_1} = \frac{1}{p!}\left\{\delta_{a_1}^{b_1}\delta_{a_2}^{b_2}\ldots\delta_{a_p}^{b_p} + \delta_{a_2}^{b_1}\delta_{a_1}^{b_2}\ldots\delta_{a_p}^{b_p} + \ldots\right\}$$



(10.6)

yields the symmetrization operator S. In birdtrack notation, a white bar drawn across p lines [6] will always denote symmetrization of the lines crossed. A factor of $1/p!$ has been introduced in order for S to satisfy the projection operator normalization

$$S^2 = S$$



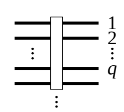
(10.7)

You have already seen such "fully-symmetric representation," in the discussion of discrete Fourier transforms, ChaosBook Example A24.3 *'Configuration-momentum' Fourier space duality*, but you are not likely to recognize it. There the average was not over all permutations, but the zero-th Fourier mode $\tilde{\phi}_0$ was the average over only cyclic permutations. Every finite discrete group has such fully-symmetric representation, and in statistical mechanics and quantum mechanics this is often the most important state (the 'ground' state).

A subset of indices $a_1, a_2, \ldots a_q$, $q < p$ can be symmetrized by symmetrization matrix $S_{12\ldots q}$

$$(S_{12\ldots q})_{a_1 a_2 \ldots a_q \ldots a_p,}{}^{b_p \ldots b_q \ldots b_2 b_1} = \frac{1}{q!}\left\{\delta_{a_1}^{b_1}\delta_{a_2}^{b_2}\ldots\delta_{a_q}^{b_q} + \delta_{a_2}^{b_1}\delta_{a_1}^{b_2}\ldots\delta_{a_q}^{b_q} + \ldots\right\}\delta_{a_q+1}^{b_q+1}\ldots\delta_{a_p}^{b_p}$$

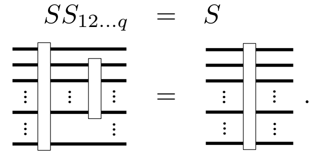
$$S_{12\ldots q} =$$



(10.8)

Overall symmetrization also symmetrizes any subset of indices:

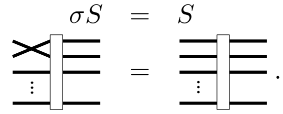
$$SS_{12\ldots q} = S$$



(10.9)

Any permutation has eigenvalue 1 on the symmetric tensor space:

$$\sigma S = S$$



(10.10)

Diagrammatically this means that legs can be crossed and uncrossed at will.

One can construct a projection operator onto the fully antisymmetric space in a similar manner [2]. Other representations are trickier - that's precisely what the theory of finite groups is about.

10.2 It's all about class

You might want to have a look at Harter [4] *Double group theory on the half-shell* (click here). Read appendices B and C on spectral decomposition and class algebras. Article works out some interesting examples.

See also remark 1.1 *Projection operators* and perhaps watch Harter's online lecture from Harter's online course.

There is more detail than what we have time to cover here, but I find Harter's Sect. *3.3 Second stage of non-Abelian symmetry analysis* particularly illuminating. It shows how physically different (but mathematically isomorphic) higher-dimensional irreps are constructed corresponding to different subgroup embeddings. One chooses the irrep that corresponds to a particular sequence of physical symmetry breakings.

10.3 Literature

The exposition (or the corresponding chapter in Tinkham [11]) that we follow here largely comes from Wigner's classic *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* [12], which is a harder going, but the more group theory you learn the more you'll appreciate it. Eugene Wigner got the 1963 Nobel Prize in Physics, so by mid 60's gruppenpest was accepted in finer social circles.

The structure of finite groups was understood by late 19th century. A full list of finite groups was another matter. The complete proof of the classification of all finite groups takes about 3 000 pages, a collective 40-years undertaking by over 100 mathematicians, read the wiki. Not all finite groups are as simple or easy to figure out as D_3. For example, the order of the Ree group ${}^2F_4(2)'$ is $212(26+1)(24-1)(23+1)(2-1)/2 = 17\,971\,200$.

From Emory Math Department: A pariah is real! The simple finite groups fit into 18 families, except for the 26 sporadic groups. 20 sporadic groups AKA the Happy Family are parts of the Monster group. The remaining six loners are known as the pariahs.

Question 10.2. Henriette Roux asks

Q What did you do this weekend?

A The same as every other weekend - prepared week's lecture, with my helpers Avi the Little, Edvard the Nordman, and Malbec el Argentino, under Master Roger's watchful eye, see here.

References

[1] R. Brauer, "On algebras which are connected with the semisimple continuous groups", Ann. Math. **38**, 857 (1937).

[2] P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Groups* (Princeton Univ. Press, Princeton NJ, 2008).

[3] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group Theory: Application to the Physics of Condensed Matter* (Springer, New York, 2007).

[4] 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).

[5] R. Penrose, "Angular momentum: An approach to combinatorical space-time", in *Quantum Theory and Beyond*, edited by T. Bastin (Cambridge Univ. Press, Cambridge, 1971).

[6] R. Penrose, "Applications of negative dimensional tensors", in *Combinatorial mathematics and its applications*, edited by D. J. A. Welsh (Academic, New York, 1971), pp. 221–244.

[7] R. Penrose, *The Road to Reality - A Complete Guide to the Laws of the Universe* (A. A. Knopf, New York, 2005).

[8] R. Penrose and M. A. H. MacCallum, "Twistor theory: An approach to the quantisation of fields and space-time", Phys. Rep. **6**, 241–315 (1973).

[9] D. S. Silver, "The new language of mathematics", Amer. Sci. **105**, 364 (2017).

[10] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

[11] M. Tinkham, *Group Theory and Quantum Mechanics* (Dover, New York, 2003).

[12] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1931).

Exercises

10.1. **1-dimensional representation of anything.** Let $D(g)$ be a representation of a group G. Show that $d(g) = \det D(g)$ is one-dimensional representation of G as well.

(B. Gutkin)

10.2. **2–dimensional representation of S_3.**

(i) Show that the group S_3 of permutations of 3 objects can be generated by two permutations, a transposition and a cyclic permutation:

$$a = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad d = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

(ii) Show that matrices:

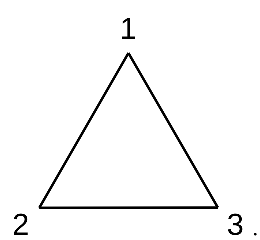
$$D(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ D(a) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ D(d) = \begin{pmatrix} z & 0 \\ 0 & z^2 \end{pmatrix},$$

with $z = e^{i2\pi/3}$, provide proper (faithful) representation for these elements and find representation for the remaining elements of the group.

(iii) Is this representation irreducible?
One of those tricky questions so simple that one does not necessarily get them. If it were reducible, all group element matrices could be simultaneously diagonalized. A motivational (counter)example: as multiplication tables for D_3 and S_3 are the same, consider D_3. Is the above representation of its C_3 subgroup irreducible?

(B. Gutkin)

10.3. **D_3: symmetries of an equilateral triangle.** Consider group $D_3 \cong C_{3v} \cong S_3$, the symmetry group of an equilateral triangle:



(a) List the group elements and the corresponding geometric operations

(b) Find the subgroups of the group D_3.

(c) Find the classes of D_3 and the number of elements in them, guided by the geometric interpretation of group elements. Verify your answer using the definition of a class.

(d) List the conjugacy classes of subgroups of D_3. (continued as exercise 11.2 and exercise 11.3)

10.4. **Permutation of three objects.** Consider S_3, the group of permutations of 3 objects.

(a) Show that S_3 is a group.

(b) List the equivalence classes of S_3?

(c) Give an interpretation of these classes if the group elements are substitution operations on a set of three objects.

(c) Give a geometrical interpretation in case of group elements being symmetry operations on equilateral triangle.

10.5. **3-dimensional representations of $\mathbf{D}_3$.** The group D_3 is the symmetry group of the equilateral triangle. It has 6 elements

$$\mathrm{D}_3 = \{E, C, C^2, \sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}\} ,$$

where C is rotation by $2\pi/3$ and $\sigma^{(i)}$ is reflection along one of the 3 symmetry axes.

(i) Prove that this group is isomorphic to S_3

(ii) Show that matrices

$$D(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, D(C) = \begin{pmatrix} z & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z^2 \end{pmatrix}, D(\sigma^{(1)}) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{10.11}$$

generate a 3-dimensional representation $D(g)$ of D_3. Hint: Calculate products for representations of group elements and compare with the group table (see lecture).

(iii) Show that this is a reducible representation which can be split into one dimensional A and two-dimensional representation Γ. In other words find a matrix R such that

$$\mathbf{R}D(g)\mathbf{R}^{-1} = \begin{pmatrix} A(g) & 0 \\ 0 & \Gamma(g) \end{pmatrix}$$

for all elements g of D_3. (Might help: D_3 has only one (non-equivalent) 2-dim irreducible representation).

(B. Gutkin)

Chapter 10 solutions: Discrete symmetries

Solution 10.1 - 1-dimensional representation of anything.

$$\begin{aligned} \det D(g_1)\det D(g_2) &= \det\left(D(g_1)\,D(g_2)\right) = \det D(g_1 g_2) \\ (\det D(g))^{-1} &= \det D(g^{-1}) \end{aligned} \tag{10.12}$$

(B. Gutkin)

Solution 10.2 - 2-dimensional representation of S_3.

(i) *It is straightforward to check that:*

$$d^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = f.$$

The 3 elements $C_3 = \{e, d, g\}$ provide the cyclic subgroup of S_3. Since ae, ad, af are all different and do not belong to C_3 (otherwise, a would necessarily belong to C_3) these three elements together with C_3 exhaust all S_3.

(ii) *First we need to check that representation matrices keep the same algebra as group elements:*

$$D(a)D(a) = D(a^2) = D(e)\,, \quad D(d)^3 = D(d)^{-1}\,.$$

Indeed

$$D(d^2) = D(d)D(d) = \begin{pmatrix} z^2 & 0 \\ 0 & z \end{pmatrix} = D(d)^{-1}\,.$$

For other group elements:

$$D(b) = D(ad) = D(a)D(d) = \begin{pmatrix} 0 & z^2 \\ z & 0 \end{pmatrix}, \quad D(c) = D(af) = D(a)D(f) = \begin{pmatrix} 0 & z \\ z^2 & 0 \end{pmatrix}$$

(iii) *Yes. Otherwise $D(d)$ and $D(a)$ would be diagonal and commute.*

(B. Gutkin)

Solution 10.3 - $\mathbf{D}_3$: symmetries of an equilateral triangle.

(a) *The group elements are $\{e, \sigma_1, \sigma_2, \sigma_3, r_{1/3}, r_{2/3}\}$. They correspond to the identity, flips which leave one element invariant but switch the other two, and rotations which map the entire set onto itself with no invariants in $1/3$ and $2/3$ rotations.*

(b) *The subgroups of D_3 are $\{e\}$, $\{e, r_{1/3}, r_{2/3}\}$, and $\{e, \sigma_i\}_{i\in\{1,2,3\}}$, and D_3. Rotations by one-third of a circle are the inverse of rotations by two-thirds of a circle. Flips from $i \to j$ can be inverted by the flip $j \to i$.*

(c) *The classes of D_3 are $\{e\}$, $\{r\}$, and $\{\sigma\}$, segmented by the identity, rotations, and flips. Note that $\prod |c(D_3)| = |D_3|$. The definition for a conjugacy class is equivalent to $ag = gb$, for a and b. Each rotation can be "undone" by a flipping sequence.*

(d) *The conjugacy to the subgroup $\{e, r_{1/3}, r_{2/3}\}$ is the set $\{e, \sigma_i\}_{i\in\{1,2,3\}}$, as all flips change the action of the rotation but are their own inverse.*

(Chris Marcotte)

Solution 10.3 - $\mathbf{D}_3$: symmetries of an equilateral triangle.

(a) List the group elements and the corresponding geometric operations

- *Identity* e
- *Reflection along line joining 1* $\to$ *bisection of 2,3:* σ_{12}*;*
  Other reflections σ_{13}*,* σ_{23}
- *Rotation by* $2\pi/3$*: Permutation* (123)*:* $C^{1/3}$
- *Rotation by* $4\pi/3$*: Permutation* (132)*:* $C^{2/3}$

(b) List the subgroups of D_3

- $\{e\}$*,* D_3
- $\{e, \sigma_{12}\}$*,* $\{e, \sigma_{13}\}$*,* $\{e, \sigma_{23}\}$
- $\{e, C^{1/3}, C^{2/3}\}$

(c) Find the classes of D_3 and the number of elements in them, guided by the geometric interpretation of group elements. Verify your answer using the definition of a class. *We can think of* D_3 *as two types of geometrical actions on the triangle: reflect and rotate. We may consider a reflection such as* σ_{12} *and generate the other reflections by rotating the labels into a new frame, performing the reflection, and then transforming back. The same goes for a rotation except the conjugating action is a reflection.*

- *Trivial class:* $\{e\}$
- *Reflections:* $\{\sigma_{12}, \sigma_{23}, \sigma_{13}\}$*.* $\sigma_{23} = \left(C^{1/3}\right)^{-1} \sigma_{12} C^{1/3}$*;* $\sigma_{13} = \left(C^{2/3}\right)^{-1} \sigma_{12} C^{2/3}$
- *Rotations:* $\{C^{1/3}, C^{2/3}\}$*.* $C^{2/3} = \sigma_{12} C^{1/3} \sigma_{12}$

(d) List the conjugacy classes of subgroups of D_3 *To begin, we make the following definitions:*

$$
\begin{aligned}
I &= \{e\} \\
\Sigma_{12} &= \{e, \sigma_{12}\} \\
\Sigma_{13} &= \{e, \sigma_{13}\} \\
\Sigma_{23} &= \{e, \sigma_{23}\} \\
R &= \{e, C^{1/3}, C^{2/3}\}
\end{aligned}
$$

The conjugate of a subgroup $H \subset G$ *is* $g^{-1} H g$ *for* $g \in G$*.*

- *The conjugate of* I *is* I
- *Since we may rotate a reflection into any other reflection,* Σ_{12}*,* Σ_{13}*, and* Σ_{23} *are mutually conjugate.*
- *Finally, since* $C^{1/3}$ *and* $C^{2/3}$ *are conjugate,* R *is only trivially conjugate to* R*.*

(Michael Dimitriyev)

Solution 10.3 - D_3: symmetries of an equilateral triangle.

(a) $D_3 = \{e, \sigma_{12}, \sigma_{13}, \sigma_{23}, C^{1/3}, C^{2/3}\}$
$\{\sigma_{12}, \sigma_{13}, \sigma_{23}\}$ *are reflections across symmetry axes.* $\{C^{1/3}, C^{2/3}\}$ *are rotations by* $2\pi/3$ *and* $4\pi/3$ *respectively.*

(b) *Subgroups are:* $\{e\}$, $\{e, \sigma_{12}\}$,$\{e, \sigma_{13}\}$,$\{e, \sigma_{23}\}$, $\{e, C^{1/3}, C^{2/3}\}$, D_3 *itself.*

(c) *The classes of* D_3 *are* $\{e\}$,$\{\sigma_{12}, \sigma_{23}, \sigma_{13}\}$ *and* $\{C^{1/3}, C^{2/3}\}$.
The first class is just identity. The second class is all the reflections. The third class contains all the rotations.

