Phys 6124 zipped!

World Wide Quest to Tame Math Methods

Predrag Cvitanović

January 2, 2020

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 the (unrecorded) 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 are the two weeks dedicated to data analysis. Let me know whether you would have preferred the more traditional math methods fare, like Bessels and such.

If you have taken this course live, you might have noticed a pattern: Every week we start with something obvious that you already know, let mathematics lead us on, and then suddenly end up someplace surprising and highly non-intuitive.

And indeed, in the last lecture (that never took place), we turn Coleman on his head, and abandon harmonic potential for the *inverted* harmonic potential, "spatiotemporal cats" and chaos, arXiv:1912.02940.

Contents

mathematical methods - week 1

Linear algebra

Georgia Tech PHYS-6124

Homework HW #1 due Monday, August 26, 2019

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

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

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

edited September 9, 2019

Week 1 syllabus

Monday, August 19, 2019

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

- Sect. 1.2 *Matrix-valued functions*
  Grigoriev notes pages 6.1, 6.2 (skip *1. Mechanics inertia tensor*), 6.4, 6.5 up to "right and left eigenvectors."

- sect. 1.3 *A linear diversion*
  There are two representations of exponential of constant matrix, the Taylor series and the compound interest (Euler product) formulas (1.23). If the matrix (for example, the Hamiltonian) changes with time, the exponential has to be time-ordered. The Taylor series generalizes to the nested integrals formula (1.50), and the Euler product to time-ordered product (1.24). The first is used in formal quantum-mechanical calculations, the second in practical, numerical calculations.

- sect. 1.4 *Eigenvalues and eigenvectors*
  Hamilton-Cayley equation, projection operators (1.34), any matrix function is evaluated by spectral decomposition (1.37). Work through example 1.3.

- Optional reading: Stone & Goldbart Appendix A; Arfken & Weber [1] (click here) Chapter 3

- Feel free to add your topics suggestions at any time: suggestions by students and faculty.

- you got a new idea?

Question 1.1. Henriette Roux find 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 [1].

1.1 Literature

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

1.2 Matrix-valued functions

What is a matrix?
—Werner Heisenberg (1925)

What is the matrix?
—-Keanu Reeves (1999)

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

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

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

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

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

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

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

Derivatives of products of matrices are evaluated by the chain rule

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

A matrix and its derivative matrix in general do not commute

$$\frac{d}{dx}\mathbf{A}^2 = \frac{d\mathbf{A}}{dx}\mathbf{A} + \mathbf{A}\frac{d\mathbf{A}}{dx}\,. \tag{1.5}$$

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

$$\frac{d}{dx}\mathbf{A}^{-1} = -\frac{1}{\mathbf{A}}\frac{d\mathbf{A}}{dx}\frac{1}{\mathbf{A}}\,. \tag{1.6}$$

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

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

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}^k\,, \qquad \mathbf{A}^0 = \mathbf{1} \tag{1.7}$$

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

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

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

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

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

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

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

$$\det\left(e^{\mathbf{A}}\right) = \lim_{N\to\infty} \left(\det\,(\mathbf{1} + \mathbf{A}/N)\right)^N \,.$$

To the leading order in $1/N$

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

hence

$$\det e^{\mathbf{A}} = \lim_{N\to\infty}\left(1 + \frac{1}{N}\mathrm{tr}\,\mathbf{A} + O(N^{-2})\right)^N = \lim_{N\to\infty}\left(1 + \frac{\mathrm{tr}\,\mathbf{A}}{N}\right)^N = e^{\mathrm{tr}\,\mathbf{A}} \tag{1.10}$$

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

$$\ln\det\mathbf{M} = \mathrm{tr}\,\ln\mathbf{M}\,, \tag{1.11}$$

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

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

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \mathbf{B} + t[\mathbf{A},\mathbf{B}] + \frac{t^2}{2}[\mathbf{A},[\mathbf{A},\mathbf{B}]] + \frac{t^3}{3!}[\mathbf{A},[\mathbf{A},[\mathbf{A},\mathbf{B}]]] + \cdots \tag{1.12}$$

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

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

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

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

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

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

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

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

where we have set $\hbar = 1$.

1.3 A linear diversion

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

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

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

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

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

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

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

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

$$\frac{x(\delta t) - x(0)}{\delta t} = A\,x(0)\,. \tag{1.19}$$

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

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

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

Now that we have some feeling for the qualitative behavior of a linear flow, we are ready to return to the nonlinear case. Consider an infinitesimal perturbation of the initial state, $x_0 + \delta x(x_0, 0)$. How do we compute the exponential (1.20) that describes linearized perturbation $\delta x(x_0, t)$?

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

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

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

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

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

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

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

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

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

where $\mathbf{T}$ stands for time-ordered integration, *defined* as the continuum limit of the successive multiplications (1.24). This integral formula for J^t is the finite time companion of the differential definition exercise 1.3

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

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

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

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

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

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

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

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

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

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

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

From (1.29) it follows that

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

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

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

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

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

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

$$P_i = \prod_{j\neq i}\frac{\mathbf{M}-\lambda_j\mathbf{1}}{\lambda_i-\lambda_j}\,, \tag{1.32}$$

which are *orthogonal* and *complete*:

$$P_i P_j = \delta_{ij} P_j \,, \quad (\text{no sum on } j)\,, \qquad \sum_{i=1}^{r} P_i = \mathbf{1}\,, \tag{1.33}$$

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

$$(\mathbf{M} - \lambda_j \mathbf{1}) P_j = P_j (\mathbf{M} - \lambda_j \mathbf{1}) = 0\,, \tag{1.34}$$

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 (1.34) that λ_i is the eigenvalue of $\mathbf{M}$ on P_i subspace:

$$\mathbf{M}\, P_i = \lambda_i P_i \qquad (\text{no sum on } i)\,. \tag{1.35}$$

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

$$\mathbf{M} = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_d P_d\,. \tag{1.36}$$

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*

$$f(\mathbf{M}) = \sum_i f(\lambda_i) P_i\,. \tag{1.37}$$

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

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

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

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

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

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

We distinguish two cases:

$\mathbf{M}$ can be brought to diagonal form. *The characteristic equation (1.39) can be replaced by the minimal polynomial,*

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

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

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

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

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*

$$\mathbf{M} = \left[\begin{array}{cc} M_{11} & M_{12} \\ M_{21} & M_{22} \end{array}\right].$$

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

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

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

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

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

$$P_1 = \frac{\mathbf{M} - \lambda_2 \mathbf{1}}{\lambda_1 - \lambda_2}\,, \qquad P_2 = \frac{\mathbf{M} - \lambda_1 \mathbf{1}}{\lambda_2 - \lambda_1}\,, \qquad \lambda_1 \neq \lambda_2\,. \tag{1.43}$$

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

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

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

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

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

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

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 $[2\times 2]$ 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 (1.42):

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

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 *(1.43)*

$$P_1 = \frac{1}{4}(\mathbf{M}-\mathbf{1}) = \frac{1}{4}\left[\begin{array}{cc} 3 & 1 \\ 3 & 1 \end{array}\right] \tag{1.45}$$

$$P_2 = \frac{1}{4}(\mathbf{M}-5\cdot\mathbf{1}) = \frac{1}{4}\left[\begin{array}{cc} 1 & -1 \\ -3 & 3 \end{array}\right]. \tag{1.46}$$

Matrices P_i are orthonormal and complete, The dimension of the ith subspace is given by $d_i = \operatorname{tr} P_i$; 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 (1.38), by inspection. (Continued in example 2.2.)