(d) *The conjugate class of subgroup* $\{e\}$ *is :* $\{e\}$.
The conjugate class of subgroup $\{e, \sigma_{12}\}$ *is :* $\{e\}$, *and* $\{\sigma_{12}\}$.
The conjugate class of subgroup $\{e, \sigma_{23}\}$ *is :* $\{e\}$, *and* $\{\sigma_{23}\}$.
The conjugate class of subgroup $\{e, \sigma_{31}\}$ *is :* $\{e\}$, *and* $\{\sigma_{31}\}$.
The conjugate class of subgroup $\{e, C^{1/3}, C^{2/3}\}$ *is :* $\{e\}$, *and* $\{e, C^{1/3}, C^{2/3}\}$.
The conjugate class of subgroup D_3 *is :* $\{e\}$, $\{e, C^{1/3}, C^{2/3}\}$ *and* $\{\sigma_{12}, \sigma_{23}, \sigma_{31}\}$.

(Xiong Ding)

Solution 10.4 - Permutation of three objects.

1. *In order to show that* S_3 *is a group, it must have (i) closure under its group operation; (ii) associativity; (iii) an identity element; and (iv) have an inverse.*

S_3 *permutes a set of three objects. Let these three objects be* a, b, c. *We can then define the identity* E *to be the ordered set that has not undergone any operations; i.e.,* $E \equiv [abc]$.
The group operation is the transposition of two objects, e.g., $[abc] \mapsto [bac]$. *A matrix* g *that represents this operation is*

$$g = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{10.13}$$

since the operation of this matrix on the column vector $[abc]^T$ *yields* $[bac]^T$, *where* T *denotes the transpose. We further note that*

$$g^2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{I} \tag{10.14}$$

where $\mathbb{I}$ *is the* 3×3 *identity matrix. This suggests that* $g = g^{-1}$ *for this particular group element. In fact, four of the six group elements of* S_3 *are their own inverses:*

$$\mathbb{I}^{-1} = [abc]^{-1} = [abc] = \mathbb{I}, \quad [acb]^{-1} = [acb], \quad [bac]^{-1} = [bac], \quad [cba]^{-1} = [cba] \tag{10.15}$$

The remaining two elements are not their own inverses:

$$[bca]^{-1} = [cab], \quad [cab]^{-1} = [bca] \tag{10.16}$$

Closure can be shown by considering all permutations of the three objects. All possible combinations of permutations yields one of the possible configurations; that is, the group operation (the transposition of two objects) can generate all six configurations of the objects (“group elements”), and all the elements can be generated by the group operation.
Associativity is most readily proved when one represents the group as matrices since matrix multiplication is associative (though, it must be pointed not, matrix multiplication is not, in general, commutative).

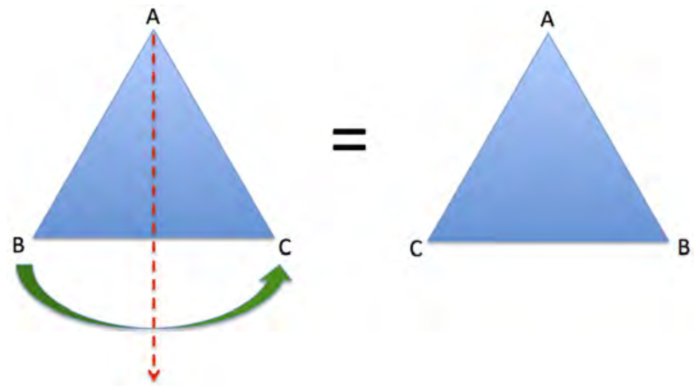


Figure 10.3: Permutation of two vertices of an equilateral triangle.

2. *There are three conjugacy classes of S_3: the identity E (in a class $\mathcal{C}_1$ by itself) and two non-trivial classes.*

   The existence of two non-trivial classes was hinted at in the first part of the problem: there are four group elements that are their own inverses—which suggests that these four elements could form a subgroup of S_3—and two elements that are not. Upon further inspection, we observe that the first non-trivial conjugacy class, which we will call $\mathcal{C}_2$, corresponds to transpositions of 2 objects (i.e., permutations of 2 objects while keeping the remaining objects fixed); this class shares the feature that each element is its own inverse.

   The remaining conjugacy class $\mathcal{C}_3$ corresponds to the remaining two elements that are inverses of each other, but are not inverses of themselves.

3. *The three conjugacy classes may be interpreted as substitution operations of three objects where the specific substitution operation may be transposition or permutation. $\mathcal{C}_1$ is the class described by no substitution, permutation, or transposition; the original ordered set is unchanged. $\mathcal{C}_2$ is the class described by the permutation of two of the three objects. $\mathcal{C}_3$ is the class described by the permutation of all three objects.*

4. *$\mathcal{C}_1$ corresponds to an unchanged configuration, e.g., an equilateral triangle whose vertices are fixed in their original locations.*

   $\mathcal{C}_2$ transposes two vertices of an equilateral triangle; this is equivalent to a rotation of π about an axis passing through the fixed vertex and the center of the triangle. For example, consider $[ABC] \mapsto [ACB]$. Transposing vertices B and C is equivalent to flipping the triangle about the line connecting vertex A to the center (i.e., about an axis of rotation coplanar with the triangle). See figure 10.3.

   $\mathcal{C}_3$ permutes all three vertices of the triangle. This operation is equivalent to rotating the triangle by either $2\pi/3$ or $4\pi/3$ about an axis passing through the center and normal to the plane of the triangle.

Solution 10.5 - 3-dimensional representations of $\mathbf{D}_3$. ***(i)*** *Table 10.1, the group multiplication table for D_3 is the same one as for S_3.*

(ii) *Straightforward check, as in exercise 10.2.*

(iii) *Since the group has only one 2-dimensional irreducible representation Γ should be equivalent to one from exercise 10.2. We therefore can find a matrix R such that:*

$$RD(C)R^{-1} = R\begin{pmatrix} z & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z^2 \end{pmatrix}R^{-1} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & z & 0 \\ 0 & 0 & z^2 \end{pmatrix},$$

where second block in the right hand side is 2-dim representation from exercise 10.2, and α some one dimensional irreducible representation of C. We find that:

$$R = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \alpha = 1.$$

Applying then R to other group elements we get a decomposition of $\mathcal{D}$ into irreducible blocks. For example:

$$RD(\sigma^{(1)})R^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Note: do not panic (yet) if you got a different answer. There are plenty of equivalent irreps.

(B. Gutkin)

mathematical methods - week 11

Continuous symmetries

Georgia Tech PHYS-6124

Homework HW #11 due Thursday, November 5, 2020

== 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 11.1 *Decompose a representation of* S_3
(a) 2; (b) 2; (c) 3; and (d) 3 points
(e) 2 and (f) 3 points bonus points

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

edited October 30, 2020

Week 11 syllabus

Tuesday, October 27, 2020

> I have given up Twitter in exchange for Tacitus & Thucydides, for Newton & Euclid; & I find myself much the happier.
> — Thomas Jefferson to John Adams, 21 January 1812

Clip 1 - They still do not get it!

This week's lectures are related to AWH Chapter 17 *Group Theory, Sect. 17.7 Continuous groups* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- Lie groups, sect. 11.2:
  Definition of a Lie group
  Cyclic group $C_N \to$ continuous $SO(2)$ plane rotations
  Infinitesimal transformations, $SO(2)$ generator of rotations
  $SO(2)$ (group element) = $\exp$(generator)

Clip 2 - What is a symmetry?

Clip 3 - Group element; transformation generator

Clip 4 - What is a symmetry group?

Clip 5 - What is a group orbit?

Clip 6 - What is dynamics?

Clip 7 - Group SO(2)

- The $N \to \infty$ limit of C_N gets you to the continuous Fourier transform as a representation of $SO(2)$, but from then on this way of thinking about continuous symmetries gets to be increasingly awkward. A fresh restart is afforded by matrix groups, and in particular the mother unitary group $U(n) = U(1) \otimes SU(n)$, which contains all other compact groups, finite or continuous, as subgroups.

Clip 10 - Unitary groups are mothers of all finite / compact symmetries. (1 h 4 min)

Discussion 1 - How did we get the Lie algebra? Why is (almost) every symmetry we care about a subgroup of an unitary group? (9 min)

Discussion 2 - How did we get the SO(2) generator? (2 min)

Optional viewing and reading

- What's the payback? While for you the geometrically intuitive representation is the set of rotation $[2{\times}2]$ matrices, group theory says no! They split into pairs of 1-dimensional irreps, and the basic building blocks of *our* 2-dimensional rotations on our kitchen table (forget quantum mechanics!) are the $\mathrm{U}(1)$ $[1\times 1]$ complex unit vector phase rotations.

  Reading: C. K. Wong *Group Theory* notes, Chap 6 *1D continuous groups*, Sects. 6.1-6.3 Irreps of $\mathrm{SO}(2)$.

  Reading: C. K. Wong *Group Theory* notes, Chap 6 *1D continuous groups*, Sect. 6.6 completes discussion of Fourier analysis as continuum limit of cyclic groups C_n, compares $\mathrm{SO}(2)$, $\mathrm{O}(2)$, discrete translations group, and continuous translations group.

Clip 8 - Infinitesimal symmetries: Lie derivative

Clip 9 - Symmetries of solutions. (18 min)

Clip 11 Special orthogonal group SO(n). (9 min)

Clip 12 Symplectic group Sp(n). (9 min)

Discussion 3 - Orthogonal and unitary transformations

Rant 1 - Is beauty symmetry? The first piece of art found in China is a perfect disk carved out of jade. All of Bach is symmetries. (9 min)

Rant 2 - students find letter A beautifully symmetric, but Predrag finds zero 'O' the most beautiful grade. (1 min)

Rant 3 - SO(3) & SU(2) preview and a long rant - listen to it at your own risk. Roger Penrose thoughts on quantum spacetime and quantum brain. Are laws of physics time invariant? Waiting for dark energy to go away. Arrow of time. (17 min)

Rant 4 - SO(3) & SU(2) preview and a long rant - listen to it at your own risk. Get this: math uses 2d complex vectors (spinors) to build our real 3d space. And all we see - starlight, graphene, greenhouse effect, helioseismography, gravitational wave detectors - it is all irreps! (12 min)

Rant 5 - Help me, I'm bullied by a mathematician. (3 min)

Rant 6 - you can always count on Prof. Z. (1/2 min)

Question 11.1. Henriette Roux, pondering exercise 11.1, writes

Q I want to make sure I understand the concept of irreducible representations. In the last homework, we saw that

1. if a representation (which can be thought of as a sort of basis) is reducible, all group element matrices can be simultaneously diagonalized. I want to be able to see how this definition of reducibility matches with the notion of block diagonalizability of an overall representation D(g).

2. AWH p. 822-823 has a discussion of this, but I'm wondering if there's an intuitive way to connect these two definitions or if it's just linear algebra.
3. We familiarized ourselves with the concept of (conjugacy) classes in the last homework. Here, we now add in the concept of character, which, according to AWH, is just the trace of any matrix in a given class (and every matrix of the same class will have the same trace b/c of the properties of classes/traces).
4. So to find the characters for a given representation (part c), we just need to find the classes and then take the trace of a matrix representation in each class?
5. My next and related question then concerns what character means conceptually. Does it relate classes to other classes within a given representation, or different representations (whether reducible or not), or both? AWH says that "the set of characters for all elements and irreducible representations of a finite group defines an orthogonal finite-dimensional vector space."
6. How does a vector space come about from a set of traces, each of which I normally think of as just a number, like the determinant? And finally,
7. how can we use our knowledge of classes/character to find irreducible representations, since that seems to be an important goal in examining a group.
8. exercise 11.1 (c) says to find the characters for this representation, which seems to imply that character depends on representation. But I would've thought that character, which is a trace of a matrix, is invariant under any similarity transform, which is how you get from a reducible representation to an irreducible representation.
9. Also, this is more of a guess than anything, but do the multiplicities of irreducible representations correspond to the multiplicity of characters (i.e. the number of elements in each class)? If so, why? (Or if not, why not?)
10. Same thing for classes, correct?. Classes shouldn't depend on representation b/c they can be thought of as corresponding to a physical operation (e.g. transposition or cyclic permutation), something which is independent of basis.

.

A Great framing for a discussion, thanks! I'll probably reedit this post several times, everybody's input is very welcome. Items numbered as in above:

(2) My favorite step-by-step, pedagogical exposition are the chapters 2 *Representation Theory and Basic Theorems* and 3 *Character of a Representation* of Dresselhaus *et al.* [2]. There is too much material for our course, but if you want to understand it once for all times, it's worth your time.

(3) Correct.

(4) Correct. Note, however, that while every matrix representation has a trace, and thus a character, you want to decompose this character into the sum of irrep characters, as it is obvious after the block diagonalization has been attained.

(5) The unitary diagonalization matrix, whose entries are characters, takes character-weighted sums of classes in order to project them onto irreps, just like what the Fourier representation does. The result, as we know from projection operators of weeks 1 & 2, are mutually orthogonal sub-spaces.

(6) Whenever you do not understand something about finite groups, ask yourself - how does it work for finite lattice Fourier representation?

There the vector space comes via a unitary transformation from the configuration coordinates (where each group element is represented by a full matrix) to the diagonalized, irreducible subspaces coordinates (Fourier modes).

The unitary $\mathcal{F}$ matrix is full of ω^{ij}, ie, characters of the cyclic group C_n. That's where the characters come from.

Now mess up C_3 by adding a reflection. Dihedral group D_3, the group of rotations and reflections, has more symmetry constrains, it cannot have 6 irreps, as reflection invariance mixes together the two senses of rotation. Now there are 3 classes, ie, kinds of things the group does: nothing, flip, rotate. The unitary transformation that diagonalizes group element matrices is now morally a smaller unitary $[3 \times 3]$ matrix from 'classes' in configuration space to 'irreps' in the diagonalized representation, where some sub-spaces must have dimension higher than one.

The surprise, for me, is that the entries in the unitary diagonalization matrix can still be written as traces of irreps, ie, characters. For me it is a calculation, a beautiful example of mathematics leading us somewhere where our intuition falls short. If you find a good intuitive explanation somewhere, please let us all know.

(7) That's automatic, now. Each irrep has a projection operator associated with it. In weeks 1 & 2 we constructed it as a sub-product of factors in Hamilton-Cayley formula. Now we know we can write it -just as we did with the Fourier representation- as sum over all class group actions, each weighted by a the irrep's character.

(8) Characters are elements of the unitary matrix with one index running over classes, the other over irreps. So you expect character to differ from representation to representation; very clear from D_3 character table. As always, you already know that from the Fourier representation example.

(9) Good question. The do not. Dresselhaus *et al.* [2] has the answer - enter it here once you understand it.

(10) Correct.

11.1 Lie groups

In week 1 we introduced projection operators (1.20). How are they related to the character projection operators constructed in the group theory lectures? While the character orthogonality might be wonderful, it is not very intuitive - it's a set of solutions to a set of symmetry-consistent orthogonality relations. You can learn a set of rules that enables you to construct a character table, but it does not tell you what it means. Similar thing will happen again when we turn to the study of continuous groups: all semisimple Lie groups will be classified by Killing and Cartan by a more complex set of orthogonality and integer-dimensionality (Diophantine) constraints. You obtain all possible Lie algebras, but have no idea what their geometrical significance is.

In my own Group Theory book [1] I (almost) get all simple Lie algebras using projection operators constructed from invariant tensors. What that means is easier to understand for finite groups, and here I like the Harter's exposition [4] best. Harter constructs 'class operators', shows that they form a basis for the algebra of 'central' or 'all-commuting' operators, and uses their characteristic equations to construct the projection operators (1.21) from the 'structure constants' of the finite group, i.e., its class multiplication tables. Expanded, these projection operators are indeed the same as the ones obtained from character orthogonality.

11.2 Continuous symmetries : unitary and orthogonal

This week's lectures are not taken from any particular book, they are about basic ideas of how one goes from finite groups to the continuous ones that any physicist should know. We have worked one example out earlier, in week 9 and ChaosBook Sect. A24.4. It gets you to the continuous Fourier transform as a representation of $\mathrm{U}(1) \simeq \mathrm{SO}(2)$, but from then on this way of thinking about continuous symmetries gets to be increasingly awkward. So we need a fresh restart; that is afforded by matrix groups, and in particular the unitary group $\mathrm{U}(n) = \mathrm{U}(1) \otimes \mathrm{SU}(n)$, which contains all other compact groups, finite or continuous, as subgroups.