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

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

If A is diagonalizable, $A = FDF^{-1}$, where D is the diagonal matrix of the eigenvalues of A and F is the matrix of corresponding eigenvectors, the result is simple: $A^n = (FDF^{-1})(FDF^{-1})\dots(FDF^{-1}) = FD^nF^{-1}$. 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 $[2\times 2]$ matrices. For any linear system in $\mathbb{R}^2$, there is a similarity transformation

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

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

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

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

$$e^{Bt} = \left[\begin{array}{cc} e^{\lambda t} & 0 \\ 0 & e^{\mu t} \end{array}\right], \qquad e^{Bt} = e^{\lambda t}\left[\begin{array}{cc} 1 & t \\ 0 & 1 \end{array}\right], \qquad e^{Bt} = e^{at}\left[\begin{array}{cc} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{array}\right],$$

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

1.4.1 Yes, but how do you really do it?

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

If two eigenvalues form a complex conjugate pair, $\{\lambda_k, \lambda_{k+1}\} = \{\mu + i\omega, \mu - i\omega\}$, they are in a sense degenerate: while a real λ_k characterizes a motion along a line, a complex λ_k characterizes a spiralling motion in a plane. We determine this plane by replacing the corresponding complex eigenvectors by their real and imaginary parts, $\{\mathbf{e}^{(k)}, \mathbf{e}^{(k+1)}\} \to \{\mathrm{Re}\,\mathbf{e}^{(k)}, \mathrm{Im}\,\mathbf{e}^{(k)}\}$, or, in terms of projection operators:

$$P_k = \frac{1}{2}(R + iQ)\,, \qquad P_{k+1} = P_k^*\,,$$

where $R = P_k + P_{k+1}$ is the subspace decomposed by the kth complex eigenvalue pair, and $Q = (P_k - P_{k+1})/i$, both matrices with real elements. Substitution

$$\begin{bmatrix} P_k \\ P_{k+1} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}\begin{bmatrix} R \\ Q \end{bmatrix}\,,$$

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

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

which brings the $\lambda_k P_k + \lambda_{k+1} P_{k+1}$ complex eigenvalue pair in the spectral decomposition into the real form,

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

$$[P_k, P_{k+1}]\begin{bmatrix} \lambda & 0 \\ 0 & \lambda^* \end{bmatrix}\begin{bmatrix} P_k \\ P_{k+1} \end{bmatrix} = [R, Q]\begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix}\begin{bmatrix} R \\ Q \end{bmatrix}\,, \tag{1.48}$$

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

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

Commentary

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

References

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

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

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

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

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

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

[7] W. G. Harter and N. dos Santos, "Double-group theory on the half-shell and the two-level system. I. Rotation and half-integral spin states", Amer. J. Phys. **46**, 251–263 (1978).

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

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

Exercises

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

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

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

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

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

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

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

may be written as

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

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

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

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

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

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

$$P = P_k\,, \qquad P^* = P_{k+1}\,,$$

where we denote P_k by P for notational brevity.

(b) P can be written as

$$P = \frac{1}{2}(R + iQ)\,,$$

where $R = P_k + P_{k+1}$ and Q are matrices with real elements.

(c)

$$\begin{bmatrix} P_k \\ P_{k+1} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}\begin{bmatrix} R \\ Q \end{bmatrix}.$$

(d) The $\cdots + \lambda_k P_k + \lambda_k^* P_{k+1} + \cdots$ complex eigenvalue pair in the spectral decomposition (1.36) is now replaced by a real $[2\times 2]$ matrix

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

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

mathematical methods - week 2

Eigenvalue problems

Georgia Tech PHYS-6124

Homework HW #2 due Wednesday, September 4, 2019

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort

Exercise 2.1 *Three masses on a loop* 8 points

Bonus points
Exercise 2.2 *A simple stable/unstable manifolds pair* 4 points

edited September 9, 2019

Week 2 syllabus

Monday, August 26, 2019

- Intro to normal modes: example 2.1 *Vibrations of a classical CO_2 molecule*
- Work through Grigoriev notes 8 Normal modes
- Linear stability : example 2.2 *Stable/unstable manifolds*
- Optional reading: Stone & Goldbart Appendix A; Arfken & Weber Arfken and Weber [1] (click here) Chapter 3
- Optional: Work through Grigoriev notes p. 6.6 crankshaft;

The big idea 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 (1.37) 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.

2.1 Normal modes

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 in figure 2.1. Newton's second law says*

$$\begin{aligned}
\ddot{x}_1 &= -\frac{k}{M}(x_1 - x_2) \\
\ddot{x}_2 &= -\frac{k}{m}(x_2 - x_3) - \frac{k}{m}(x_2 - x_1) \\
\ddot{x}_3 &= -\frac{k}{M}(x_3 - x_2)\,.
\end{aligned} \tag{2.3}$$

The normal modes, with time dependence $x_j(t) = x_j \exp(it\omega)$, are the common frequency ω vibrations that satisfy (2.3),

$$\mathbf{H}\mathbf{x} = \begin{pmatrix} A & -A & 0 \\ -a & 2\,a & -a \\ 0 & -A & A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \omega^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\,, \tag{2.4}$$

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.

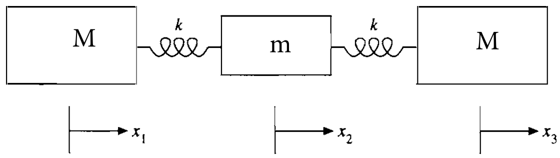


Figure 2.1: A classical colinear CO_2 molecule [1].

You might be tempted to stick this $[3\times3]$ *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 (2.3) 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* $\mathbf{H}$*,*

$$\sigma \;=\; \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} , \qquad \sigma\mathbf{H} = \mathbf{H}\sigma \;=\; \begin{pmatrix} 0 & -A & A \\ -a & 2\,a & -a \\ A & -A & 0 \end{pmatrix} . \tag{2.5}$$

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 *(1.43) are*

$$P_+ = \frac{1}{2}(\mathbf{1}+\sigma) , \qquad P_- = \frac{1}{2}(\mathbf{1}-\sigma) . \tag{2.6}$$

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

$$d_+ = 2 , \qquad d_- = 1 . \tag{2.7}$$

As σ *and* $\mathbf{H}$ *commute, we can now use spectral decomposition (1.37) to block-diagonalize* $\mathbf{H}$ *to a 1-dimensional and a 2-dimensional matrix.*

On the 1-dimensional antisymmetric subspace, the trace of a $[1\times1]$ *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 (2.7) explicitly. Its non-vanishing columns are proportional to the sole eigenvector*

$$P_- \;=\; \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \;\Rightarrow\; \mathbf{e}^{(-)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} . \tag{2.8}$$

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 (2.3) do not change their form,*

$$\mathbf{H}\,\mathbf{x} \;=\; \mathbf{H}\,\mathbf{x} + \mathbf{H}\,\delta\mathbf{x} \;=\; \omega^2\mathbf{x} \quad\Rightarrow\quad \mathbf{H}\,\delta\mathbf{x} \;=\; 0\,. \tag{2.9}$$

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 (2.5), 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.*

2.2 Stable/unstable manifolds

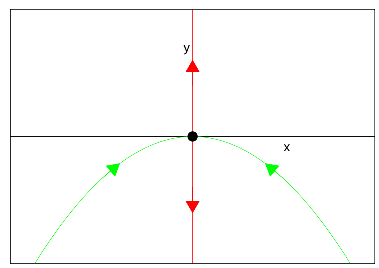


Figure 2.2: The stable/unstable manifolds of the equilibrium $(x_q, x_q) = (0, 0)$ of 2-dimensional flow (2.10).

Example 2.2. ***A simple stable/unstable manifolds pair:*** *Consider the 2-dimensional ODE system*

$$\frac{dx}{dt} = -x\,, \quad \frac{dy}{dt} = y + x^2\,, \tag{2.10}$$

The flow through a point $x(0) = x_0, y(0) = y_0$ *can be integrated*

$$x(t) = x_0\,e^{-t}, \quad y(t) = (y_0 + x_0^2/3)\,e^t - x_0^2\,e^{-2t}/3\,. \tag{2.11}$$

Linear stability of the flow is described by the stability matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 2x & 1 \end{pmatrix}. \tag{2.12}$$

The flow is hyperbolic, with a real expanding/contracting eigenvalue pair $\lambda_1 = 1$*,* $\lambda_2 = -1$*, and area preserving. The right eigenvectors at the point* (x, y)

$$\mathbf{e}^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathbf{e}^{(2)} = \begin{pmatrix} 1 \\ -x \end{pmatrix}. \tag{2.13}$$

can be obtained by acting with the projection operators (see example 1.2 Decomposition of 2-dimensional vector spaces*)*

$$P_i = \frac{\mathbf{A} - \lambda_j \mathbf{1}}{\lambda_i - \lambda_j} \;: \qquad P_1 = \begin{bmatrix} 0 & 0 \\ x & 1 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 1 & 0 \\ -x & 0 \end{bmatrix} \tag{2.14}$$

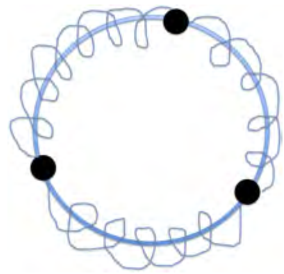


Figure 2.3: Three identical masses are constrained to move on a hoop, connected by three identical springs such that the system wraps completely around the hoop. Find the normal modes.

on an arbitrary vector. Matrices P_i are orthonormal and complete.