The main idea in a way comes from discrete groups: the cyclic group C_N is generated by the powers of the smallest rotation by $\Delta\theta = 2\pi/N$, and in the $N \to \infty$ limit one only needs to understand the commutation relations among T_ℓ, generators of infinitesimal transformations,

$$D(\Delta\theta) = 1 + i \sum_\ell \Delta\theta_\ell T_\ell + O(\Delta\theta^2)\,. \tag{11.1}$$

These thoughts are spread over chapters of my book *Group Theory - Birdtracks, Lie's, and Exceptional Groups* [1] that you can steal from my website, but the book itself is too sophisticated for this course. If you ever want to learn some group theory in depth, you'll have to petition the School to offer it.

11.2.1 Lie groups for pedestrians

> [...] which is an expression of consecration of angular momentum.
>
> — Mason A. Porter's student

Definition: A Lie group is a topological group G such that (i) G has the structure of a smooth differential manifold, and (ii) the composition map $G \times G \to G : (g,h) \to gh^{-1}$ is smooth, i.e., $\mathbb{C}^\infty$ differentiable.

Do not be mystified by this definition. Mathematicians also have to make a living. The compact Lie groups that we will deploy here are a generalization of the theory of $\mathrm{SO}(2) \simeq \mathrm{U}(1)$ rotations, i.e., Fourier analysis. By a 'smooth differential manifold' one means objects like the circle of angles that parameterize continuous rotations in a plane, figure 11.1, or the manifold swept by the three Euler angles that parameterize $\mathrm{SO}(3)$ rotations.

By 'compact' one means that these parameters run over finite ranges, as opposed to parameters in hyperbolic geometries, such as Minkowsky $\mathrm{SO}(3,1)$. The groups we focus on here are compact by default, as their representations are linear, finite-dimensional matrix subgroups of the unitary matrix group $\mathrm{U}(d)$.

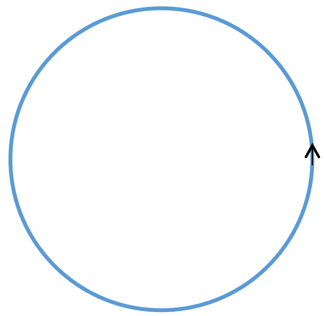


Figure 11.1: Circle group $S^1 = \mathrm{SO}(2)$, the symmetry group of a circle with directed rotations, is a compact group, as its natural parametrization is either the angle $\phi \in [0, 2\pi)$, or the perimeter $x \in [0, L)$.

Example 1. Circle group. A circle with a direction, figure 11.1, is invariant under rotation by any angle $\theta \in [0, 2\pi)$, and the group multiplication corresponds to composition of two rotations $\theta_1 + \theta_2 \mod 2\pi$. The natural representation of the group action is by a complex numbers of absolute value 1, i.e., the exponential $e^{i\theta}$. The composition rule is then the complex multiplication $e^{i\theta_2} e^{i\theta_1} = e^{i(\theta_1+\theta_2)}$. The circle group is a *continuous group*, with infinite number of elements, parametrized by the continuous parameter $\theta \in [0, 2\pi)$. It can be thought of as the $n \to \infty$ limit of the cyclic group C_n. Note that the circle divided into n segments is *compact*, in distinction to the infinite lattice of integers $\mathbb{Z}$, whose limit is a *line* (noncompact, of infinite length).

An element of a $[d \times d]$ -dimensional matrix representation of a *Lie group* continuously connected to identity can be written as

$$g(\boldsymbol{\phi}) = e^{i\boldsymbol{\phi}\cdot T}\,, \qquad \boldsymbol{\phi}\cdot T = \sum_{a=1}^{N} \phi_a T_a\,, \tag{11.2}$$

where $\boldsymbol{\phi}\cdot T$ is a *Lie algebra* element, T_a are matrices called ‘generators’, and $\boldsymbol{\phi} = (\phi_1, \phi_2, \cdots, \phi_N)$ are the parameters of the transformation. Repeated indices are summed throughout, and the dot product refers to a sum over Lie algebra generators. Sometimes it is convenient to use the Dirac bra-ket notation for the Euclidean product of two real vectors $x, y \in \mathbb{R}^d$, or the product of two complex vectors $x, y \in \mathbb{C}^d$, i.e., indicate complex x-transpose times y by

$$\langle x|y\rangle = x^\dagger y = \sum_{i}^{d} x_i^* y_i\,. \tag{11.3}$$

Finite unitary transformations $\exp(i\boldsymbol{\phi}\cdot T)$ are generated by sequences of infinitesimal steps of form

$$g(\boldsymbol{\delta\phi}) \simeq 1 + i\boldsymbol{\delta\phi}\cdot T\,, \quad \boldsymbol{\delta\phi} \in \mathbb{R}^N\,, \quad |\boldsymbol{\delta\phi}| \ll 1\,, \tag{11.4}$$

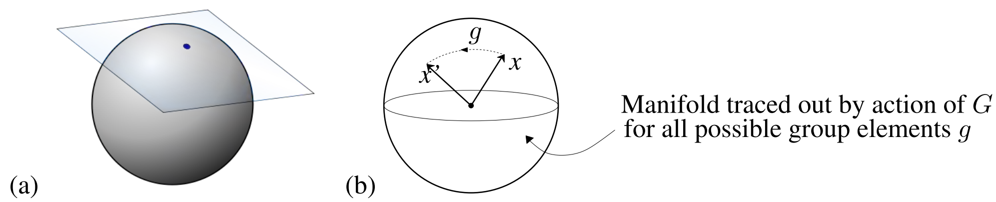


Figure 11.2: (a) Lie algebra fields $\{t_1, \cdots, t_N\}$ span the tangent space of the group orbit $\mathcal{M}_x$ at state space point x, see (11.6) (figure from WikiMedia.org). (b) A global group transformation $g : x \to x'$ can be pieced together from a series of infinitesimal steps along a continuous trajectory connecting the two points. The group orbit of state space point $x \in \mathbb{R}^d$ is the N-dimensional manifold of all actions of the elements of group G on x.

where T_a, the *generators* of infinitesimal transformations, are a set of linearly independent $[d \times d]$ hermitian matrices (see figure 11.2 (b)).

The reason why one can piece a global transformation from infinitesimal steps is that the choice of the "origin" in coordinatization of the group manifold sketched in figure 11.2 (a) is arbitrary. The coordinatization of the tangent space at one point on the group manifold suffices to have it everywhere, by a coordinate transformation g, i.e., the new origin y is related to the old origin x by conjugation $y = g^{-1}xg$, so all tangent spaces belong the same class, they are geometrically equivalent.

Unitary and orthogonal groups are defined as groups that preserve 'length' norms, $\langle gx|gx\rangle = \langle x|x\rangle$, and infinitesimally their generators (11.4) induce no change in the norm, $\langle T_a x|x\rangle + \langle x|T_a x\rangle = 0$, hence the Lie algebra generators T_a are hermitian for,

$$T_a^\dagger = T_a \,. \tag{11.5}$$

The flow field at the state space point x induced by the action of the group is given by the set of N tangent fields

$$t_a(x)_i = (T_a)_{ij} x_j \,, \tag{11.6}$$

which span the d-dimensional *group tangent space* at state space point x, parametrized by $\boldsymbol{\delta\phi}$.

For continuous groups the Lie algebra, i.e., the algebra spanned by the set of N generators T_a of infinitesimal transformations, takes the role that the $|G|$ group elements play in the theory of discrete groups (see figure 11.2).

References

[1] P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Groups* (Princeton Univ. Press, Princeton NJ, 2008).

[2] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group Theory: Application to the Physics of Condensed Matter* (Springer, New York, 2007).

[3] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Dover, New York, 1962).

[4] W. G. Harter, *Principles of Symmetry, Dynamics, and Spectroscopy* (Wiley, New York, 1993).

[5] M. Tinkham, *Group Theory and Quantum Mechanics* (Dover, New York, 2003).

Exercises

11.1. **Decompose a representation of S_3.** Consider a reducible representation $D(g)$, i.e., a representation of group element g that after a suitable similarity transformation takes form

$$D(g) = \begin{pmatrix} D^{(a)}(g) & 0 & 0 & 0 \\ 0 & D^{(b)}(g) & 0 & 0 \\ 0 & 0 & D^{(c)}(g) & 0 \\ 0 & 0 & 0 & \ddots \end{pmatrix},$$

with character for class $\mathcal{C}$ given by

$$\chi(\mathcal{C}) = c_a\,\chi^{(a)}(\mathcal{C}) + c_b\,\chi^{(b)}(\mathcal{C}) + c_c\,\chi^{(c)}(\mathcal{C}) + \cdots ,$$

where c_a, the multiplicity of the ath irreducible representation (colloquially called "irrep"), is determined by the character orthonormality relations,

$$c_a = \overline{\chi^{(a)*}\,\chi} = \frac{1}{h}\sum_k^{class} N_k \chi^{(a)}(\mathcal{C}_k^{-1})\;\chi(\mathcal{C}_k)\,. \tag{11.7}$$

Knowing characters is all that is needed to figure out what any reducible representation decomposes into!

As an example, let's work out the reduction of the matrix representation of S_3 permutations. The identity element acting on the three objects $(a, b, c)^\top$, arranged as components of a 3-vector, is a $[3\times3]$ identity matrix,

$$D(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Transposing the first and second object yields $(b, a, c)^\top$, represented by the matrix

$$D(A) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

since

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} b \\ a \\ c \end{pmatrix}$$

a) Find all six matrices for this representation.

b) Split this representation into its conjugacy classes.

c) Evaluate the characters $\chi(\mathcal{C}_j)$ for this representation.

d) Determine multiplicities c_a of irreps contained in this representation.

e) Construct explicitly all irreps.

f) Explain whether any irreps are missing in this decomposition, and why.

11.2. **Invariance under fractional rotations.** Argue that if the discrete cyclic subgroup

$$\mathrm{C}_N = \{e, C^{1/N}, C^{2/N}, \cdots, (C^{1/N})^{N-1}\}, \qquad (C^{1/N})^N = e$$

of $\mathrm{SO}(2)$ rotations about an axis (let's say the 'z-axis') is a symmetry group of the 'equations of motion' $\dot{x} = v(x)$,

$$C^{1/N} v(x) = v(C^{1/N} x) = v(x)\,,$$

the only non-zero components of Fourier-transformed equations of motion are a_{jN} for $j = 1, 2, \cdots$. Argue that the Fourier representation is then the 'quotient map' of the dynamics, $\mathcal{M}/\mathrm{C}_N$. (Hint: this sounds much fancier than what is - think first of how it applies to the 2- and 3-disk pinballs.)

11.3. **Characters of $\mathbf{D}_3$.** (continued from exercise 10.3) $\mathrm{D}_3 \cong \mathrm{C}_{3v}$, the group of symmetries of an equilateral triangle: has three irreducible representations, two one-dimensional and the other one of multiplicity 2.

(a) All finite discrete groups are isomorphic to a permutation group or one of its subgroups, and elements of the permutation group can be expressed as cycles. Express the elements of the group D_3 as cycles. For example, one of the rotations is (123), meaning that vertex 1 maps to 2, $2 \to 3$, and $3 \to 1$.

(b) Use your representation from exercise 10.3 to compute the D_3 character table.

(c) Use a more elegant method from the group-theory literature to verify your D_3 character table.

(d) Two D_3 irreducible representations are one dimensional and the third one of multiplicity 2 is formed by $[2\times 2]$ matrices. Find the matrices for all six group elements in this representation.

(Hint: get yourself a good textbook, like Dresselhaus *et al.* [2], Tinkham [5] or Hamermesh [3], and read up on classes and characters.)

Chapter 11 solutions: Continuous symmetries

Solution 11.1 - Decompose a representation of S_3

a) *The six matrices*

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix},\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix},\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix},\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$

correspond to permutations $[a\ b\ c], [b\ a\ c], [a\ c\ b], [b\ c\ a], [c\ b\ a], [c\ a\ b]$.

b) *The conjugacy classes are the identity* $\mathcal{C}_1$*, transpositions* $\mathcal{C}_2$ *(the permutation of two objects while keeping the remaining objects fixed), and 3-cycles* $\mathcal{C}_3$ *, such as* $c \to b \to a \to c$*.*

c) *The characters of this representation are*

Class	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$
$\chi(\mathcal{C}_j)$	3	0	1

d) *The representation* D *may be decomposed into irreducible representations* $D^{(i)}$ *as*

$$D = c_1 D^{(1)} \oplus c_2 D^{(2)} \oplus c_3 D^{(3)}$$

where multiplicities c_i *are determined by the character orthonormality relations (11.7):*

$$\begin{aligned}
c_1 &= \overline{\chi^{(1)*}\,\chi} \\
&= \frac{1}{6}\left(1\times1\times3\,+\,2\times1\times0\,+\,3\times1\times1\right)=1 \\
c_2 &= \overline{\chi^{(2)*}\,\chi} \\
&= \frac{1}{6}\left(1\times1\times3\,+\,2\times1\times0\quad+\quad3\times(-1)\times1\right)=0 \\
c_3 &= \overline{\chi^{(3)*}\,\chi} \\
&= \frac{1}{6}\left(1\times2\times3\,+\,2\times(-1)\times0\,+\,3\times0\times1\right)=1
\end{aligned}$$

Thus,

$$D = D^{(1)} \oplus D^{(3)}$$

We observe that $D^{(2)}$ *is missing, but have no clue why.*

e) *The identity element is the only element in its conjugacy class* $\mathcal{C}_1$*, corresponding to the irreducible "trivial" representation with* $\chi^{(j)}(\mathcal{C}_1) = 1$*. Thus the first row of the character table 11.1 is* $[1\ 1\ 1]$*.*

The 2-cycles (ab)*,* (ac)*, and* (bc) *in* $\mathcal{C}_2$ *conjugacy class have alternating character* $\chi = \pm 1$*, depending if the permutation is odd or even; this irreducible representation is sometimes called the "sign representation". Thus the second row of the character table is* $[1\ -1\ 1]$*.*

The 3-cycles (abc) *and* (acb) *make up* $\mathcal{C}_3$ *conjugacy class. In order to fill out the remaining row of the character table, we recall that the order of the group is equal to the sum of the squares of the orders of its irreps; that is,*

$$|S_3| = 6 = 1^1 + 1^2 + f_3^2 \quad \to \quad f_3 = \chi^{(3)}((1)) = 2 \tag{11.8}$$

	(a)	(ab)	(abc)
$\chi^{(1)}$	1	1	1
$\chi^{(2)}$	1	-1	1
$\chi^{(3)}$	2	0	-1

Table 11.1: S_3 character table.