The flow has a degenerate pair of equilibria at $(x_q, y_q) = (0, 0)$, with eigenvalues (stability exponents), $\lambda_1 = 1$, $\lambda_2 = -1$, eigenvectors $\mathbf{e}^{(1)} = (0, 1)$, $\mathbf{e}^{(2)} = (1, 0)$. The unstable manifold is the y axis, and the stable manifold is given by (see figure 2.2)

$$y_0 + \frac{1}{3}x_0^2 = 0 \Rightarrow y(t) + \frac{1}{3}x(t)^2 = 0\,. \tag{2.15}$$

(N. Lebovitz)

References

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

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. **A simple stable/unstable manifolds pair.** Integrate flow (2.10), verify (2.11). Check that the projection matrices P_i (2.14) are orthonormal and complete. Use them to construct right and left eigenvectors; check that they are mutually orthogonal. Explain why is (2.15) the equation for the stable manifold. (N. Lebovitz)

mathematical methods - week 3

Go with the flow

Georgia Tech PHYS-6124

Homework HW #3 due Monday, September 9, 2019

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort

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 4, 2019

Week 3 syllabus

Wednesday, September 4, 2019

- Sect. 3.1 *Linear flows*
- Sect. 3.2 *Stability of linear flows*
- Optional reading: Sect. 3.3 *Nonlinear flows*
- Sect. 3.4 *Optional listening*

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.

3.1 Linear flows

Linear is good, nonlinear is bad.
—Jean Bellissard

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

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

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

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.

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 (1.7):

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

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).

3.2 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.*

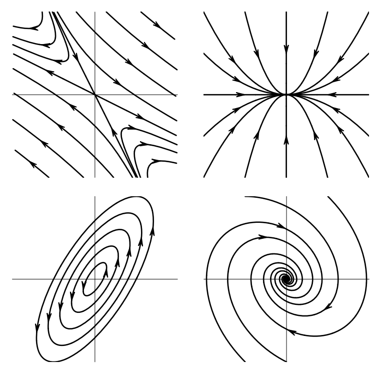


Figure 3.1: Streamlines for several typical 2-dimensional flows: saddle (hyperbolic), in node (attracting), center (elliptic), in spiral.

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 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*

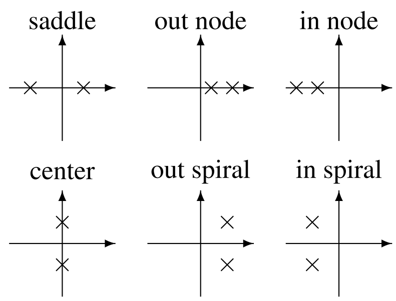


Figure 3.2: Qualitatively distinct types of exponents $\{\lambda_1, \lambda_2\}$ of a $[2\times 2]$ Jacobian matrix.

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.

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

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)

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

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

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

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 $10^{-2}T$).

3.3 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.

We have already made a foray into nonlinearity in example 2.2 *A simple stable/unstable manifolds pair*, but that was a bit of a cheat - it is really an example of a non-autonomous flow in variable $y(t)$, driven by external forcing by $x(t)$. Duffing flow of

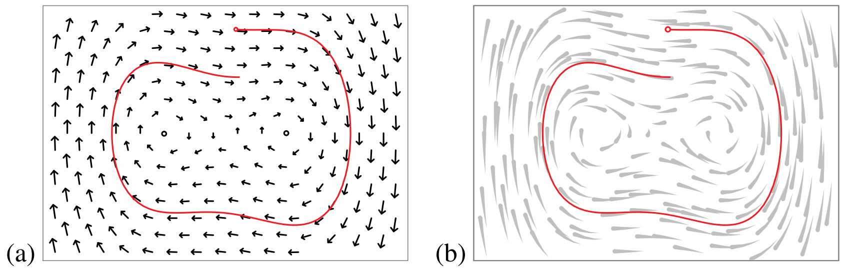


Figure 3.3: (a) The 2-dimensional vector field for the Duffing system (3.14), together with a short trajectory segment. (b) The flow lines. Each 'comet' represents the same time interval of a trajectory, starting at the tail and ending at the head. The longer the comet, the faster the flow in that region.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

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* $\mathcal{M}$*.*

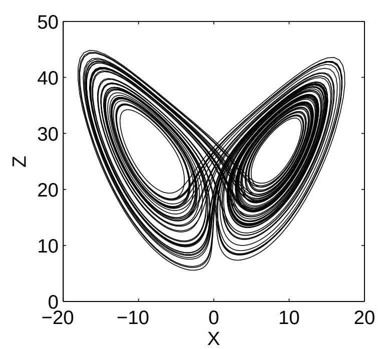


Figure 3.4: Lorenz "butterfly" strange attractor. (From ChaosBook [1])

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.

3.4 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.

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 .

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, 2019).

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 use any integration routine to integrate numerically, 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 (1.42), 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

mathematical methods - week 4

Complex differentiation

Georgia Tech PHYS-6124

Homework HW #4 due Monday, September 16, 2019

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort
== if you are LaTeXing, here is the 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 22, 2019

Week 4 syllabus

Monday, September 9, 2019

Complex variables; History; algebraic and geometric insights; De Moivre's formula; roots of unity; functions of complex variables as mappings; differentiation of complex functions; Cauchy-Riemann conditions; analytic functions; Riemann surfaces; conformal mappings.

Mon Goldbart pages 1/10 - 1/120

Wed Goldbart pages 1/130 - 1/140 (skipped: Riemann sphere)
Goldbart pages 1/200 - 1/260 (complex differentiation)

Fri Goldbart pages 1/270 - 1/340

Optional reading

- Grigoriev notes pages 2.1 - 2.3
- Stone and Goldbart [4] (click here) Chapter 17 sect. 17.1
- Arfken and Weber [2] (click here) Chapter 6 sects. 6.1 - 6.2,
- Ahlfors [1] (click here)
- Needham [3] (click here)

From now on, copyright-protected references are on a password protected site. What password? Have your ears up in the class; the password will be posted on the Canvas for a week or so, so remember to write it down.

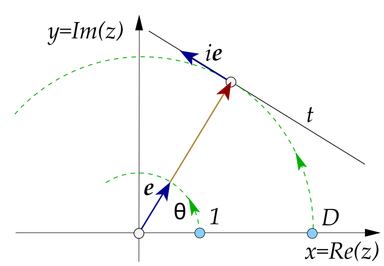


Figure 4.1: A unit vector $\mathbf{e}$ multiplied by a real number D traces out a circle of points in the complex plane. Multiplication by the imaginary unit i rotates a complex vector by 90^0, so $D\mathbf{e}+it\mathbf{e}$ is a tangent to this circle, a line parametrized by a real number t.

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 and H. J. Weber, *Mathematical Methods for Physicists: A Comprehensive Guide*, 6th ed. (Academic, New York, 2005).