Further, column orthogonality requires

$$0 = \sum_{i=1}^{3} f_i \chi^{(i)}\left((ab)\right) = 1 \cdot 1 + 1 \cdot (-1) + 2\,\chi^{(3)}((ab)) \tag{11.9}$$

which forces $\chi^{(3)}((ab)) = 0$*. Similarly,*

$$0 = \sum_{i=1}^{3} f_i \chi^{(i)}\left((abc)\right) = 1 \cdot 1 + 1 \cdot 1 + 2\,\chi^{(3)}((ab)) \tag{11.10}$$

yields $\chi^{(3)}((abc)) = -1$*, and the full character table 11.1 for* S_3*.*

f) We observe that the character is equal to the number of points fixed by the permutations. A transposition of any two objects (i.e., leading to a 2-cycle) in a two-object set leaves no fixed points; consequently, $\chi^{(3)}((ab)) = 0$*.*

Solution 11.2 - Invariance under fractional rotations. *Consider a system which is equivariant with respect to* $SO(2)$ *rotations about some axis, either all rotations, or rotations by discrete angle* $2\pi/m$ *(a cyclic subgroup* $C_m \in SO(2)$*). Chose two of the coordinates to be in a plane normal to the axis of rotation, and rewrite these variables in terms of polar coordinates* (r, ϕ)*. Take the Fourier transform the azimuthal angle* ϕ*. The system we are considering has the discrete rotational symmetry*

$$C^{1/m} v(x) = v(C^{1/m} x) = v(x), \quad (C^{1/m})^m = e. \tag{11.11}$$

Because of this symmetry the velocity field $v(x)$ *is not only periodic with period* 2π *(as required by continuity at* $\phi = 0$*), it is also periodic with period* $2\pi/m$*, hence it can be represented by a Fourier expansion on the interval* $(0, 2\pi/m)$*. This gives me a complete representation of the function in terms of* $\cos(2\pi mjx)$*,* $\sin(2\pi mjx)$*, where* j *is a nonnegative integer. If I Fourier transform the function on the full interval* $(0, 2\pi)$*, I can evaluate the resulting integrals using the series representation on the interval* $(0, 2\pi/m)$*. The orthonormality of the trigonometric polynomials guarantees that I will get the same result as before. This means that the the non-vanishing coefficients will only involve Fourier modes whose wave numbers are multiples of* m*. (J.M. Heninger)*

Solution 11.3 - Characters of D_3. *No solution available.*

mathematical methods - week 12

SO(3) and SU(2)

Georgia Tech PHYS-6124

Homework HW #12 due Thursday, November 12, 2020

== 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 12.1 *Irreps of SO*(2) 3 points
Exercise 12.2 *Conjugacy classes of SO*(3) 4 points (+ 2 bonus points, if complete)
Exercise 12.3 *The character of SO*(3) *3-dimensional representation* 3 points

Bonus points
Exercise 12.4 *The orthonormality of SO*(3) *characters* 2 points

Total of 10 points = 100 % score.

edited November 5, 2020

Week 12 syllabus

Tuesday, November 3, 2020

This week's lectures are related to AWH Chapter 3 *Vector Analysis* (click here) and Chapter 16 *Angular Momentum* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

There is way too much material in this week's notes. Watch the main sequence of video clips, that and recommended reading should suffice. The rest is optional. You can glance through sect. 12.2 *Linear algebra*, and sect. 12.3 *SO(3) character orthogonality*, but I do not expect you to master this material.

Clip 1 - Rotations in 3 dimensions (30 min)

– OK, I see that formally $\mathrm{SU}(2) \simeq \mathrm{SO}(3)$, but who ordered "spin?"

Clip 4 - Rotations in 2 complex dimensions (42 min)

∘ Read sect. 12.4 *SU(2) Pauli matrices*

∘ Read sect. 12.5 $\mathrm{SU}(2) \simeq \mathrm{SO}(3)$

Optional reading

Clip 2 - Lie algebra (21 min)

Clip 3 - Birdtracks (6 min)

∘ Sect. 10.1.1 *Permutations in birdtracks*

For overall clarity and pleasure of reading, I like Schwichtenberg [8] (click here) discussion best. If you read anything for this week's lectures, read Schwichtenberg.

Reading: Chen, Ping and Wang [2] *Group Representation Theory for Physicists*, Sect 5.2 *Definition of a Lie group, with examples* (click here).

- Dirac belt trick applet

If still anxious, maybe this helps: Mark Staley, *Understanding quaternions and the Dirac belt trick* arXiv:1001.1778.

I have enjoyed reading Mathews and Walker [7] Chap. 16 *Introduction to groups* (click here). Goldbart writes that the book is "based on lectures by Richard Feynman at Cornell University." Very clever. In particular, work through the example of fig. 16.2: it is very cute, you get explicit eigenmodes from group theory alone. The main message is that if you think things through first, you never have to go through using explicit form of representation matrices - thinking in terms of invariants, like characters, will get you there much faster.

Any book, of 100s available, like Cornwell [3] *Group Theory in Physics: An introduction* that covers group theory might be more to your taste.

Hamilton's quaternions

Stone and Goldbart [9] (click here) Chapter 17 Sect 17.6 *Analytic functions and topology* (wherein stereographic projection is revealed to be the geometric origin of the spinor representations of the rotation group)

12.1 Nobel Prize in Physics 2020

Students –really, anybody who has learned some physics– often ask me: is space continuous or discrete?

We do not know, but this week's $\mathrm{SO}(3) \approx \mathrm{SU}(2)$ correspondence is one of the gateway drugs to speculations about quantum underpinnings of the observed spacetime. It start's with Hamilton's quaternions - the discovery that the building blocks of our apparent 3 Euclidian dimensions are 2-dimensional complex spin 1/2 'spinors', and it leads different people to different theories of quantum spacetime - one direction is the one taken by David Ritz Finkelstein, another one leads to Roger Penrose's description of Minkowski spacetime in terms of twistors.

In what follows, Erin Wells Bonning from Emory University and Predrag Cvitanović from the Georgia Tech explain the 2020 Nobel Prize in physics in terms accessible to all.

A half of the 2020 Nobel Prize in Physics was awarded to Roger Penrose, for the discovery that black hole formation is a robust prediction of the general theory of relativity. In 1957 Penrose, then a graduate student, met Georgia Tech's late David Ritz Finkelstein in a fateful meeting that changed both men's lives forever after. It was Finkelstein's extension of the Schwarzschild metric which provided Penrose with an opening into general relativity and set him on the path to his 1965 discovery celebrated by this year's prize.

A half of the 2020 Nobel Prize in Physics was awarded jointly to Reinhard Genzel and Andrea Ghez for the discovery of –in Ghez's words- "The Monster at the heart of the Milky Way," a black hole whose existence had been hypothesized since the early 1970s. In order to visually observe an object that famously does not emit any light, precise measurements of stars moving in the black hole's gravitational field had to be carried out. The independent work of Genzel and Ghez mapping the positions of these stars over many years has led to the clearest evidence yet that the center of our Milky Way galaxy contains "The Monster", that possibly every galaxy contains a black hole, and that the environment near it looks nothing like what was expected.

Nobel Lecture: Roger Penrose, Nobel Prize in Physics 2020 (34 min)

Nobel Prize in Physics 2020 (56 min)

- Abstract
- Penrose slides for Predrag's 1/2 of the presentation
- 2020 Nobel Prizes in Chemistry and Physics, Explained

Roger Penrose gets Nobel Prize. How David Ritz Finkelstein and Roger Penrose met, and exchanged their lives' paths.

Clip 6 - negative dimensions (6 min)

Andrea Ghez: "The Monster at the Heart of our Galaxy"

Veritasium: "The Infinite Pattern That Never Repeats"

Question 12.1. Predrag asks

Q You are graduate students now. Are you ready for The Talk?

A Henriette Roux: I'm ready!

12.1.1 Quaternionic speculations

Predrag: putting this here for a further re-examination - safely ignored:)

Marek Danielewski (AGH), December 29, 2020, and L. Sapa: Foundations of the Quaternion Quantum Mechanics *Foundations of the Quaternion Quantum Mechanics*, Entropy, **2020**, 22, 1424:

"We show that quaternion quantum mechanics has well-founded mathematical roots and can be derived from the model of the elastic continuum by Cauchy, i.e., it can be regarded as representing the physical reality of elastic continuum. Starting from the Cauchy theory (classical balance equations for isotropic Cauchy-elastic material) and using the Hamilton quaternion algebra, we present a rigorous derivation of the quaternion form of the non- and relativistic wave equations. The family of the wave equations and the Poisson equation are a straightforward consequence of the quaternion representation of the Cauchy model of the elastic continuum. This is the most general kind of quantum mechanics possessing the same kind of calculus of assertions as conventional quantum mechanics. The problem of the Schrödinger equation, where imaginary 'i' should emerge, is solved. This interpretation is an attempt to describe the ontology of quantum mechanics, and demonstrates that, besides Bohmian mechanics, the complete ontological interpretations of quantum theory exists."

It has a quack feel to it, but should be easy to work through...

For a different approach, straightforward, no quackery, see Pavel A. Bolokhov *Quaternionic wave function* arXiv:1712.04795: " quaternions form a natural language for the description of quantum-mechanical wave functions with spin. We use the quaternionic spinor formalism which is in one-to-one correspondence with the usual spinor language. No unphysical degrees of freedom are admitted, in contrast to the majority of literature on quaternions. We build a Dirac Lagrangian in the quaternionic form, derive the Dirac equation and take the nonrelativistic limit to find the Schrödinger's equation. We show that the quaternionic formalism is a natural choice to start with, while in the transition to the noninteracting nonrelativistic limit, the quaternionic description effectively reduces to the regular complex wave function language. We provide an easy-to-use grammar for switching between the ordinary spinor language and the description in terms of quaternions. As an illustration of the broader range of the formalism, we also derive the Maxwell's equation from the quaternionic Lagrangian of Quantum Electrodynamics. In order to derive the equations of motion, we develop the variational calculus appropriate for this formalism. "

Commentary:

Manfried Faber, Richard Gill Quaternions were invented by Benjamin Olinde Rodrigues, before Hamilton. (He is also known for Rodrigues formula for Legendre polynomials.) In 1840 he published a result on transformation groups,[4] which applied Leonhard Euler's four squares formula, a precursor to the quaternions of William Rowan Hamilton, to the problem of representing rotations in space.[5] In 1846 Arthur Cayley acknowledged[6] Euler's and Rodrigues' priority describing orthogonal transformations.

Manfried Faber MathsHistory.st-andrews: In 1840 he published a mathematical paper which contains the second result for which he is known today, namely his work on transformation groups where he derived the formula for the composition of successive finite rotations by an entirely geometric method. Rodrigues' composition of rotations is basically the composition of unit quaternions. The paper appeared in volume five of the *Annales de mathématique pures et appliquées* which was perhaps better known as *Annales de Gergonne* and is described in detail in [4].

Predrag I teach my students that SU(2) is double cover of SO(3) and do not do more with quaternions. Octonions is another story...

Richard Gill According to Stigler's law of eponomy, everything worth remembering is associated with the name of someone we want to remember, who did something else.

12.2 Linear algebra

In this section we collect a few basic definitions. A sophisticated reader might prefer skipping straight to the definition of the Lie product (12.8), the big difference between the group elements product used so far in discussions of finite groups, and what is needed to describe continuous groups.

Vector space. A set V of elements $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$ is called a *vector* (or *linear*) *space* over a field $\mathbb{F}$ if

(a) *vector addition* "+" is defined in V such that V is an abelian group under addition, with identity element $\mathbf{0}$;

(b) the set is *closed* with respect to *scalar multiplication* and vector addition

$$\begin{aligned} a(\mathbf{x}+\mathbf{y}) &= a\mathbf{x}+a\mathbf{y}\,, \qquad a,b\in\mathbb{F}\,, \quad \mathbf{x},\mathbf{y}\in V \\ (a+b)\mathbf{x} &= a\mathbf{x}+b\mathbf{x} \\ a(b\mathbf{x}) &= (ab)\mathbf{x} \\ 1\,\mathbf{x} &= \mathbf{x}\,, \qquad 0\,\mathbf{x}=\mathbf{0}\,. \end{aligned} \tag{12.1}$$

Here the field $\mathbb{F}$ is either $\mathbb{R}$, the field of reals numbers, or $\mathbb{C}$, the field of complex numbers. Given a subset $V_0 \subset V$, the set of all linear combinations of elements of V_0, or the *span* of V_0, is also a vector space.

A basis. $\{\mathbf{e}^{(1)}, \cdots, \mathbf{e}^{(d)}\}$ is any linearly independent subset of V whose span is V. The number of basis elements d is the *dimension* of the vector space V.

Dual space, dual basis. Under a general linear transformation $g \in GL(n, \mathbb{F})$, the row of basis vectors transforms by right multiplication as $\mathbf{e}^{(j)} = \sum_k (\mathbf{g}^{-1})^j{}_k \, \mathbf{e}^{(k)}$, and the column of x_a's transforms by left multiplication as $x' = \mathbf{g}x$. Under left multiplication the column (row transposed) of basis vectors $\mathbf{e}_{(k)}$ transforms as $\mathbf{e}_{(j)} = (\mathbf{g}^\dagger)_j{}^k \mathbf{e}_{(k)}$, where the *dual rep* $\mathbf{g}^\dagger = (\mathbf{g}^{-1})^\top$ is the transpose of the inverse of $\mathbf{g}$. This observation motivates introduction of a *dual* representation space $\bar{V}$, the space on which $GL(n, \mathbb{F})$ acts via the dual rep $\mathbf{g}^\dagger$.

Definition. If V is a vector representation space, then the *dual space* $\bar{V}$ is the set of all linear forms on V over the field $\mathbb{F}$.

If $\{\mathbf{e}^{(1)}, \cdots, \mathbf{e}^{(d)}\}$ is a basis of V, then $\bar{V}$ is spanned by the *dual basis* $\{\mathbf{e}_{(1)}, \cdots, \mathbf{e}_{(d)}\}$, the set of d linear forms $\mathbf{e}_{(k)}$ such that

$$\mathbf{e}_{(j)} \cdot \mathbf{e}^{(k)} = \delta_j^k ,$$

where δ_j^k is the Kronecker symbol, $\delta_j^k = 1$ if $j = k$, and zero otherwise. The components of dual representation space vectors $\bar{y} \in \bar{V}$ will here be distinguished by upper indices

$$(y^1, y^2, \ldots, y^n) . \tag{12.2}$$

They transform under $GL(n, \mathbb{F})$ as

$$y'^a = (\mathbf{g}^\dagger)^a{}_b y^b . \tag{12.3}$$

For $GL(n, \mathbb{F})$ no complex conjugation is implied by the † notation; that interpretation applies only to unitary subgroups $U(n) \subset GL(n, \mathbb{C})$. In the index notation, $\mathbf{g}$ can be distinguished from $\mathbf{g}^\dagger$ by keeping track of the relative ordering of the indices,

$$(\mathbf{g})_a^b \to g_a{}^b , \qquad (\mathbf{g}^\dagger)_a^b \to g^b{}_a . \tag{12.4}$$

Algebra. A set of r elements $\mathbf{t}_\alpha$ of a vector space $\mathcal{T}$ forms an algebra if, in addition to the vector addition and scalar multiplication,

(a) the set is *closed* with respect to multiplication $\mathcal{T} \cdot \mathcal{T} \to \mathcal{T}$, so that for any two elements $\mathbf{t}_\alpha, \mathbf{t}_\beta \in \mathcal{T}$, the product $\mathbf{t}_\alpha \cdot \mathbf{t}_\beta$ also belongs to $\mathcal{T}$:

$$\mathbf{t}_\alpha \cdot \mathbf{t}_\beta = \sum_{\gamma=0}^{r-1} \tau_{\alpha\beta}{}^\gamma \mathbf{t}_\gamma , \qquad \tau_{\alpha\beta}{}^\gamma \in \mathbb{C} ; \tag{12.5}$$

(b) the multiplication operation is *distributive*:

$$\begin{aligned}(\mathbf{t}_\alpha + \mathbf{t}_\beta) \cdot \mathbf{t}_\gamma &= \mathbf{t}_\alpha \cdot \mathbf{t}_\gamma + \mathbf{t}_\beta \cdot \mathbf{t}_\gamma \\ \mathbf{t}_\alpha \cdot (\mathbf{t}_\beta + \mathbf{t}_\gamma) &= \mathbf{t}_\alpha \cdot \mathbf{t}_\beta + \mathbf{t}_\alpha \cdot \mathbf{t}_\gamma \,.\end{aligned}$$

The set of numbers $\tau_{\alpha\beta}{}^{\gamma}$ are called the *structure constants*. They form a matrix rep of the algebra,

$$(\mathbf{t}_\alpha)_\beta{}^\gamma \equiv \tau_{\alpha\beta}{}^\gamma \,, \tag{12.6}$$

whose dimension is the dimension r of the algebra itself.

Depending on what further assumptions one makes on the multiplication, one obtains different types of algebras. For example, if the multiplication is associative

$$(\mathbf{t}_\alpha \cdot \mathbf{t}_\beta) \cdot \mathbf{t}_\gamma = \mathbf{t}_\alpha \cdot (\mathbf{t}_\beta \cdot \mathbf{t}_\gamma) \,,$$

the algebra is *associative*. Typical examples of products are the *matrix product*

$$(\mathbf{t}_\alpha \cdot \mathbf{t}_\beta)^c_a = (t_\alpha)^b_a (t_\beta)^c_b \,, \qquad \mathbf{t}_\alpha \in V \otimes \bar{V} \,, \tag{12.7}$$

and the *Lie product*

$$(\mathbf{t}_\alpha \cdot \mathbf{t}_\beta)^c_a = (t_\alpha)^b_a (t_\beta)^c_b - (t_\alpha)^b_c (t_\beta)^a_b \,, \qquad \mathbf{t}_\alpha \in V \otimes \bar{V} \tag{12.8}$$

which defines a *Lie algebra*.

12.3 SO(3) character orthogonality

In 3 Euclidean dimensions, a rotation around z axis is given by the SO(2) matrix

$$R_3(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \exp\varphi \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{12.9}$$

An arbitrary rotation in $\mathbb{R}^3$ can be represented by

$$R_{\boldsymbol{n}}(\varphi) = e^{-i\varphi\,\boldsymbol{n}\cdot\boldsymbol{L}} \qquad \boldsymbol{L} = (L_1, L_2, L_3) \,, \tag{12.10}$$

where the unit vector $\boldsymbol{n}$ determines the plane and the direction of the rotation by angle φ. Here L_1, L_2, L_3 are the generators of rotations along x, y, z axes respectively,

$$L_1 = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} , \quad L_2 = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} , \quad L_3 = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \tag{12.11}$$

with Lie algebra relations

$$[L_i, L_j] = i\varepsilon_{ijk} L_k \,. \tag{12.12}$$

All SO(3) rotations by the same angle θ around different rotation axis $\boldsymbol{n}$ are conjugate to each other,

$$e^{i\phi \boldsymbol{n}_2\cdot \boldsymbol{L}} e^{i\theta \boldsymbol{n}_1\cdot \boldsymbol{L}} e^{-i\phi \boldsymbol{n}_2\cdot \boldsymbol{L}} = e^{i\theta \boldsymbol{n}_3\cdot \boldsymbol{L}}, \tag{12.13}$$

with $e^{i\phi \boldsymbol{n}_2\cdot \boldsymbol{L}}$ and $e^{-i\theta \boldsymbol{n}_2\cdot \boldsymbol{L}}$ mapping the vector $\boldsymbol{n}_1$ to $\boldsymbol{n}_3$ and back, so that the rotation around axis $\boldsymbol{n}_1$ by angle θ is mapped to a rotation around axis $\boldsymbol{n}_3$ by the same θ. The conjugacy classes of SO(3) thus consist of rotations by the same angle about all distinct rotation axes, and are thus labelled the angle θ. As the conjugacy class depends only on θ, the characters can only be a function of θ. For the 3-dimensional special orthogonal representation, the character is

exercise 12.3

$$\chi = 2\cos(\theta) + 1\,. \tag{12.14}$$

For an irrep labeled by j, the character of a conjugacy class labeled by θ is

$$\chi^{(j)}(\theta) = \frac{\sin(j+1/2)\theta}{\sin(\theta/2)} \tag{12.15}$$

To check that these characters are orthogonal to each other, one needs to define the group integration over a parametrization of the SO(3) group manifold. A group element is parametrized by the rotation axis $\boldsymbol{n}$ and the rotation angle $\theta \in (-\pi, \pi]$, with $\boldsymbol{n}$ a unit vector which ranges over all points on the surface of a unit ball. Note however, that a π rotation is the same as a $-\pi$ rotation ($\boldsymbol{n}$ and $-\boldsymbol{n}$ point along the same direction), and the $\boldsymbol{n}$ parametrization of SO(3) is thus a 2-dimensional surface of a unit-radius ball with the opposite points identified.

The Haar measure for SO(3) requires a bit of work, here we just note that after the integration over the solid angle (characters do not depend on it), the Haar measure is

$$dg = d\mu(\theta) = \frac{d\theta}{2\pi}(1-\cos(\theta)) = \frac{d\theta}{\pi}\sin^2(\theta/2)\,. \tag{12.16}$$

exercise 12.4

With this measure the characters are orthogonal, and the character orthogonality theorems follow, of the same form as for the finite groups, but with the group averages replaced by the continuous, parameter dependant group integrals

$$\frac{1}{|G|}\sum_{g\in G} \to \int_G dg\,.$$

The good news is that, as explained in ChaosBook.org Chap. *Relativity for cyclists* (and in *Group Theory - Birdtracks, Lie's, and Exceptional Groups* [5]), one never needs to actually explicitly construct a group manifold parametrizations and the corresponding Haar measure.

12.4 SU(2) Pauli matrices

A lightning, bullet points review.

- U(n): unitary transformation $U = e^{iH}$

- Unitarity: $U^\dagger U = \mathbf{1} \Rightarrow H^\dagger = H$, the generator is hermitian.
- SU(n): special unitary transformation det $U = 1$
- Must know: $\ln \det\, = \mathrm{tr}\, \ln$ for any matrix, so the generator is traceless
  $\ln \det U = \mathrm{tr}\, \ln U \, = \mathrm{tr}\, H = 0$
- $\mathrm{SU}(2) : H = \begin{pmatrix} a & c \\ e & b \end{pmatrix}\,, \quad a, b, c, e \in \mathbb{C}$, eight real numbers in all.
- H is hermitian: $H = \begin{pmatrix} a & c+id \\ c-id & b \end{pmatrix}\,, \quad a, b, c, d \in \mathbb{R}\,,$
- H is traceless: $0 = \mathrm{tr}\, H \Rightarrow a + b = 0$, three real rotation parameters in all, so

$$\begin{aligned} H &= c\sigma_x + d\sigma_y + a\sigma_z \\ &= c\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + a\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \qquad (12.17)$$

  where σ_j are known as Pauli matrices.

12.5 SU(2) and SO(3)

K. Y. Short

An element of SU(2) can be written as

$$U_{\mathbf{n}}(\phi) = e^{i\phi\, \sigma\cdot\hat{\mathbf{n}}/2} \qquad (12.18)$$

where σ_j is a Pauli matrix and ϕ is a real number. What is the significance of the $1/2$ factor in the argument of the exponential?

Consider a generic position vector $\boldsymbol{x} = (x, y, z)$ and construct a Hermitian matrix of the form

$$\begin{aligned} \sigma\cdot\boldsymbol{x} &= \sigma_x x + \sigma_y y + \sigma_z z \\ &= \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} + \begin{pmatrix} 0 & -iy \\ iy & 0 \end{pmatrix} + \begin{pmatrix} z & 0 \\ 0 & -z \end{pmatrix} \\ &= \begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix} \end{aligned} \qquad (12.19)$$

Its determinant

$$\det\begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix} = -(x^2+y^2+z^2) = -\boldsymbol{x}^2 \qquad (12.20)$$

gives the length of a vector. Consider a SU(2) transformation (12.18) of this matrix, $U^\dagger(\sigma\cdot\boldsymbol{x})U$. Taking the determinant, we find the same expression as before:

$$\det U(\sigma\cdot\boldsymbol{x})U^\dagger = \det U\, \det(\sigma\cdot\boldsymbol{x})\, \det U^\dagger = \det(\sigma\cdot\boldsymbol{x})\,. \qquad (12.21)$$

Just as SO(3), SU(2) preserves the lengths of vectors.

To make the correspondence between SO(3) and SU(2) more explicit, consider a SU(2) transformation on a complex two-component *spinor*

$$\psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{12.22}$$

related to $\boldsymbol{x}$ by

$$x = \frac{1}{2}(\beta^2 - \alpha^2), \quad y = -\frac{i}{2}(\alpha^2 + \beta^2), \quad z = \alpha\beta \tag{12.23}$$

Check that a SU(2) transformation of ψ is equivalent to a SO(3) transformation on $\boldsymbol{x}$. From this equivalence, one sees that a SU(2) transformation has three real parameters that correspond to the three rotation angles of SO(3). If we label the "angles" for the SU(2) transformation by α, β, and γ, we observe, for a "rotation" about $\hat{x}$

$$U_x(\alpha) = \begin{pmatrix} \cos\alpha/2 & i\sin\alpha/2 \\ i\sin\alpha/2 & \cos\alpha/2 \end{pmatrix}, \tag{12.24}$$

for a "rotation" about $\hat{y}$,

$$U_y(\beta) = \begin{pmatrix} \cos\beta/2 & \sin\beta/2 \\ -\sin\beta/2 & \cos\beta/2 \end{pmatrix}, \tag{12.25}$$

and for "rotation" about $\hat{z}$,

$$U_z(\gamma) = \begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix}. \tag{12.26}$$

Compare these three matrices to the corresponding SO(3) rotation matrices:

$$R_x(\zeta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\zeta & \sin\zeta \\ 0 & -\sin\zeta & \cos\zeta \end{pmatrix}, \quad R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix}$$
$$R_z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{12.27}$$

They're equivalent! Result: *Half the rotation angle generated by SU(2) corresponds to a rotation generated by SO(3).*

What does this mean? At this point, probably best to switch to Schwichtenberg [8] (click here) who explains clearly that SU(2) is a simply-connected group, and thus the "mother" or covering group, or the double cover of SO(3). This means there is a two-to-one map from SU(2) to SO(3); an SU(2) turn by 4π corresponds to an SO(3) turn by 2π. So, the building blocks of your 3-dimensional world are not 3-dimensional real vectors, but the 2-dimensional complex spinors! Quantum mechanics chose electrons to be spin 1/2, and there is nothing Fox Channel can do about it.

Question 12.2. Henriette Roux asks
Q Why is this complex 2-dimensional vector called a 'spinor'?
A Historical, as Arfken, Weber & Harris [1] explain: "It turns out that half-integral angular momentum states are needed to describe the intrinsic angular momentum of the electron and many other particles. Since these particles also have magnetic moments, an intuitive interpretation is that their charge distributions are spinning about some axis; hence the term spin. It is now understood that the spin phenomena cannot be explained consistently by describing these particles as ordinary charge distributions undergoing rotational motion, [...] "

Schwichtenberg [8]: "[...] spinors have properties that usual vectors do not have. For instance, the factor $1/2$ in the exponent. This factor shows us that a spinor1 is after a rotation by 2π not the same, but gets a minus sign. This is a pretty crazy property, because all objects we deal with in everyday life are exactly the same after a rotation by $360^o = 2\pi$.

Question 12.3. Henriette Roux asks
Q What' relation of Pauli exclusion principle to the spinor 2π rotation amounting to overall minus sign?
A I think of fermion/Grassmann statistics as Archimedes principle + linearity, see my Field Theory [4] *chap. 4 Fermions*. Basically, usually a constraint is imposed by eliminating a variable, for example, given the constraint is $x^2 + y^2 + z^2 = 1$, one gets rid of z by replacing it everywhere with $z \to \sqrt{1 - x^2 - y^2}$. This makes a fully symmetric theory asymmetric and ugly. In linear setting, another option is to keep all the variables and the symmetry, but add a new variable which by construction subtracts a degree of freedom, what I call [6] a "negative dimension". In quantum field theory such variable is called a 'ghost'; it needs to be anti-commuting or Grassmann.

12.6 What *really* happened

They do not make Norwegians as they used to. In his brief biographical sketch of Sophus Lie, Burkman writes: "I feel that I would be remiss in my duties if I failed to mention avery interesting event that took place in Lie's life. Klein (a German) and Lie had moved to Paris in the spring of 1870 (they had earlier been working in Berlin). However, in July 1870, the Franco-Prussian war broke out. Being a German alien in France, Klein decided that it would be safer to return to Germany; Lie also decided to go home to Norway. However (in a move that I think questions his geometric abilities), Lie decided that to go from Paris to Norway, he would walk to Italy (and then presumably take a ship to Norway). The trip did not go as Lie had planned. On the way, Lie ran into some trouble–first some rain, and he had a habit of taking off his clothes and putting them in his backpack when he walked in the rain (so he was walking to Italy in the nude). Second, he ran into the French military (quite possibly while walking in the nude) and they discovered in his sack (in addition to his hopefully dry clothing) letters written to Klein in German containing the words 'lines' and 'spheres' (which the French interpreted as meaning 'infantry' and 'artillery'). Lie was arrested as a (insane) German spy. However, due to intervention by Gaston Darboux, he was released four weeks later and returned to Norway to finish his doctoral dissertation."

Question 12.4. Henriette Roux asks

Q

A This is a math methods course. Why are you not teaching us Bessel functions?

Blame Feynman: On May 2, 1985 my stay at Cornell was to end, and Vinnie of college town *Italian Kitchen* made a special dinner for three of us regulars. Das Wunderkind noticed Feynman ambling down Eddy Avenue, kidnapped him, and here we were, two wunderkinds, two humans.

Feynman was a very smart, forever driven wunderkind. He naturally bonded with our very smart, forever driven wunderkind, who suddenly lurched out of control, and got very competive about at what age who summed which kind of Bessel function series. Something like age twelve, do not remember which one did the Bessels first. At that age I read " Palle Alone in the World," while my nonwunderkind friend, being from California, watched television 12 hours a day.

When Das Wunderkind taught graduate E&M, he spent hours crafting lectures about symmetry groups and their representations as various eigenfunctions. Students were not pleased.

So, fuggedaboutit! if you have not done your Bessels yet, they are eigenfunctions, just like your Fourier modes, but for a spherical symmetry rather than for a translation symmetry; wiggle like a cosine, but decay radially.

When you need them you'll figure them out. Or sue me.

References

[1] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th ed. (Academic, New York, 2013).

[2] J.-Q. Chen, J. Ping, and F. Wang, *Group Representation Theory for Physicists* (World Scientific, Singapore, 1989).

[3] J. F. Cornwell, *Group Theory in Physics: An Introduction* (Academic, New York, 1997).

[4] P. Cvitanović, *Field Theory*, Notes prepared by E. Gyldenkerne (Nordita, Copenhagen, 1983).

[5] P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Groups* (Princeton Univ. Press, Princeton NJ, 2008).

[6] P. Cvitanović and A. D. Kennedy, "Spinors in negative dimensions", Phys. Scr. **26**, 5–14 (1982).

[7] J. Mathews and R. L. Walker, *Mathematical Methods of Physics* (W. A. Benjamin, Reading, MA, 1970).

[8] J. Schwichtenberg, *Physics from Symmetry* (Springer, Berlin, 2015).

[9] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge UK, 2009).

Exercises

12.1. **Irreps of SO(2).** Matrix

$$T = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{12.28}$$

is the generator of rotations in a plane.

(a) Use the method of projection operators to show that for rotations in the kth Fourier mode plane, the irreducible $1D$ subspaces orthonormal basis vectors are

$$\mathbf{e}^{(\pm k)} = \frac{1}{\sqrt{2}} \left(\pm \mathbf{e}_1^{(k)} - i\, \mathbf{e}_2^{(k)} \right) .$$

How does T act on $\mathbf{e}^{(\pm k)}$?

(b) What is the action of the $[2\times 2]$ rotation matrix

$$D^{(k)}(\theta) = \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} , \qquad k = 1, 2, \cdots$$

on the $(\pm k)$th subspace $\mathbf{e}^{(\pm k)}$?