[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, 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

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

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

(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 $|a| \leq |c|$. 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 $\{a_j, b_j\}$

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

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

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

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

(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 $1 - \omega^H + \omega^{2H} - \cdots + (-1)^{N-1}\omega^{(N-1)H}$?

mathematical methods - week 5

Complex integration

Georgia Tech PHYS-6124

Homework HW #5 due Monday, September 23, 2019

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

Exercise 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 22, 2019

Week 5 syllabus

September 16, 2019

Mon Goldbart pages 1/400 - 1/580

Wed Goldbart pages 2/10 - 2/96 (contour integration)

Fri Goldbart pages 2/130 -2/150; 2/270 - 2/280 (derivation of Cauchy Theorem)

Optional reading

- Goldbart pages 3/10 - 3/140 (Cauchy contour integral)
- Grigoriev pages 3.1 - 3.3 (Cauchy contour integral)
- Arfken and Weber [1] (click here) Chapter 6 sects. 6.3 - 6.4, on Cauchy contour integral

Question 5.1. Henriette Roux asks
Q What do you mean when you write "Determine the possibilities" in exercise 5.1 (b)?
A Fair enough. I rewrote the text in the exercise.

References

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

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) $\operatorname{Re} f(z) = 0$; (ii) $|f(z)| = \text{const.}$; (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 $A - B \le \operatorname{Re} z \le A + B$ and $-B \le \operatorname{Im} z \le B$ 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).

mathematical methods - week 6

Cauchy - applications

Georgia Tech PHYS-6124

Homework HW #6 due Monday, September 30, 2019

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

Exercise 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 25, 2019

Week 6 syllabus

September 23, 2019

Mon Goldbart pages 3/10 - 3/30; 3/60 - 3/70 (Cauchy integral formula)

Wed Goldbart pages 3/80 - 3/110 (singularities; Laurent series)
Grigoriev pages 3.4 - 3.5b (evaluation of integrals)

Fri Grigoriev pages 3.4 - 3.5b (evaluation of integrals)
Goldbart pages 4/10 - 4/100 (linear response)

Optional reading

- Arfken and Weber [1] (click here) Chapter 6 sects. 6.3 - 6.4, on Cauchy contour integral
- Arfken and Weber [1] Chapter 6 sects. 6.5 - 6.8, on Laurent expansion, cuts, mappings
- Arfken and Weber [1] (click here) Chapter 7 sects. 7.1 - 7.2, on residues
- Stone and Goldbart [2] (click here) Chapter 17 sect. 17.2 - 17.4

Question 6.1. Henriette Roux had asked

Q You made us do exercise 4.5, but you did not cover this in class? What's up with that? I left it blank!

A Mhm. Check the discussion of this problem in the updated week 4 notes.

References

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

[2] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students* (Cambridge Univ. Press, Cambridge, 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 $I = \int_0^\infty \exp\left(ix^2\right) dx$. To do this, consider the contour integral $I_R = \int_{C(R)} \exp\left(iz^2\right) dz$, 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 $\lim_{R\to\infty} \int_{C_1(R)} \exp\left(iz^2\right) dz = 0$, 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

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

and, from the well-known result of real integration $\int_0^\infty \exp\left(-x^2\right) dx = \sqrt{\pi}/2$, 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 $I_R = \int_{C(R)} \exp\left(iz^2\right) dz$, 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 $\lim_{R\to\infty} \int_{C_1(R)} \exp\left(iz^2\right) dz = 0$, 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

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

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

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

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 $\oint_C dz\, f(z) = 0$.

mathematical methods - week 7

Method of steepest descent

Georgia Tech PHYS-6124

Homework HW #7 due Monday, October 7, 2019

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

Exercise 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 6, 2019

Week 7 syllabus

September 30, 2019

Jensen's theorem: saddle point method; Gamma, Airy function estimates;

Mon Arfken and Weber [1] (click here) Chapter 7 sect. 7.3 *Method of steepest descents*

Wed discusses Grigoriev notes: Gamma, Airy functions, Sterling formula

Fri Sect. 7.1 *Saddle-point expansions are asymptotic*

Optional reading

- Arfken and Weber [1] (click here) Chapter 8 has interesting tidbits about the Gamma function. Beta function is also often encountered.
- Branch-cut integrals, see Arfken and Weber example 7.1.6.

Apropos Jensen: the most popular Danish family names are 1. Jensen 303,089 2. Nielsen 296,850 3. Hansen 248,968. This out of population of 5.5 million.

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

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

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

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

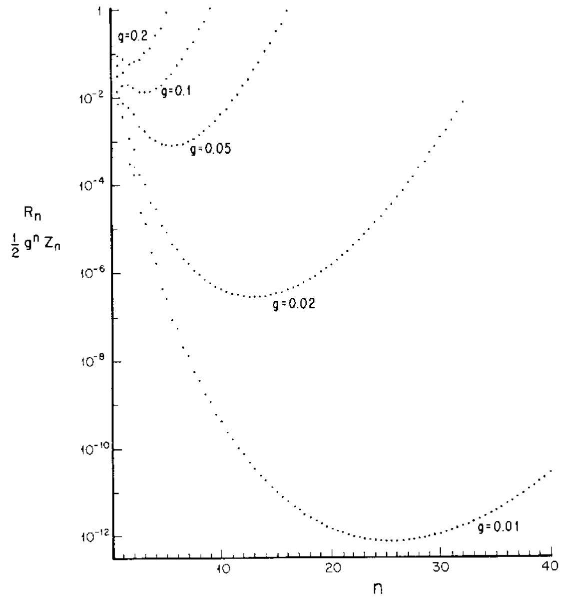


Figure 7.1: Plot of the saddle-point estimate of Z_n vs. the exact result (7.2) for $g = 0.1$, $g = 0.02$, $g = 0.01$.

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 [5]. Consider the residual error after inclusion of the first n perturbative corrections:

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

The inequality follows from the convexity of exponentials, a generalization of the inequality $e^x \geq 1 + x$. 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 $n \leq 5$.

2019-06-01 Predrag I find Córdova, Heidenreich, Popolitov and Shakirov [3] *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

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

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 [4] 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 [8] 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 [6, 7]. 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 [10] 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. [9] 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. [9] 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 and H. J. Weber, *Mathematical Methods for Physicists: A Comprehensive Guide*, 6th ed. (Academic, New York, 2005).

[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] 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).

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

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

[6] 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.

[7] J. A. Gallego, M. G. Perich, R. H. Chowdhury, S. A. Solla, and L. E. Miller, A stable, long-term cortical signature underlying consistent behavior, 2018.

[8] 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).

[9] 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).

[10] J. Hopcroft and R. Kannan, Foundations of Data Science, 2014.

[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 $d\Omega/(\text{unit hyper-sphere surface})$.

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

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

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.

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.* [9].

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.

mathematical methods - week 8

Discrete Fourier transform

Georgia Tech PHYS-6124

Homework HW #8 due Wednesday, October 16, 2019

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

Exercise 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 23, 2019

Week 8 syllabus

Monday, October 7, 2019

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

Mon Applied math version: how to discretize derivatives:
ChaosBook Appendix A24 Deterministic diffusion
Sects. A24.1 to A24.1.1 *Lattice Laplacian*.

Wed 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 transform.*
ChaosBook Example A24.3 *'Configuration-momentum' Fourier space duality.*

Fri Sect. A24.4 Fourier transform as the limit of a discrete Fourier transform.

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 transforms by computing a free propagator on a lattice.

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 version of the lattice Laplacian (**??**) 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:

$$\varphi_k^\dagger \cdot \frac{1}{m^2\mathbf{1} - \Delta} \cdot \varphi_{k'} = \frac{\delta_{kk'}}{m^2 - \frac{2}{(ma)^2}\left(\cos 2\pi k/N - 1\right)} \,, \tag{8.4}$$

where k is a site in the N-dimensional dual lattice, and $a = L/N$ is the lattice spacing.

mathematical methods - week 9

Fourier transform

Georgia Tech PHYS-6124

Homework HW #9 due Monday, October 21, 2019

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

Exercise 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 October 18, 2019

Week 9 syllabus

Wednesday, October 16, 2019

Wed Arfken and Weber [1] (click here) Chapter 14. Fourier Series.
Farazmand notes on Fourier transforms.

Fri 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

Optional reading

- Stone and Goldbart [4] (click here) Appendix B
- Roger Penrose [3] chapter on Fourier transforms is sophisticated, but too pretty to pass up.

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. What is that for discrete Fourier transforms? What is a "convolution theorem" for matrices?
A "Convolution" is a simplified matrix multiplication for translationally invariant matrix operators, see example 9.2.

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):

$$UAU^\dagger = UBU^\dagger UCU^\dagger \quad \rightarrow \quad \tilde{A}_{kk} = \tilde{B}_{kk} \tilde{C}_{kk}\,. \tag{9.6}$$

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

[2] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals* (Dover, New York, 1986).

[3] R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe* (A. A. Knopf, New York, 2005).

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

$$
\begin{aligned}
f(x) &= \frac{1}{N_1}\, e^{-\frac{1}{2} x^\top \cdot \frac{1}{\Delta_1} \cdot x}\,, \quad N_1 = \sqrt{\det\left(2\pi\Delta_1\right)} \\
g(x) &= \frac{1}{N_2}\, e^{-\frac{1}{2} x^\top \cdot \frac{1}{\Delta_2} \cdot x}\,, \quad N_2 = \sqrt{\det\left(2\pi\Delta_2\right)} \\
1 &= \int d^d k\, f(x) = \int d^d k\, g(x)\,.
\end{aligned}
$$

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

$$
[f * g](x) = \frac{(2\pi)^{-d/2}}{\sqrt{\det\left(\Delta_1 + \Delta_2\right)}}\, e^{-\frac{1}{2} x^\top \cdot \frac{1}{\Delta_1 + \Delta_2} \cdot x}\,.
$$

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.

mathematical methods - week 10

Finite groups

Georgia Tech PHYS-6124

Homework HW #10 due Monday, October 28, 2019

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

Exercise 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 23, 2019

Week 10 syllabus

Monday, October 21, 2019

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

Mon Groups, permutations, $D_3 \cong C_{3v} \cong S_3$ symmetries of equilateral triangle, rearrangement theorem, subgroups, cosets.

- Chapter 1 *Basic Mathematical Background: Introduction*
  Dresselhaus *et al.* [3] (click here)
- by Socratica:
  a delightful introduction to group multiplication (or Cayley) tables.
- ChaosBook Chapter 10. Flips, slides and turns
- For deeper insights, read Roger Penrose [8] (click here).

Wed Irreps, unitary reps and 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)

Fri "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* (1.32), 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

Optional reading

- There is no need to learn all these "Greek" words.
- Bedside crocheting.

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.

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.

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 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.

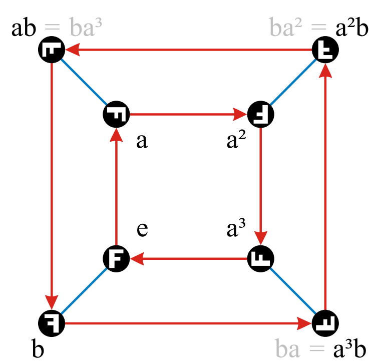


Figure 10.1: A Cayley graph presentation of the dihedral group D_4. The 'root vertex' of the graph, marked e, is here indicated by the letter $\mathbb{F}$, the links are multiplications by two generators: a cyclic rotation by left-multiplication by element a (directed red link), and the flip by b (undirected blue link). The vertices are the 8 possible orientations of the transformed letter $\mathbb{F}$.

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.

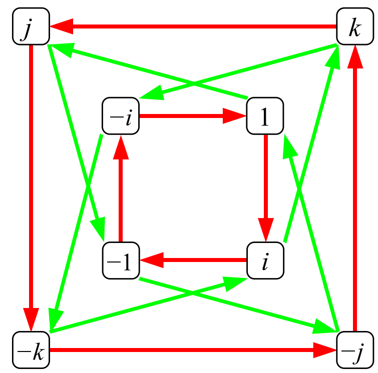
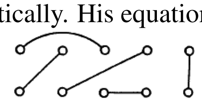


Figure 10.2: A Cayley graph presentation of the quaternion group Q_8. It is also of order 8, but distinct from D_4.

10.1.1 Permutations in 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)

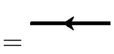


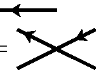
(send index 1 to 2, 2 to 4, contract ingoing (3·4), outgoing (1·3)) is the earliest published diagrammatic notation I know about. While in kindergarten (disclosure: we were too poor to afford kindergarten) I sat out to revolutionize modern group theory [2]. But I suffered a terrible setback; in early 1970's Roger Penrose pre-invented my "birdtracks," or diagrammatic notation, for symmetrization operators [7], Levi-Civita tensors [9], and "strand networks" [6]. 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:

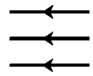
$$\sigma_\alpha^\beta = \sigma^{a_1 a_2 \dots a_q}_{b_1 \dots b_p}{}^{d_p \dots d_1}_{,c_q \dots c_2 c_1} \ . \tag{10.2}$$

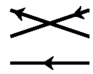
To draw this, Brauer style, it is convenient to turn his drawing on a side. For 2-index tensors, there are two permutations:

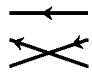
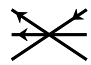
identity: $\mathbf{1}_{ab,}{}^{cd} = \delta_a^d \delta_b^c =$ 

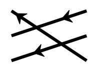
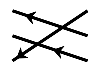
flip: $\sigma_{(12)ab,}{}^{cd} = \delta_a^c \delta_b^d =$  . (10.3)

For 3-index tensors, there are six permutations:

$$\mathbf{1}_{a_1a_2a_3},{}^{b_3b_2b_1} = \delta_{a_1}^{b_1}\delta_{a_2}^{b_2}\delta_{a_3}^{b_3} =$$ 

$$\sigma_{(12)a_1a_2a_3},{}^{b_3b_2b_1} = \delta_{a_1}^{b_2}\delta_{a_2}^{b_1}\delta_{a_3}^{b_3} =$$ 

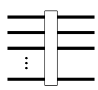
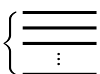
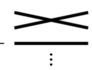
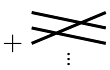
$\sigma_{(23)} =$  , $\sigma_{(13)} =$ 

$\sigma_{(123)} =$  , $\sigma_{(132)} =$  . (10.4)

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.

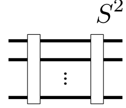
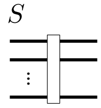
The symmetric sum of all permutations,

$$S_{a_1a_2\ldots a_p},{}^{b_p\ldots b_2b_1} = \frac{1}{p!}\left\{\delta_{a_1}^{b_1}\delta_{a_2}^{b_2}\ldots\delta_{a_p}^{b_p} + \delta_{a_2}^{b_1}\delta_{a_1}^{b_2}\ldots\delta_{a_p}^{b_p} + \ldots\right\}$$

$S =$  $= \frac{1}{p!}\Big\{$  $+$  $+$  $+ \ldots\Big\}$, (10.5)

yields the symmetrization operator S. In birdtrack notation, a white bar drawn across p lines [7] 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

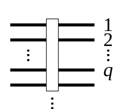
$$S^2 = S$$

 $=$  . (10.6)

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).

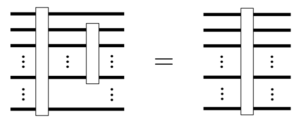
A subset of indices a_1, a_2, … a_q, $q < p$ can be symmetrized by symmetrization matrix $S_{12\ldots q}$

$$(S_{12\ldots q})_{a_1a_2\ldots a_q\ldots a_p},{}^{b_p\ldots b_q\ldots b_2b_1} = \frac{1}{q!}\left\{\delta_{a_1}^{b_1}\delta_{a_2}^{b_2}\ldots\delta_{a_q}^{b_q} + \delta_{a_2}^{b_1}\delta_{a_1}^{b_2}\ldots\delta_{a_q}^{b_q} + \ldots\right\}\delta_{a_q+1}^{b_q+1}\ldots\delta_{a_p}^{b_p}$$

$S_{12\ldots q} =$  . (10.7)

Overall symmetrization also symmetrizes any subset of indices:

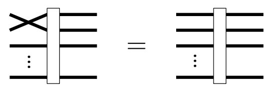
$$SS_{12\ldots q} = S$$



(10.8)

Any permutation has eigenvalue 1 on the symmetric tensor space:

$$\sigma S = S$$



(10.9)

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 Literature

It's a matter of no small pride for a card-carrying dirt physics theorist to claim full and total ignorance of group theory (read sect. A.6 *Gruppenpest* of ref. [5]). The exposition (or the corresponding chapter in Tinkham [10]) that we follow here largely comes from Wigner's classic *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* [11], 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, 2004).