(c) What are the irreducible representations characters of SO(2)?

12.2. **Conjugacy classes of SO(3):** Show that all SO(3) rotations (12.10) by the same angle θ around any rotation axis $\boldsymbol{n}$ are conjugate to each other:

$$e^{i\phi \boldsymbol{n}_2 \cdot \boldsymbol{L}} e^{i\theta \boldsymbol{n}_1 \cdot \boldsymbol{L}} e^{-i\phi \boldsymbol{n}_2 \cdot \boldsymbol{L}} = e^{i\theta \boldsymbol{n}_3 \cdot \boldsymbol{L}} \tag{12.29}$$

Check this for infinitesimal ϕ, and argue that from that it follows that it is also true for finite ϕ. Hint: use the Lie algebra commutators (12.12).

12.3. **The character of SO(3) 3-dimensional representation:** Show that for the 3-dimensional special orthogonal representation (12.10), the character is

$$\chi = 2\cos(\theta) + 1 . \tag{12.30}$$

Hint: evaluate the character explicitly for $R_x(\theta)$, $R_y(\theta)$ and $R_z(\theta)$, then explain what is the intuitive meaning of 'class' for rotations.

12.4. **The orthonormality of SO(3) characters:** Verify that given the Haar measure (12.16), the characters (12.15) are orthogonal:

$$\langle \chi(j) | \chi(j') \rangle = \int_G dg\, \chi^{(j)}(g^{-1})\, \chi^{(j')}(g) = \delta_{jj'} . \tag{12.31}$$

Chapter 12 solutions: SO(3) and SU(2)

Solution 12.1 - Irreps of SO(2). *Read D. Vvedensky group theory notes, chapter 8, sects. 8.1 and 8.2.*

Solution 12.1 - Irreps of SO(2) .
(a) *The generator $\mathbb{T}$ has two eigenvalues $+1$ and -1, corresponding to eigenvectors $e^+ = \{1, i\}$ and $e^- = \{1, -i\}$ respectively. Since e^+ and e^- are the eigenvectors of $\mathbb{T}$:*

$$\mathbb{T}e^{(\pm)} = \pm 1 e^{(\pm)}$$

(b) *Since $D^k(\theta) = e^{i\theta k \mathbb{T}}$,*

$$D^k(\theta)\mathbf{e}^{(\pm)} = e^{\pm i\theta k}\mathbf{e}^{(\pm)}$$

(c) *The SO(2) group is an Abelian group. So it will only have one-dimensional irreps. The kth irreducible representation of a rotation by angle θ is $e^{ik\theta}$, where k can be any integer. And this is also the character of the representations since they are one-dimensional.*

(Han Liang)

Solution 12.1 - Irreps of SO(2) .
(a) *We instead find eigenvalues of a generic $\mathrm{SO}(2)$ matrix with period $k \in \mathbb{Z}$:*

$$\begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix}.$$

These are easily worked out to be $\lambda_\pm = \cos k\theta \pm i \sin k\theta$. The projection operators are therefore

$$P_+ = \frac{1}{2i \sin k\theta}\left[\begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} - (\cos k\theta - i\sin k\theta)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$
$$= \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix},$$

and

$$P_- = -\frac{1}{2i \sin k\theta}\left[\begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} - (\cos k\theta + i\sin k\theta)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$
$$= \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

Therefore, the irreducible $1D$ subspaces orthonormal basis vectors are

$$\mathbf{e}^{(\pm k)} = \frac{1}{\sqrt{2}}\left(\mathbf{e}_1^{(k)} \mp i\mathbf{e}_2^{(k)}\right), \qquad \mathbf{e}_1 = (1,0)^T, \quad \mathbf{e}_2 = (0,1)^T,$$

and these are the same as the ones given in the problem statement up to a sign. The action of T on $\mathbf{e}^{(\pm k)}$ is given by

$$T\mathbf{e}^{(\pm k)} = \frac{1}{\sqrt{2}}\left(T\mathbf{e}_1^{(k)} \mp iT\mathbf{e}_2^{(k)}\right) = \frac{1}{\sqrt{2}}\left(-i\mathbf{e}_2^{(k)} \pm \mathbf{e}_1^{(k)}\right) = \frac{1}{\sqrt{2}}\left(\pm\mathbf{e}_1^{(k)} - i\mathbf{e}_2^{(k)}\right) = \pm\mathbf{e}^{(\pm k)}.$$

In other words, T leaves the subspaces invariant.

(b) *We compute the action as follows:*

$$\begin{aligned} D^{(k)}(\theta)\mathbf{e}^{(\pm k)} &= \frac{1}{\sqrt{2}}\left(D^{(k)}(\theta)\mathbf{e}_1^{(k)} \mp iD^{(k)}(\theta)\mathbf{e}_2^{(k)}\right) \\ &= \frac{1}{\sqrt{2}}\left(\cos(k\theta)\mathbf{e}_1^{(k)} + \sin(k\theta)\mathbf{e}_2^{(k)} \pm i\sin(k\theta)\mathbf{e}_1^{(k)} \mp i\cos(k\theta)\mathbf{e}_2^{(k)}\right) \\ &= \cos(k\theta)\mathbf{e}^{(\pm k)} \pm i\sin(k\theta)\mathbf{e}^{(\pm k)} \\ &= [\cos(k\theta) \pm i\sin(k\theta)]\,\mathbf{e}^{(\pm k)}. \end{aligned}$$

We could also have just note that the $\mathbf{e}^{(\pm k)}$ are the complex eigenvectors of $D^{(k)}(\theta)$.
(c) *Since $\mathrm{SO}(2)$ is abelian and, hence, irreps are 1-dimensional, the characters and the irreps are the same things. Any irrep $\chi : \mathrm{SO}(2) \to \mathbb{C}$ will vary smoothly with the rotation parameter θ and must satisfy $\chi(\theta+\theta') = \chi(\theta)\chi(\theta')$. We must also have $\chi(0) = 1$ and $\chi(\theta + 2\pi)$. The only solution to these constraints are the functions*

$$\chi^{(m)}(\theta) = e^{im\theta}, \qquad m \in \mathbb{Z}.$$

Therefore, we find that there are infinitely many one-dimensional irreps of $\mathrm{SO}(2)$, labelled by the integers.

(T. Forrest Kieffer)

Solution 12.2 - Conjugacy classes of $\mathbf{SO}(3)$. *See SO3ConjClasses.pdf.*

Solution 12.2 - Conjugacy classes of $\mathbf{SO}(3)$. *We take as our generators*

$$L_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad L_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We do this to rid ourselves of the imaginary unit i in what follows. The commutation relations are $[L_i, L_j] = \epsilon_{ijk}L_k$. We may represent a rotation about the axis $\mathbf{n}$ through the angle θ as $R_{\mathbf{n}}(\theta) = e^{\theta\mathbf{n}\cdot\mathbf{L}}$. We note that

$$e^{\theta\mathbf{n}\cdot\mathbf{L}} = I + (\mathbf{n}\cdot\mathbf{L})\sin(\theta) + 2(\mathbf{n}\cdot\mathbf{L})^2\sin^2(\theta/2).$$

Given real unit vectors $\mathbf{n}_1$ and $\mathbf{n}_3$, and an angle $\theta \in [0, 2\pi)$ we aim to construct a real unit vector $\mathbf{n}_2$ and an angle $\phi \in [0, 2\pi)$ so that

$$e^{\phi\mathbf{n}_2\cdot\mathbf{L}}e^{\theta\mathbf{n}_1\cdot\mathbf{L}}e^{-\phi\mathbf{n}_2\cdot\mathbf{L}} = e^{\theta\mathbf{n}_3\cdot\mathbf{L}}.$$

We will frequently use the identity

$$(\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L}) = n_1^i n_2^j L_i L_j = (\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L}) + \epsilon_{ijk}n_1^i n_2^j L_k = (\mathbf{n}_1\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L}) - (\mathbf{n}_1\times\mathbf{n}_2)\cdot\mathbf{L}.$$

Using the above formulas we find

$$\begin{aligned} &e^{\phi\mathbf{n}_2\cdot\mathbf{L}}e^{\theta\mathbf{n}_1\cdot\mathbf{L}} \\ &= \left(I + (\mathbf{n}_2\cdot\mathbf{L})\sin(\phi) + 2(\mathbf{n}_2\cdot\mathbf{L})^2\sin^2(\phi/2)\right)\left(I + (\mathbf{n}_1\cdot\mathbf{L})\sin(\theta) + 2(\mathbf{n}_1\cdot\mathbf{L})^2\sin^2(\theta/2)\right) \\ &= I + (\mathbf{n}_1\cdot\mathbf{L})\sin(\theta) + 2(\mathbf{n}_1\cdot\mathbf{L})^2\sin^2(\theta/2) + (\mathbf{n}_2\cdot\mathbf{L})\sin(\phi) + (\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L})\sin(\theta)\sin(\phi) \\ &\quad + 2(\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L})^2\sin(\phi)\sin^2(\theta/2) + 2(\mathbf{n}_2\cdot\mathbf{L})^2\sin^2(\phi/2) + 2(\mathbf{n}_2\cdot\mathbf{L})^2(\mathbf{n}_1\cdot\mathbf{L})\sin(\theta)\sin^2(\phi/2) \\ &\quad + 4(\mathbf{n}_2\cdot\mathbf{L})^2(\mathbf{n}_1\cdot\mathbf{L})^2\sin^2(\phi/2)\sin^2(\theta/2), \end{aligned}$$

and we compare this with

$$
\begin{aligned}
& e^{\theta \mathbf{n}_3\cdot\mathbf{L}} e^{\phi \mathbf{n}_2\cdot\mathbf{L}} \\
&= \left(I + (\mathbf{n}_3\cdot\mathbf{L})\sin(\theta) + 2(\mathbf{n}_3\cdot\mathbf{L})^2\sin^2(\theta/2)\right)\left(I + (\mathbf{n}_2\cdot\mathbf{L})\sin(\phi) + 2(\mathbf{n}_2\cdot\mathbf{L})^2\sin^2(\phi/2)\right) \\
&= I + (\mathbf{n}_2\cdot\mathbf{L})\sin(\phi) + 2(\mathbf{n}_2\cdot\mathbf{L})^2\sin^2(\phi/2) + (\mathbf{n}_3\cdot\mathbf{L})\sin(\theta) + (\mathbf{n}_3\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L})\sin(\phi)\sin(\theta) \\
&\quad + 2(\mathbf{n}_3\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L})^2\sin(\theta)\sin^2(\phi/2) + 2(\mathbf{n}_3\cdot\mathbf{L})^2\sin^2(\theta/2) + 2(\mathbf{n}_3\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L})\sin(\phi)\sin^2(\theta/2) \\
&\quad + 4(\mathbf{n}_3\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L})^2\sin^2(\phi/2)\sin^2(\theta/2).
\end{aligned}
$$

We choose $\mathbf{n}_2$ to be the unit vector, orthogonal to $\mathbf{n}_1$, so that $\mathbf{n}_2\times\mathbf{n}_1 = \mathbf{n}_3$ and $\phi = \pi/2$. Using the commutation relation above and taking the difference we find

$$
\begin{aligned}
& e^{\phi \mathbf{n}_2\cdot\mathbf{L}} e^{\theta \mathbf{n}_1\cdot\mathbf{L}} - e^{\theta \mathbf{n}_3\cdot\mathbf{L}} e^{\phi \mathbf{n}_2\cdot\mathbf{L}} \\
&= ((\mathbf{n}_1\cdot\mathbf{L}) - (\mathbf{n}_3\cdot\mathbf{L}))\sin(\theta) + 2\left((\mathbf{n}_1\cdot\mathbf{L})^2 - (\mathbf{n}_3\cdot\mathbf{L})^2\right)\sin^2(\theta/2) \\
&\quad + ((\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L}) - (\mathbf{n}_3\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L}))\sin(\phi)\sin(\theta) \\
&\quad + 2\left((\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L})^2 - (\mathbf{n}_3\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L})\right)\sin(\phi)\sin^2(\theta/2) \\
&\quad + 2\left((\mathbf{n}_2\cdot\mathbf{L})^2(\mathbf{n}_1\cdot\mathbf{L}) - (\mathbf{n}_3\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L})^2\right)\sin(\theta)\sin^2(\phi/2) \\
&\quad + 4\left((\mathbf{n}_2\cdot\mathbf{L})^2(\mathbf{n}_1\cdot\mathbf{L})^2 - (\mathbf{n}_3\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L})^2\right)\sin^2(\phi/2)\sin^2(\theta/2) \\
&= \left((\mathbf{n}_1\cdot\mathbf{L}) + (\mathbf{n}_1\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L}) + (\mathbf{n}_1\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L})^2 + (\mathbf{n}_2\cdot\mathbf{L})(\mathbf{n}_3\cdot\mathbf{L}) - (\mathbf{n}_3\cdot\mathbf{L})(\mathbf{n}_2\cdot\mathbf{L})^2\right)\sin(\theta) \\
&\quad + 2\left((\mathbf{n}_1\cdot\mathbf{L})^2 - (\mathbf{n}_3\cdot\mathbf{L})^2 + (\mathbf{n}_1\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L}) + (\mathbf{n}_1\cdot\mathbf{L})(\mathbf{n}_3\cdot\mathbf{L}) + (\mathbf{n}_3\cdot\mathbf{L})(\mathbf{n}_1\cdot\mathbf{L})\right. \\
&\qquad\qquad \left. - (\mathbf{n}_3\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L}) + (\mathbf{n}_2\cdot\mathbf{L})^2(\mathbf{n}_1\cdot\mathbf{L})^2 - (\mathbf{n}_3\cdot\mathbf{L})^2(\mathbf{n}_2\cdot\mathbf{L})^2\right)\sin^2(\theta/2).
\end{aligned}
$$

After a long computation using commutators and vector identities, we find that the above expression can be made identically zero. *(T. Forrest Kieffer)*

Solution 12.3 - The character of $\mathbf{SO}(3)$ 3-dimensional representation: *Pick n along z axis, take the trace of $R_3(\theta)$, given in (12.9).*

Solution 12.3 - The character of $\mathbf{SO}(3)$ 3-dimensional representation: *Let us compute the eigenvalues of $\mathbf{n}\cdot\mathbf{L} = n_1L_1 + n_2L_2 + n_3L_3$. We compute the roots of* $\det(\mathbf{n}\cdot\mathbf{L} - \lambda I)$ *to find the eigenvalues of 0 and ± 1 (we used the fact that $\mathbf{n}$ is a unit vector). Hence,*

$$
\chi(\theta) = \operatorname{tr} e^{-i\theta\mathbf{n}\cdot\mathbf{L}} = e^{-i\theta\cdot 0} + e^{-i\theta} + e^{i\theta} = 1 + 2\cos\theta.
$$

(T. Forrest Kieffer)

Solution 12.4 - The orthonormality of $\mathbf{SO}(3)$ characters: *With the Haar measure (12.16) the characters are orthogonal by the standard orthogonality of trigonometric functions:*

$$
\langle\chi(j)|\chi(j')\rangle = \int_0^{2\pi}\frac{d\theta}{\pi}\sin^2(\theta/2)\,\frac{\sin[(j+1/2)\theta]}{\sin(\theta/2)}\,\frac{\sin[(j'+1/2)\theta]}{\sin(\theta/2)} = \delta_{jj'}\,. \tag{12.32}
$$