[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. Mainieri and P. Cvitanović, "A brief history of chaos", in *Chaos: Classical and Quantum*, edited by P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay (Niels Bohr Inst., Copenhagen, 2017).

[6] R. Penrose, "Angular momentum: An approach to combinatorical space-time", in *Quantum Theory and Beyond*, edited by T. Bastin (Cambridge Univ. Press, Cambridge, 1971).

[7] R. Penrose, "Applications of negative dimensional tensors", in *Combinatorial mathematics and its applications*, edited by D. J. J.A. Welsh (Academic, New York, 1971), pp. 221–244.

[8] R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe* (A. A. Knopf, New York, 2005).

[9] 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).

[10] M. Tinkham, *Group Theory and Quantum Mechanics* (Dover, New York, 2003).

[11] 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:

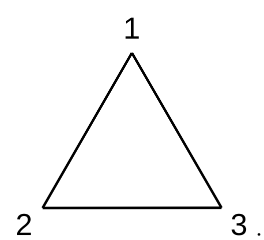
$$\rho(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; \rho(a) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \rho(d) = \begin{pmatrix} z & 0 \\ 0 & z^2 \end{pmatrix},$$

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 $\mathrm{D}_3 \cong \mathrm{C}_{3v} \cong S_3$, 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 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

$$\mathcal{D}(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \mathcal{D}(C) = \begin{pmatrix} z & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z^2 \end{pmatrix}, \mathcal{D}(\sigma^{(1)}) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{10.10}$$

generate a 3-dimensional representation $\mathcal{D}$ 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

$$\mathbf{R}\mathcal{D}(g)\mathbf{R}^{-1} = \begin{pmatrix} A(g) & 0 \\ 0 & \Gamma(g) \end{pmatrix}$$

for all elements g of D_3. (Might help: D_3 has only one (non-equivalent) 2-dim irreducible representation).

(B. Gutkin)

mathematical methods - week 11

Continuous groups

Georgia Tech PHYS-6124

Homework HW #11 due Monday, November 4, 2019

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

Exercise 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.

Week 11 syllabus

Monday, October 28, 2019

Irreps, or "irreducible representations.".

Mon 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;

Wed 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)

Fri Lie groups, sect. 11.3

- 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)

11.1 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.

11.2 Lie groups

In week 1 we introduced projection operators (1.33). 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 [3] 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 (1.34) 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.

11.3 Continuous groups: unitary and orthogonal

Friday's lecture is not taken from any particular book, it's 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 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 whole cyclic group C_N is generated by the powers of the smallest rotation by $2\pi/N$, and in the $N \to \infty$ limit one only needs to understand the algebra of T_ℓ, generators of infinitesimal transformations, $D(\theta) = 1 + i\sum_\ell \theta_\ell T_\ell$.

These thoughts are spread over chapters of my book [1] *Group Theory - Birdtracks, Lie's, and Exceptional Groups* 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.

References

[1] P. Cvitanović, *Group Theory: Birdtracks, Lie's and Exceptional Groups* (Princeton Univ. Press, Princeton NJ, 2004).

[2] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Dover, New York, 1962).

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

[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] 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,

$$c_a = \overline{\chi^{(a)*}\,\chi} = \frac{1}{h} \sum_k^{class} N_k \chi^{(a)}(\mathcal{C}_k^{-1})\, \chi(\mathcal{C}_k)\,. \tag{11.1}$$

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 three objects $[a\ b\ c]$ is a 3×3 identity matrix,

$$D(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Transposing the first and second object yields $[b\ a\ c]$, 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 isotropy group of the velocity field $v(x)$ is the discrete subgroup C_m of SO(2) rotations about an axis (let's say the 'z-axis'),

$$C^{1/m} v(x) = v(C^{1/m} x) = v(x)\,, \qquad (C^{1/m})^m = e\,,$$

the only non-zero components of Fourier-transformed equations of motion are a_{jm} for $j = 1, 2, \cdots$. Argue that the Fourier representation is then the quotient map of the dynamics, $\mathcal{M}/\mathrm{C}_m$. (Hint: this sounds much fancier than what is - think first of how it applies to the Lorenz system and the 3-disk pinball.)

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 Hamermesh [2] or Tinkham [5], and read up on classes and characters.)

mathematical methods - week 12

SO(3) and SU(2)

Georgia Tech PHYS-6124

Homework HW #12 due Monday, November 11, 2019

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

Exercise 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 11, 2019

Week 12 syllabus

Monday, November 4, 2019

Mon The $N \to \infty$ limit of C_N gets you to the continuous Fourier transform as a representation of $\mathrm{SO}(2)$, 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.

- Reading: Chen, Ping and Wang [2] *Group Representation Theory for Physicists*, Sect 5.2 *Definition of a Lie group, with examples* (click here).
- 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.

Wed 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 $\mathrm{SO}(2)$.

Fri OK, I see that formally $\mathrm{SU}(2) \simeq \mathrm{SO}(3)$, but who ordered "spin?"

- For overall clarity and pleasure of reading, I like Schwichtenberg [6] (click here) discussion best. If you read anything for this week's lectures, read Schwichtenberg.
- Read sect. 12.3 $\mathrm{SU}(2) \simeq \mathrm{SO}(3)$

Optional reading

- We had started the discussion of continuous groups last Friday - you might want to have a look at the current version of week 11 notes.
- 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 [5] 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, like Arfken & Weber [1], or Cornwell [3] *Group Theory in Physics: An introduction* that covers group theory might be more in your taste.

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 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

$$(y^1, y^2, \dots, y^n) . \tag{12.2}$$

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*

$$(\mathbf{t}_\alpha \cdot \mathbf{t}_\beta)_a^c = (t_\alpha)_a^b (t_\beta)_b^c , \qquad \mathbf{t}_\alpha \in V \otimes \bar{V} , \tag{12.7}$$

and the *Lie product*

$$(\mathbf{t}_\alpha \cdot \mathbf{t}_\beta)_a^c = (t_\alpha)_a^b (t_\beta)_b^c - (t_\alpha)_c^b (t_\beta)_b^a , \qquad \mathbf{t}_\alpha \in V \otimes \bar{V} \tag{12.8}$$

which defines a *Lie algebra*.

12.2 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

$$dg = d\mu(\theta) = \frac{d\theta}{2\pi}(1 - \cos(\theta)) = \frac{d\theta}{\pi}\sin^2(\theta/2)\,. \tag{12.16}$$

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 exercise 12.4

$$\frac{1}{|G|}\sum_{g\in G} \to \int_G dg\,.$$

The good news is that, as explained in ChaosBook.org Chap. *Relativity for cyclists* (and in *Group Theory - Birdtracks, Lie's, and Exceptional Groups* [4]), one never needs to actually explicitly construct a group manifold parametrizations and the corresponding Haar measure.

12.3 SU(2) and SO(3)

K. Y. Short

An element of SU(2) can be written as

$$U_{vec\hat{n}}(\phi) = e^{i\phi\,\sigma\cdot\hat{\boldsymbol{n}}/2} \tag{12.17}$$

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 matrix of the form

$$\begin{aligned}\sigma\cdot\boldsymbol{x} &= \sigma_x x + \sigma_y y + \sigma_z z \\ &= \begin{pmatrix}0 & x\\ x & 0\end{pmatrix} + \begin{pmatrix}0 & -iy\\ iy & 0\end{pmatrix} + \begin{pmatrix}z & 0\\ 0 & -z\end{pmatrix} \\ &= \begin{pmatrix}z & x-iy\\ x+iy & -z\end{pmatrix}\end{aligned} \tag{12.18}$$

Its determinant

$$\det\begin{pmatrix}z & x-iy\\ x+iy & -z\end{pmatrix} = -(x^2+y^2+z^2) = -\boldsymbol{x}^2 \tag{12.19}$$

gives the length of a vector. Consider a SU(2) transformation (12.17) of this matrix, $U^\dagger(\sigma\cdot\boldsymbol{x})U$. Taking the determinant, we find the same expression as before:

$$\det U(\sigma\cdot\boldsymbol{x})U^\dagger = \det U\,\det(\sigma\cdot\boldsymbol{x})\,\det U^\dagger = \det(\sigma\cdot\boldsymbol{x})\,. \tag{12.20}$$

Just as SO(3), SU(2) preserves the lengths of vectors.

To make the correspondence between $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ more explicit, consider a $\mathrm{SU}(2)$ transformation on a complex two-component *spinor*

$$\psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{12.21}$$

related to $\boldsymbol{x}$ by

$$x = \frac{1}{2}(\beta^2 - \alpha^2), \quad y = -\frac{i}{2}(\alpha^2 + \beta^2), \quad z = \alpha\beta \tag{12.22}$$

Check that a $\mathrm{SU}(2)$ transformation of ψ is equivalent to a $\mathrm{SO}(3)$ transformation on $\boldsymbol{x}$. From this equivalence, one sees that a $\mathrm{SU}(2)$ transformation has three real parameters that correspond to the three rotation angles of $\mathrm{SO}(3)$. If we label the "angles" for the $\mathrm{SU}(2)$ transformation by α, β, and γ, we observe, for a "rotation" about $\hat{x}$

$$U_x(\alpha) = \begin{pmatrix} \cos\alpha/2 & i\sin\alpha/2 \\ i\sin\alpha/2 & \cos\alpha/2 \end{pmatrix}, \tag{12.23}$$

for a "rotation" about $\hat{y}$,

$$U_y(\beta) = \begin{pmatrix} \cos\beta/2 & \sin\beta/2 \\ -\sin\beta/2 & \cos\beta/2 \end{pmatrix}, \tag{12.24}$$

and for "rotation" about $\hat{z}$,

$$U_z(\gamma) = \begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix}. \tag{12.25}$$

Compare these three matrices to the corresponding $\mathrm{SO}(3)$ rotation matrices:

$$R_x(\zeta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\zeta & \sin\zeta \\ 0 & -\sin\zeta & \cos\zeta \end{pmatrix}, \quad R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix}$$
$$R_z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{12.26}$$

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 [6] (click here) who explains clearly that $\mathrm{SU}(2)$ is a simply-connected group, and thus the "mother" or covering group, or the double cover of $\mathrm{SO}(3)$. This means there is a two-to-one map from $\mathrm{SU}(2)$ to $\mathrm{SO}(3)$; an $\mathrm{SU}(2)$ turn by 4π corresponds to an $\mathrm{SO}(3)$ 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.

12.4 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 12.2. Henriette Roux asks
Q This is a math methods course. Why are you not teaching us Bessel functions?
A 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 creating 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 and H. J. Weber, *Mathematical Methods for Physicists: A Comprehensive Guide*, 6th ed. (Academic, New York, 2005).

[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, *Group Theory: Birdtracks, Lie's and Exceptional Groups* (Princeton Univ. Press, Princeton NJ, 2004).

[5] J. Mathews and R. L. Walker, *Mathematical Methods of Physics* (W. A. Benjamin, Reading, MA, 1970).

[6] J. Schwichtenberg, *Physics from Symmetry* (Springer, Berlin, 2015).

Exercises

12.1. **Irreps of SO(2).** Matrix

$$T = \left[\begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right] \tag{12.27}$$

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:

$$e^{i\phi \boldsymbol{n}_2 \cdot \boldsymbol{L}} e^{i\theta \boldsymbol{n}_1 \cdot \boldsymbol{L}} e^{-i\phi \boldsymbol{n}_2 \cdot \boldsymbol{L}} = e^{i\theta \boldsymbol{n}_3 \cdot \boldsymbol{L}} \tag{12.28}$$

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

$$\chi = 2\cos(\theta) + 1 \,. \tag{12.29}$$

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:

$$\langle \chi(j) | \chi(j') \rangle = \int_G dg\, \chi^{(j)}(g^{-1})\, \chi^{(j')}(g) = \delta_{jj'} \,. \tag{12.30}$$

mathematical methods - week 13

Probability

Georgia Tech PHYS-6124

Homework HW #13 due Monday, November 18, 2019

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

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 15, 2019

Week 13 syllabus

November 11, 2019

Mon A summary of key concepts

- ChaosBook appendix A20.1 Moments, cumulants

Wed Why Gaussians again?

- ChaosBook 33.2 Brownian diffusion
- ChaosBook 33.3 Noisy trajectories

Fri 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

13.1 Literature

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, *How well can one resolve the state space of a chaotic map?*, Phys. Rev. Lett. 104, 014101 (2010); arXiv:0902.4269
- P. Cvitanović and D. Lippolis, *Knowing when to stop: How noise frees us from determinism*, in M. Robnik and V.G. Romanovski, eds., *Let's Face Chaos through Nonlinear Dynamics* (Am. Inst. of Phys., 2012); arXiv:1206.5506
- J. M. Heninger, D. Lippolis and P. Cvitanović, *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.

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

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 $0 \leq \tau \leq t$ 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?

mathematical methods - week 14

Math for experimentalists

Georgia Tech PHYS-6124

Homework HW #14 due Monday, November 25, 2019

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

Exercise 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

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

edited November 22, 2019

Week 14 syllabus November 18, 2019

Lecturer: Ignacio Taboada

Mon Uncertainty, probability, probability density functions, error matrix, etc.

- Ignacio's lecture notes.
- Predrag's summary of key concepts for a physicist: ChaosBook appendix A20.1 Moments, cumulants.

Wed Distributions: binomial; Poisson; normal; uniform. Moments; quantiles.

- Ignacio's lecture notes.
- For binomial theorem, Poisson and Gaussian distributions, see Arfken and Weber [1] (click here) *Mathematical Methods for Physicists: A Comprehensive Guide*, Chapter 19.

Fri Monte Carlo (why you need the uniform ditribution); central limit theorem (why you need normal dist). Multi-dimensional PDFs; correlation error propagation.

- Ignacio's lecture notes.
- Predrag's Noise is your friend and ChaosBook 33.3 Noisy trajectories derive the closely related covariance matrix evolution formula

14.1 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 [3], *Analysis methods for results in gamma-ray astronomy*, and Feldman and Cousins [2] *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 $\rightarrow$ 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 $\rightarrow$ 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

References

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

[2] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals", Phys. Rev. D **57**, 3873–3889 (1998).

[3] T.-P. Li and Y.-Q. Ma, "Analysis methods for results in gamma-ray astronomy", Astrophys. J. **272**, 317–324 (1983).

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.

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.

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.

mathematical methods - week 15

(Non)linear dimensionality reduction

Georgia Tech PHYS-6124

Homework HW #15 due Monday, December 2, 2019

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

Exercise 15.1 *Unbiased sample variance* 5 points

Bonus points
Exercise 15.2 *Standard error of the mean* 5 points
Exercise 15.3 *Bayesian statistics*, by Sara A. Solla 10 points

Total of 10 points = 100 % score. Extra points accumulate, can help you still if you had missed a few problems.

edited December 2, 2019

Week 15 syllabus

Monday, November 25, 2019

Linear and nonlinear dimensionality reduction:
applications to neural data
Lecturer: Sara A. Solla

Mon Neural recordings; Principal Components Analysis (PCA); Singular Value Decomposition (SVD); ISOMAP nonlinear dimensionality reduction; Multidimensional scaling

- Sara's lecture notes.
- Predrag's summary of key concepts for a physicist: ChaosBook appendix A20.1 Moments, cumulants.

15.1 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 2019 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 [1], *Bayes and Frequentism: a particle physicist's perspective* (click here)

References

[1] L. Lyons, "Bayes and Frequentism: a particle physicist's perspective", Contemporary Physics **54**, 1–16 (2013).