Solution 12.4 - The orthonormality of $\mathbf{SO}(3)$ characters: *If $j \neq j'$ we have:*

$$\begin{aligned}
\langle \chi^{(j)} | \chi^{(j')} \rangle &= \int_G dg \chi^{(j)}(g^{-1}) \chi^{(j')}(g) = \frac{1}{\pi} \int_0^{2\pi} \sin^2(\theta/2) \frac{\sin[(j+1/2)\theta]}{\sin(\theta/2)} \frac{\sin[(j'+1/2)\theta]}{\sin(\theta/2)} d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} \big(\cos\big((j-j')\theta\big) - \cos\big((j+j'+1)\theta\big)\big) \, d\theta \\
&= \frac{1}{2\pi} \left(\frac{1}{j-j'} \sin\big((j-j')\theta\big) - \frac{1}{j+j'+1} \sin\big((j+j'+1)\theta\big) \right) \Big|_0^{2\pi} d\theta = 0.
\end{aligned}$$

If $j = j'$ we have:

$$\begin{aligned}
\langle \chi^{(j)} | \chi^{(j)} \rangle &= \int_G dg \chi^{(j)}(g^{-1}) \chi^{(j)}(g) = \frac{1}{\pi} \int_0^{2\pi} \sin^2(\theta/2) \frac{\sin[(j+1/2)\theta]}{\sin(\theta/2)} \frac{\sin[(j+1/2)\theta]}{\sin(\theta/2)} d\theta \\
&= \frac{1}{2\pi} \int_0^{2\pi} (1 - \cos((2j+1)\theta)) \, d\theta \\
&= \frac{1}{2\pi} \left(2\pi - \frac{1}{2j+1} \sin((2j+1)\theta) \right) \Big|_0^{2\pi} d\theta = 1.
\end{aligned}$$

(T. Forrest Kieffer)

mathematical methods - week 13

Probability

Georgia Tech PHYS-6124

Homework HW #13 due Thursday, November 19, 2020

== 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

Bonus points
Exercise 13.1 *Lyapunov equation* 12 points

This week there are no required exercises. Whatever you do, you get bonus points.

edited November 12, 2020

Week 13 syllabus

November 10, 2020

This week's lectures are related to AWH Chapter 23 *Probability and Statistics* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- A summary of key concepts
  - ChaosBook appendix A20.1 Moments, cumulants
  - *Clip 1 - Averages, moments (45 min)*
  - *Clip 2 - Why a Gaussian? It's the maximum entropy distribution (8 min)*
- Why Gaussians again?
  - ChaosBook 33.2 Brownian diffusion
  - *Clip 3 - diffusion; Fokker-Planck density evolution (43 min)*
  - ChaosBook 33.3 Noisy trajectories
  - *Clip 4 - I don't like Langevin equation (3 min)*
- A glimpse of Orstein-Uhlenbeck, the "harmonic oscillator" of the theory of stochastic processes. And the one "Lyapunov" thing Lyapunov actually did:)
  - Noise is your friend
  - ChaosBook 33.4 Noisy maps
  - ChaosBook 33.5 All nonlinear noise is local
  - *Clip 5 - noise is your friend (23 min)*

Optional reading

Discussion 1 - Density is evaluated in configuration space, but the Laplacian is diagonalized in the Fourier space; a random walk from stochastic to quantum mechanics; Wiener path integrals; Schrödinger harmonic oscillator is imaginary time relative of the Ornstein-Uhlenbeck; Liouville theorem; Predrag's lecturing is a Gaussian process - on average you learn zero. (20 min)

13.1 Other sources

- MIT 16-90 Computational methods is a typical mathematical methods in engineering course. Probabilistic methods and optimization are discussed here.

Really going into the Ornstein-Uhlenbeck equation might take too much of your time, so this week we skip doing exercises, and if you are curious, and want to try your hand at solving exercise 13.1 *Lyapunov equation*, you probably should first skim through our lectures on the Ornstein-Uhlenbeck spectrum, Sect. 4.1 and Appen. B.1

here. Finally! we get something one expects from a math methods course, an example of why orthogonal polynomials are useful, in this case the Hermite polynomials :) .

The reason why I like this example is that again the standard 'physics' intuition misleads us. Brownian noise spreads with time as $\sqrt{t}$, but the diffusive dynamics of nonlinear flows is fundamentally different - instead of spreading, in the Ornstein-Uhlenbeck example the noise contained and balanced by the nonlinear dynamics.

- D. Lippolis and P. Cvitanović [4], *How well can one resolve the state space of a chaotic map?*; arXiv:0902.4269
- P. Cvitanović and D. Lippolis [1], *Knowing when to stop: How noise frees us from determinism*; arXiv:1206.5506
- J. M. Heninger, D. Lippolis and P. Cvitanović [3], *Neighborhoods of periodic orbits and the stationary distribution of a noisy chaotic system*; arXiv:1507.00462

Question 13.1. Henriette Roux asks

Q What percentage score on problem sets is a passing grade?

A That might still change, but currently it looks like 60% is good enough to pass the course. 70% for C, 80% for B, 90% for A. Very roughly - will alert you if this changes. Here is the percentage score as of week 10 in the 2019 course.

Question 13.2. Henriette Roux asks

Q How do I subscribe to the nonlinear and math physics and other seminars mailing lists?

A click here

References

[1] P. Cvitanović and D. Lippolis, Knowing when to stop: How noise frees us from determinism, in Let's Face Chaos through Nonlinear Dynamics, edited by M. Robnik and V. G. Romanovski (2012), pp. 82–126.

[2] Z. Gajić and M. Qureshi, *Lyapunov Matrix Equation in System Stability and Control* (Academic, New York, 1995).

[3] J. M. Heninger, P. Cvitanović, and D. Lippolis, "Neighborhoods of periodic orbits and the stationary distribution of a noisy chaotic system", Phys. Rev. E **92**, 062922 (2015).

[4] D. Lippolis and P. Cvitanović, "How well can one resolve the state space of a chaotic map?", Phys. Rev. Lett. **104**, 014101 (2010).

[5] H. Rome, "A direct solution to the linear variance equation of a time-invariant linear system", IEEE Trans. Automatic Control **14**, 592–593 (1969).

Exercises

13.1. **Lyapunov equation.** Consider the following system of ordinary differential equations,

$$\dot{Q} = A\,Q + Q\,A^{\top} + \Delta\,, \tag{13.1}$$

in which $\{Q, A, \Delta\} = \{Q(t), A(t), \Delta(t)\}$ are $[d\times d]$ matrix functions of time t through their dependence on a deterministic trajectory, $A(t) = A(x(t))$, etc., with stability matrix A and noise covariance matrix Δ given, and density covariance matrix Q sought. The superscript $(\ )^{\top}$ indicates the transpose of the matrix. Find the solution $Q(t)$, by taking the following steps:

(a) Write the solution in the form $Q(t) = J(t)[Q(0) + W(t)]J^{\top}(t)$, with Jacobian matrix $J(t)$ satisfying

$$\dot{J}(t) = A(t)\,J(t)\,, \qquad J(0) = \mathbf{1}\,, \tag{13.2}$$

with $\mathbf{1}$ the $[d\times d]$ identity matrix. The Jacobian matrix at time t

$$J(t) = \hat{T} e^{\int_0^t d\tau\, A(\tau)}\,, \tag{13.3}$$

where $\hat{T}$ denotes the 'time-ordering' operation, can be evaluated by integrating (13.2).

(b) Show that $W(t)$ satisfies

$$\dot{W} = \frac{1}{J}\,\Delta\,\frac{1}{J^{\top}}\,, \qquad W(0) = 0\,. \tag{13.4}$$

(c) Integrate (13.1) to obtain

$$Q(t) = J(t)\left[Q(0) + \int_0^t d\tau\,\frac{1}{J(\tau)}\,\Delta(\tau)\,\frac{1}{J^{\top}(\tau)}\right] J^{\top}(t)\,. \tag{13.5}$$

(d) Show that if $A(t)$ commutes with itself throughout the interval $0 \le \tau \le t$ then the time-ordering operation is redundant, and we have the explicit solution $J(t) = \exp\left\{\int_0^t d\tau\, A(\tau)\right\}$. Show that in this case the solution reduces to

$$Q(t) = J(t)\;Q(0)\,J(t)^{\top} + \int_0^t d\tau'\, e^{\int_{\tau'}^t d\tau\, A(t)}\,\Delta(\tau')\, e^{\int_{\tau'}^t d\tau\, A^{\top}(t)}\,. \tag{13.6}$$

(e) It is hard to imagine a time dependent $A(t) = A(x(t))$ that would be commuting. However, in the neighborhood of an equilibrium point x^* one can approximate the stability matrix with its time-independent linearization, $A = A(x^*)$. Show that in that case (13.3) reduces to

$$J(t) = e^{t\,A}\,,$$

and (13.6) to what?

Chapter 13 solutions: Probability

Solution 13.1 - Lyapunov equation. *The continuous Lyapunov equation is given by*

$$\dot{Q}(t) = A\,Q + Q\,A^\top + \Delta\,,$$

where $\{Q, A, \Delta\} = \{Q(t), A(t), \Delta(t)\}$ *are* $[d\times d]$ *matrix functions of time* t *through their dependence on a deterministic trajectory,* $A(t) = A(x(t))$*, etc., with stability matrix* A *and noise covariance matrix* Δ *given, and density covariance matrix* Q *sought. The solution is given by the initial covariance* $Q(0)$ *transported by the linearized flow, plus the noise experienced at any intermediate time, transported by the linearized flow,*

$$Q(t) = J(t,t_0)\,Q(0)\,J(t,t_0)^\top + \int_{,t_0}^{t} d\tau\,J(t,\tau)\,\Delta(\tau)\,J(t,\tau)^\top\,,$$

That this is indeed a solution is verified by differentiating both sides and using the equation for Jacobian matrix $J(t)$ *in terms of the stability matrix* A*,*

$$\dot{J}(t) = A(t)\,J(t)\,,\qquad J(0) = \mathbf{1}\,,$$

and $\mathbf{1}$ *is the* $[d\times d]$ *identity matrix. The solution exists for any finite time, as long as* $\Delta(\tau)$ *and* $A(\tau)$ *are matrices of bounded functions; the* $t\to\infty$ *limit, independent of the initial* $Q(0)$*, exists if* A *is strictly contracting,* $A < 0$*.*

If the flow is evaluated in the linearized neighborhood of an equilibrium point x_q*, the stability matrix* A *is a time-independent constant matrix, and the Jacobian matrix is simply* $J(t,t_0) = \exp((t-t_0)A)$ *. In that case it is possible to write an analytic, explicit formula [5] for* $Q(t)$ *in terms of stability exponents (eigenvalues of* A*). See Gajić and M. Qureshi [2]* Differential Lyapunov Equation, *Chapter 4 for this solution, and a discussion of numerical methods to solve time-dependent Lyapunov equations.*

mathematical methods - week 14

Math for experimentalists

Georgia Tech PHYS-6124

Homework HW #14 due Tuesday, November 24, 2020

== 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 14.1 *A "study" of stress and life satisfaction* a) to d) 10 points

Bonus points
Exercise 14.1 *A "study" of stress and life satisfaction* e) 4 points
Exercise 14.2 *Unbiased sample variance* 5 points
Exercise 14.3 *Standard error of the mean* 5 points
Exercise 14.4 *Bayesian statistics*, by Sara A. Solla 10 points

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

edited November 25, 2020

Week 14 syllabus

November 17, 2020

For this week's lectures read about the binomial theorem, Poisson and Gaussian distributions in AWH Chapter 23 *Probability and statistics* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run *the course playlist*.

- Sara A Solla: *Linear and nonlinear dimensionality reduction: applications to neural data, or - The unreasonable effectiveness of linear algebra. (1:08 hour)*

- Sara A. Solla's lecture notes: Neural recordings; Principal Components Analysis (PCA); Singular Value Decomposition (SVD); ISOMAP nonlinear dimensionality reduction; Multidimensional scaling.

- *Clip 2 Ignacio Taboada- Probability, Uncertainty, probability density functions, error matrix (20 min)*

- Ignacio Taboada — lecture notes.

- *Clip 3 Distributions: binomial, normal, uniform, moments, quantiles. Monte Carlo (why you need the uniform distribution) (19 min)*

- Ignacio Taboada — lecture notes.

- *Clip 4 central limit theorem (why you need normal dist) (2 min)*

- *Clip 5 Multi-dimensional PDFs (13 min)*

- *Clip 6 Error propagation: Covariances add! Covariances add! Covariances add! Three times :) (18 min)*

- Ignacio Taboada — lecture notes.

Optional reading

- Discussion 1, Sara A Solla: *A physicist turns neuroscientist. You can do anything. Progress in brain science. What is consciousness. What we know. The rest is speculation. Much speculation. (57 min)*

- *Rant - This is a kindergarten course. A professional should teach it, so I can teach you stuff that nobody teaches you here. (4 min)*

- *Sermon - Thanksgiving is upon us, don't be stupid (3 min)*

14.1 Optional reading: Bayesian statistics

Sara A. Solla

Natural sciences aim at abstracting general principles from the observation of natural phenomena. Such observations are always affected by instrumental restrictions and limited measurement time. The available information is thus imperfect and to some extent unreliable; scientists in general and physicists in particular thus have to face the task of extracting valid inferences from noisy and incomplete data. Bayesian probability theory provides a systematic framework for quantitative reasoning in the face of such uncertainty.

In this lecture (not given in the Fall 2020 course) we will focus on the problem of inferring a probabilistic relationship between a dependent and an independent variable. We will review the concepts of joint and conditional probability distributions, and justify the commonly adopted Gaussian assumption on the basis of maximal entropy arguments. We will state Bayes' theorem and discuss its application to the problem of integrating prior knowledge about the variables of interest with the information provided by the data in order to optimally update our knowledge about these variables. We will introduce and discuss Maximum Likelihood (ML) and Maximum A Posteriori (MAP) for optimal inference. These methods provide a solution to the problem of specifying optimal values for the parameters in a model for the relationship between independent and dependent variables. We will discuss the general formulation of this framework, and demonstrate that it validates the method of minimizing the sum-of-squared-errors in the case of Gaussian distributions.

- A quick but superficial read: Matthew R. Francis, *So what's all the fuss about Bayesian statistics?*
- Reading: Lyons [3], *Bayes and Frequentism: a particle physicist's perspective* (click here)

14.2 Statistics for experimentalists: desiderata

I have solicited advice from my experimental colleagues. You tell me how to cover this in less than two semesters :)

2012-09-24 Ignacio Taboada Cover least squares. To me, this is the absolute most basic thing you need to know about data fitting - and usually I use more advanced methods.

For a few things that particle and astroparticle people do often for hypothesis testing, read Li and Ma [2], *Analysis methods for results in gamma-ray astronomy*, and Feldman and Cousins [1] *Unified approach to the classical statistical analysis of small signals*. Both papers are too advanced to cover in this course, but the idea of hypothesis testing can be studied in simpler cases.

2012-09-24 Peter Dimon thoughts on how to teach math methods needed by experimentlists:

1. Probability theory
   (a) Inference
   (b) random walks
   (c) Conditional probability
   (d) Bayes rule (another look at diffusion)
   (e) Machlup has a classic paper on analysing simple on-off random spectrum. Hand out to students. (no Baysians use of information that you do not have) (Peter takes a dim view)
2. Fourier transforms
3. power spectrum - Wiener-Kitchen for correlation function
   (a) works for stationary system
   (b) useless on drifting system (tail can be due to drift only)
   (c) must check whether the data is stationary
4. measure: power spectrum, work in Fourier space
   (a) do this always in the lab
5. power spectra for processes: Brownian motion,
   (a) Langevin $\to$ get Lorenzian
   (b) connect to diffusion equation
6. they need to know:
   (a) need to know contour integral to get from Langevin power spectrum, to the correlation function
7. why is power spectrum Lorenzian - look at the tail $1/\omega^2$
   (a) because the cusp at small times that gives the tails
   (b) flat spectrum at origin gives long time lack of correlation
8. position is not stationary
   (a) diffusion
9. Green's function
   (a) δ fct $\to$ Gaussian + additivity
10. Nayquist theorem
    (a) sampling up to a Nayquist theorem (easy to prove)
11. Other processes:
    (a) what signal you expect for a given process
12. Fluctuation-dissipation theorem
    (a) connection to response function (lots of them measure that)
    (b) for Brownian motion power spectrum related to imaginary part of response function
13. Use *Numerical Recipes* (stupid on correlation functions)

   (a) zillion filters (murky subject)
   (b) Kalman (?)