Exercises

15.1. **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,

$$\mathbb{E}[\hat{\mu}] = \mu\,, \qquad \mathbb{E}[\hat{\sigma}^2] = \sigma^2\,. \tag{15.1}$$

(a) Verify that the empirical mean

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} a_i \tag{15.2}$$

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

$$\left(\sum_{i=1}^{N} a_i\right)^2 = \sum_{i=1}^{N} a_i^2 + \sum_{i\neq j}^{N} a_i a_j\,. \tag{15.3}$$

(c) Show that the empirical estimate of form

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(a_i - \hat{\mu})^2\,, \tag{15.4}$$

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

15.2. **Standard error of the mean.**
Now, estimate the empirical mean (15.2) of observable a by $j = 1, 2, \cdots, N$ attempts to estimate the mean $\hat{\mu}_j$, each based on M data samples

$$\hat{\mu}_j = \frac{1}{M}\sum_{i=1}^{M} a_i\,. \tag{15.5}$$

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

$$\mathrm{Var}[\hat{\mu}] = \mathbb{E}[(\hat{\mu} - \mu)^2] = \mathbb{E}[\hat{\mu}^2] - \mu^2$$

as a function of variance expectation (15.1) and N, the number of $\hat{\mu}_j$ estimates? Hint; one way to do this is to repeat the calculations of exercise 15.1, this time for $\hat{\mu}_j$ rather than a_i.

(c) The quantity $\sqrt{\mathrm{Var}[\hat{\mu}]} = \sigma/\sqrt{N}$ 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

15.3. **Bayes.** *Bayesian statistics*.

mathematical methods - week 16

Calculus of variations

Georgia Tech PHYS-6124

Homework HW #16 due whenever in 2020

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort

Optional exercises

Problem set #0 *Farmer and the pig* 0 points

Problem set #1 *Fermat's principle* 0 points

Problem set #2 *Lagrange multipliers* 0 points

No points for any of these, but solutions available upon request.

Perhaps, by now, you have gotten drift of what this course is about. We always start with something intuitively simple and obvious; all topics are picked from current, on-going research across physics and engineering. Then we let mathematics take over, and bang! math takes you someplace where your intuition fails you, and the surprising path that Nature had picked instead is stunningly beautiful. Like quantum mechanics, or electron spin.

There is so much more essential mathematical physics we never got to... For example, I would have loved to discuss the calculus of variations, and whether the God does exist?

In either case, I wish you a great, stimulating, rewarding 2020!

Mon Calculus of variations.

- ChaosBook sect. 34.3 *Least action method*
- P. Goldbart notes
- M. Stone & P. Goldbart Chapter 1

Wed (No class meeting) Lagrange-Euler equations; Lagrange multipliers.

- M. Stone & P. Goldbart Chapter 1
- M. Stone & P. Goldbart Section 1.5
- Example 16.1 *Gaussian minimizes information*
- Read only if things nonlinear interest you: *A variational principle for periodic orbit searches*; read *Turbulent fields and their recurrences* first.

Fri (No class meeting) From Lagrange to Hamilton; Constrained Hamiltonian systems; Dirac constraints.

- Cristel Chandre *Lagrange to Hamilton* notes
- Cristel Chandre *Dirac constraints* notes

16.1 Calculus of variations

Think globally, act locally.
— Patrick Geddes

It all started by pondering elementary geometrical problems, such as ChaosBook 34.3 Least action method reflection and refraction of light rays, but by the end of 18th century d'Alembert (1742), Maupertuis (1744), and an Italian immigrant to France, Joseph-Louis de La Grange (1788) knew that classical mechanics in phase-space of any dimensions can be recast -in a demonstration of divine elegance- as a variational condition on a single scalar (!) function. In physics that is known as “the principle of least action,” in engineering as “constrained optimization,” and the function is known as respectively the “Lagrangian” or the “cost” function.

In the regulation physics indoctrination classes, Lagrangians are always presented as an partial derivative gymnastics orgy relatives of the cuddly old Hamiltonians. But that totally misses the point. Hamiltonian formulation obtains solutions of natural laws by integration in time for given, *locally* specified initial conditions. Lagrangian formulation seeks *global* solutions, to be obtained without any time or space integrations.

Maupertuis saw this as a proof of the existence of God, Lagrange as merely a bag of useful mathematical tricks to solve mechanical problems with constraints. Neither could dream that in 20th century the masters of Lagrangians would be not physicists, but robotics engineers, and that Lagrangians would be central to the formulation of special and general relativity, quantum mechanics and quantum field theory (Feynman path integrals), a succinct statement of the myriad spatiotemporal and internal symmetries of modern particle physics.

In 21st century this goes by name of “Machine Learning" or even “Deep Learning;" I have invited our friend from Minsk to tell us all about it. “Machine learning" is typically a gradient-descent method or a neural network for searching for desired answers, and that would be vastly improved if we knew how to impose natural laws as constraints on the spaces one searches in. That turns out to be surprisingly hard. In the class I tried to explain how one imposes constraints in the Lagrangian constrained-optimization formalism.

We illustrate the principle by answering the question that might have bugged you in the recent weeks - what’s so special about Gaussians?

Example 16.1. ***Gaussian minimizes information.*** *Shannon information entropy is given by*

$$S[\rho] = -\langle \ln \rho \rangle = -\int_{\mathcal{M}} dx\, \rho(x)\, \ln \rho(x)\,, \qquad (16.1)$$

where ρ is a probability density. Shannon thought of $-\ln \rho$ as ‘information’ in the sense that if -for example- $\rho(x) = 2^{-6}$, it takes $-\ln \rho = 6$ bits of ‘information’ to specify the probability density ρ at the point x. Information entropy (16.1) is the expectation value of (or average) information.

A probability density $\rho \geq 0$ is an arbitrary function, of which we only require that it is normalized as a probability,

$$\int_{\mathcal{M}} dx\, \rho(x) = 1\,, \qquad (16.2)$$

has a mean value,

$$\int_{\mathcal{M}} dx\, x\, \rho(x) = \mu\,, \qquad (16.3)$$

and has a variance

$$\int_{\mathcal{M}} dx\, x^2 \rho(x) = \mu^2 + \sigma^2\,. \qquad (16.4)$$

As ρ can be arbitrarily wild, it might take much “information” to describe it. Is there a function $\rho(x)$ that contains the least *information, i.e., that minimizes the information entropy (16.1)?*

To find it, we minimize (16.1) subject to constraints (16.2)-(16.4), implemented by *adding Lagrange multipliers* λ_j

$$\begin{aligned} C[\rho] &= \int_{\mathcal{M}} dx\, \rho(x)\, \ln \rho(x) \\ &+\lambda_0 \left(\int_{\mathcal{M}} dx\, \rho(x) - 1 \right) + \lambda_1 \left(\int_{\mathcal{M}} dx\, x\, \rho(x) - \mu \right) \\ &+\lambda_2 \left(\int_{\mathcal{M}} dx\, (x-\mu)^2 \rho(x) - \sigma^2 \right), \end{aligned} \tag{16.5}$$

and looking for the extremum $\delta C = 0$*,*

$$\frac{\delta C[\rho]}{\delta \rho(x)} = (\ln \rho(x) + 1) + \lambda_0 + \lambda_1 x + \lambda_2 x^2 = 0\,, \tag{16.6}$$

so

$$\rho(x) = e^{-(1+\lambda_0+\lambda_1 x+\lambda_2 x^2)}\,. \tag{16.7}$$

The Lagrange multipliers λ_j *can be expressed in terms of distribution parameters* μ *and* σ *by substituting this* $\rho(x)$ *into the constraint equations (16.2)-(16.4), and we find that the probability density that minimizes information entropy is the Gaussian*

$$\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,. \tag{16.8}$$

In what sense is that the distribution with the 'least information'? As we saw in the derivation of the ChaosBook appendix A20.1 cumulant expansion eq. (A20.7), for a Gaussian distribution all cumulants but the mean μ *and the variance* σ^2 *vanish, it is a distribution specified by only two 'informations', the location of its peak and its width.*

Sara A. Solla

And that's all, folks.

mathematical methods - week 17

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 : 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.