14. (last 3 lecturs)
   (a) how to make a measurement
   (b) finite time sampling rates (be intelligent about it)

PS: Did I suggest all that? I thought I mentioned, like, three things.

Did you do the diffusion equation? It's an easy example for PDEs, Green's function, etc. And it has an unphysically infinite speed of information, so you can add a wave term to make it finite. This is called the Telegraph Equation (it was originally used to describe damping in transmission lines).

What about Navier-Stokes? There is a really cool exact solution (stationary) in two-dimensions called Jeffery-Hamel flow that involves elliptic functions and has a symmetry-breaking. (It's outlined in Landau and Lifshitz, *Fluid Dynamics*).

2012-09-24 Mike Schatz .

1. 1D bare minimum:
   (a) temporal signal, time series analysis
   (b) discrete Fourier transform, FFT in 1 and 2D - exercises
   (c) make finite set periodic
2. Image processing:
   (a) Fourier transforms, correlations,
   (b) convolution, particle tracking
3. PDEs in 2D (Matlab): will give it to Predrag (Predrag is still waiting)

References

[1] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals", Phys. Rev. D **57**, 3873–3889 (1998).

[2] T.-P. Li and Y.-Q. Ma, "Analysis methods for results in gamma-ray astronomy", Astrophys. J. **272**, 317–324 (1983).

[3] L. Lyons, "Bayes and Frequentism: a particle physicist's perspective", Contemporary Physics **54**, 1–16 (2013).

Participant	Stress score (X)	Life Satisfaction (Y)
1	11	7
2	25	1
3	19	4
4	7	9
5	23	2
6	6	8
7	11	8
8	22	3
9	25	3
10	10	6

Table 14.1: Stress vs. satisfaction for a sample of 10 individuals.

Exercises

14.1. **A "study" of stress and life satisfaction.**

Participants completed a measure on how stressed they were feeling (on a 1 to 30 scale) and a measure of how satisfied they felt with their lives (measured on a 1 to 10 scale). Participants' scores are given in table 14.1.

You can do this homework with pencil and paper, in Excel, Python, whatever:

a) Calculate the average stress and satisfaction.

b) Calculate the variance of each.

c) Plot Y vs. X.

d) Calculate the correlation coefficient matrix and indicate the value of the covariance.

e) Bonus: Read the article on "The Economist" (if you can get past the paywall), or, more seriously, D. Kahneman and A. Deaton -the 2002 Nobel Memorial Prize in Economic Sciences- about the correlation between income and happiness. Report on your conclusions.

14.2. **Unbiased sample variance.** Empirical estimates of the mean $\hat{\mu}$ and the variance $\hat{\sigma}^2$ are said to be "unbiased" if their expectations equal the exact values,

$$\mathbb{E}[\hat{\mu}] = \mu\,, \qquad \mathbb{E}[\hat{\sigma}^2] = \sigma^2\,. \tag{14.1}$$

(a) Verify that the empirical mean

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} a_i \tag{14.2}$$

is unbiased.

(b) Show that the naive empirical estimate for the *sample variance*

$$\bar{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(a_i - \hat{\mu})^2 = \frac{1}{N}\sum_{i=1}^{N} a_i^2 - \frac{1}{N^2}\left(\sum_{i=1}^{N} a_i\right)^2$$

is biased. Hint: note that in evaluating $\mathbb{E}[\cdots]$ you have to separate out the diagonal terms in

$$\left(\sum_{i=1}^{N} a_i\right)^2 = \sum_{i=1}^{N} a_i^2 + \sum_{i\neq j}^{N} a_i a_j \,. \tag{14.3}$$

(c) Show that the empirical estimate of form

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(a_i - \hat{\mu})^2 \,, \tag{14.4}$$

is unbiased.

(d) Is this empirical sample variance unbiased for any finite sample size, or is it unbiased only in the $n \to \infty$ limit?

Sara A. Solla

14.3. **Standard error of the mean.**

Now, estimate the empirical mean (14.2) of observable a by $j = 1, 2, \cdots, N$ attempts to estimate the mean $\hat{\mu}_j$, each based on M data samples

$$\hat{\mu}_j = \frac{1}{M}\sum_{i=1}^{M} a_i \,. \tag{14.5}$$

Every attempt yields a different sample mean.

(a) Argue that $\hat{\mu}_j$ itself is an idd random variable, with unbiased expectation $\mathbb{E}[\hat{\mu}] = \mu$.

(b) What is its variance

$$\text{Var}[\hat{\mu}] = \mathbb{E}[(\hat{\mu} - \mu)^2] = \mathbb{E}[\hat{\mu}^2] - \mu^2$$

as a function of variance expectation (14.1) and N, the number of $\hat{\mu}_j$ estimates? Hint; one way to do this is to repeat the calculations of exercise 14.2, this time for $\hat{\mu}_j$ rather than a_i.

(c) The quantity $\sqrt{\text{Var}[\hat{\mu}]} = \sigma/\sqrt{N}$ is called the *standard error of the mean* (SEM); it tells us that the accuracy of the determination of the mean μ. How does SEM decrease as the N, the number of estimate attempts, increases?

Sara A. Solla

14.4. **Bayes.** *Bayesian statistics*.

Chapter 14 solutions: Math for experimentalists

Solution 14.1 - A “study” of stress and life satisfaction.
(a) The average stress and satisfaction:

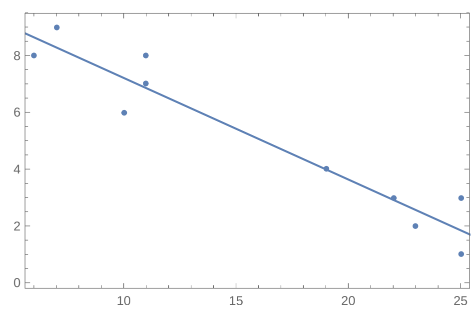
$$\hat{\mu}_X = \frac{1}{10}\sum_{i=1}^{10} X_i = 15.9\,, \qquad \hat{\mu}_Y = \frac{1}{10}\sum_{i=1}^{10} Y_i = 5.1\,.$$

(b) The unbiased *variances and standard deviations:*

$$\hat{\sigma}_X^2 = \frac{1}{10-1}\sum_{i=1}^{10}(X_i-\mu_X)^2 = 58.1\,, \quad \hat{\sigma}_X = 7.6\,.$$

$$\hat{\sigma}_Y^2 = \frac{1}{10-1}\sum_{i=1}^{10}(Y_i-\mu_Y)^2 = 8.1\,, \quad \hat{\sigma}_Y = 2.8\,.$$

(c)



The averages (a) are misleading - the subjects are either unhappy or happy, there is nobody in between. The standard deviations (b) of such bimodal distribution are not helpful either, as they are measuring deviations from the non-existent average participant. However, the linear fit

$$Y = 11 - .36\,X \tag{14.6}$$

is pretty good.

(d) The stressed / the satisfied covariance is

$$V_{XY} = \frac{1}{10-1}\sum_{i=1}^{10}(X_i-\hat{\sigma}_X)(Y_i-\hat{\sigma}_Y) = -20.8\,.$$

The ellipsoid given by the covariance matrix

$$V = \begin{pmatrix} \hat{\sigma}_X^2 & V_{XY} \\ V_{XY} & \hat{\sigma}_Y^2 \end{pmatrix}$$

singular values (square roots of eigenvalues) and eigenvectors

$$\{\sigma_1, \sigma_2\} = \{8.10, 0.76\}\ : \quad \mathbf{e}_{(1)} = (0.94, -0.34)\,, \quad \mathbf{e}_{(2)} = (0.34, 0.94)$$

gives a pretty good description of the data, aligned along $\mathbf{e}_{(1)}$ *(of slope close to the linear fit (14.6)), with small fluctuations along* $\mathbf{e}_{(2)}$*. The only problem is that we are plotting lemons vs. roses.*

For that reason one considers instead correlation coefficients matrix

$$\textit{Corr}(X,Y) = \begin{pmatrix} 1 & \rho_{XY} \\ \rho_{XY} & 1 \end{pmatrix} \,,$$

where

$$\rho_{XY} = \frac{V_{XY}}{\hat{\sigma}_X \hat{\sigma}_Y} = -0.9573 \,.$$

Its singular values, eigenvectors are a dimensionless *least-squares fit to the data, with the ellipsoid's principal axes along the diagonals*

$$\{\sigma_1, \sigma_2\} = \{1.40, 0.207\} \;:\quad \mathbf{e}_{(1)} = (1/\sqrt{2}, -1/\sqrt{2}) \,,\quad \mathbf{e}_{(2)} = (1/\sqrt{2}, 1/\sqrt{2}) \,.$$

The transformation from the covariance matrix to the correlations matrix is not a similarity transformation, so I do not see why is it legal to describe the data covariances by the correlation matrix singular values.

P. Cvitanović

Solution 14.2 - Unbiased sample variance. *See ChaosBook Example A20.3*

Solution 14.3 - Standard error of the mean. *See ChaosBook Example A20.4*

Solution 14.4 - Bayes. *No solution available.*

mathematical methods - week 15

What is 'chaos'?

Georgia Tech PHYS-6124

No Homework HW #15 due never, not graded this week

---

edited November 25, 2020

Week 15 syllabus

Tuesday, November 24, 2020

The fastest way to watch any week’s lecture videos is by letting YouTube run *the course playlist*.

- *Clip 1 - what is ‘chaos’? how an applied mathematicians thinks about it (12 min)*
- *Clip 2 - chaos for field theorists, 3rd millennium, lattice formulation: exponentially many distinct walks through Bernoullistan. (22 min)*
- *Clip 3 - periodic orbit theory. How come Hill determinant counts periodic points? (12 min)*
- *Clip 4 - chaos for a field theorist : think globally, act locally (6 min)*
- coin toss

Optional reading

- *Spatiotemporal cat and the end of time (2 min)*
- *Spatiotemporal cat - a chaotic field theory* (55 min seminar)
- *Herding cats: a chaotic field theory (1h 10 min seminar)*
- Turbulence in spacetime : website, talks
- *Rutgers seminar Q & A - Harper’s model; why periodic boundary conditions; Spatiotemporal chaos for continuous theories; chaos as anti-integrability.* (7 min)
- *Rant - E&M exam traumatized; on the necessity of E&M exams; Bologna and the necessity of child abuse, generation to generation; no Jonestown Colony, please; Jaques Laskar’s miracle - we are here because we have our Moon; working for the “industry", by intimidating children with zeta functions; working for organized crime AKA hedge funds. (23 min)*

mathematical methods - week 16

The epilogue, and what next?

If I had had more time, I would have written less
— Blaise Pascal, a remark made to a correspondent

Student evaluations November 2019: 11 out of 21 students filled in the questionnaire, with a bimodal distribution, typically 4 at the “Exceptional” end, and 4 at the other, “Very Poor” end.

Positive evaluations along lines of “This course’s best aspect was the breadth of material covered,” “This expended effort for this course was proportional to the amount of material I wanted to learn,” are not included in what follows; we focus on criticisms, and how to improve the course for the future students.

Several students have written in depth about the problems with the course. These valuable comments are merged (for student privacy) and addressed below, to assist future instructors in preparing this annual physics Fall course aimed at the incoming graduate students.

Structure of the course comments :

> The course is taught as a very rushed sweep of complicated mathematical concepts, trying to go much **more in-depth with the topics than we had time for**, and, as result, I understood almost nothing. The beginning of the course was on topics some had seen in other physics or math courses, but from the start the course often felt inaccessible if you did not already have some familiarity with whichever topic was being lectured on. By the last third, we faced quite advanced math topics that only a student with a degree in mathematics would have possibly seen, so attempting to research these topics without any background or any semblance of a direction to start was mind-numbingly frustrating at best, and a complete waste of time at worst.
>
> There is no **consistent textbook** for the course. The recommended multiple texts for each individual topic lead to a disorienting mess of information hunting that ends up with the student cutting their losses and giving up.
>
> The **homework** was extremely abstractly related to the lecture and did not touch upon the aspects we talked about in class. Homework problems varied between

very easy to wildly difficult (or difficultly worded). One had to research for hours to figure out the material necessary to do the homework, as it was never addressed in class nor was there any dedicated textbook on which to rely.

The most important issue of this course is **consistency**, the severe lack of correlation between lecture, study, and homework. Lectures are inconsistent with homework assignments, and often one finds that the information required to do a problem is revealed the same day the homework is due, maybe even days later.

The **workload** for this course was not appropriate for a pass/fail class. It was unclear what grade constitutes a pass until about 2/3 through the course. The best part of the course was that it was **pass/fail**. This reduced the overall stress of the ineffectiveness of the course, so it did not impact my other courses at all.

With the **lecturing to the board**, *many* minutes can pass with your hand up before the instructor turns to the class to notice you might have a question.

If I'm getting next to nothing of value from **lecture** and have to do all this research on my own just to stand a chance at completing a homework problem, why show up? I was so lost in this course almost all the time that I eventually found it useless to attend class, and learned much more by reading textbooks not assigned by the course, in order to hopefully glean something useful to solve the esoteric problems. These severe issues encouraged **skipping class**; minimal practice/learning was actually achieved.

Action :

In the first semester of graduate school, and as a required course, the incoming class of graduate students needs a traditional, clearly structured textbook course, with clearly spelled out expectations for each learning step, and much better learning practices than what this version of the course offered.

Use **one consistent textbook** that guides the entire course (lecture, study, and homework).

Assign **homework** directly relevant to what students learned in class (the lecture taking place before the homework is due), and requiring no outside research. Have someone read over the homework questions before they are assigned to make sure that what is being asked is clear. For **advanced topics**, make homework **optional**.

State on course homepage the **grade** required for a pass.

Keep the course **pass/fail** if it remains a required course. (However, starting Fall 2020 Math Methods will no longer be required for all 1st year physics grad students. It will be an elective, letter grade course, not pass/fail.)

Have a deep look at how this course was taught and **what students found difficult**; try to relate to the average learner. Understand better what a student is asking. To facilitate that, at the beginning of a class go through a bullet-point list of concepts covered in the previous class, ask for questions related to each. On the days the homework is due, go through problems, ask what difficulties were there with each one.

Integrate the students into teaching by asking them more questions. Allow for more time to fully discuss the topics. Give clear explanations, not slowing down but actually taking ideas step by step, with students contributing.

Instructor comment :

This is one of the first graduate school courses encountered by incoming students, of vastly different backgrounds. Not only should I have not assumed a high level of prior knowledge, but at this point students do not need to be taught in a style that reflects the ways knowledge is acquired in actual research, involving multiple sources, approaches, and notations.

Advanced approach is better suited to second or later years of graduate study. Indeed, the School of Physics plans to offer such research oriented course (PHYS 4740/6740, to be initially taught by Grigoriev) as an advanced elective.

Course content comments :

The course topics were very interesting, and it is a shame that there was not enough time to explore them in depth. Great balance on the wide scope and enough difficulty of the course. Such a class is very useful and I would still be interested to learn more about the topics covered.

Cover less group theory.

Dedicate one week to the calculus of variations.

Action :

Teach fewer topics; spend more time on each topic.

Instructor comment :

There were no detailed students comments on course content, except the two listed above.

The choice of course topics was quite different from what is covered in traditional mathematical methods courses, in order to reflect the current research in the School of Physics and in the engineering schools; fewer topics preparatory to E&M and QM courses, more topics related to physics of living systems, soft condensed matter and the analysis of experimental data. I am not aware of any textbook that covers this ground.