Math 218 — Extra problems 2

Alex Cowan
alex.cowan@uwaterloo.ca

The matrix exponential

1.

1.1) Let A be an $n \times n$ matrix defined over the complex numbers. Show that $|A^k_{i,j}| \leq C^k$ for some constant C depending only on A (this means dependence on various features of A, such as n and its entries $A_{i,j}$, are okay too).

1.2) Show that, for any $n \times n$ matrix A defined over the complex numbers, the series defining $\exp(tA)$ converges for all $t \in \mathbb{C}$.

2. Let A be an $n \times n$ matrix defined over $\mathbb{C}$.

2.1) Show that, for any $t_0 \in \mathbb{R}$ fixed, both $\exp(tA)\exp(t_0 A)$ and $\exp((t+t_0)A)$ solve the initial value problem $M'(t) = AM(t)$ with $M(0) = \exp(t_0 A)$ (here M is an $n \times n$ matrix of functions of t).

2.2) Justify the fact that $\exp(t_1 A)\exp(t_2 A) = \exp((t_1 + t_2)A)$ for all $t_1, t_2 \in \mathbb{R}$.

2.3) Show that $\exp(tA)^{-1} = \exp(-tA)$.

2.4) Show, using the series definition of exp, that "to first order" $\exp(t_1 A)\exp(t_2 A) = \exp((t_1 + t_2)A)$ for all $t_1, t_2 \in \mathbb{R}$. "To first order" means considering only the terms whose t_1-degree plus t_2-degree is 1 or less.

2.5) Show, using the series definition of exp, that $\exp(t_1 A)\exp(t_2 A) = \exp((t_1 + t_2)A)$ to second order in t_1 and t_2.

2.6) In general $e^X e^Y \neq e^{X+Y}$. Define Z to be such that $e^X e^Y = e^Z$. Compute Z to third order in X and Y.

3. Give code which numerically estimates $\exp(tA)$, and then uses this to numerically estimate the solution $\vec{x}$ to the initial value problem $\vec{x}'(t) = A\vec{x}(t)$ and $\vec{x}(0) = \vec{c}$.

4. Let $\vec{x}$ be a solution to the initial value problem $\vec{x}'(t) = A\vec{x}(t)$ and $\vec{x}(0) = \vec{c}$.

4.1) Show that $\vec{x}$ satisfies the equation

$$\vec{x}(t) = \vec{c} + \int_0^t A\vec{x}(s)\, ds.$$

Define $\vec{x}_0(t) \coloneqq \vec{c}$ and

$$\vec{x}_{k+1}(t) \coloneqq \vec{c} + \int_0^t A\vec{x}_k(s)\, ds$$

for $k = 0, 1, \ldots$.

4.2) Show that $\vec{x}_1 = (I + tA)\vec{c}$.

4.3) Show that $\vec{x}_2 = (I + tA + \frac{t^2}{2}A^2)\vec{c}$.

4.4) Show that $\vec{x}_k = (I + tA + \frac{t^2}{2}A^2 + \cdots + \frac{t^k}{k!}A^k)\vec{c}$.

4.5) Show that $\vec{x} = \lim_{k\to\infty} \vec{x}_k = \exp(tA)\vec{c}$.

5. Prove the following statements.

5.1) $\det(I + \varepsilon T) = 1 + \varepsilon \operatorname{tr}(T) + O(\varepsilon^2)$.

5.2) $\det(A + \varepsilon T) = \det(A) + \varepsilon \operatorname{tr}(A^{-1}T)\det(A) + O(\varepsilon^2)$.

5.3) $\frac{d}{dt}\det(e^{tB}) = \operatorname{tr}(B)\det(e^{tB})$. Hint: take $T = \frac{d}{dt}e^{tB}$ and $\varepsilon = \Delta t$.

5.4) $\det(e^X) = e^{\operatorname{tr}(X)}$.

Lie groups

6. This problem introduces the Lie groups of orthogonal and unitary matrices, and illustrates some techniques for computing their Lie algebras.

6.1) Show that $\exp(M^T) = \exp(M)^T$, i.e. the exponential of the transpose of a matrix is the transpose of the exponential of that matrix.

The matrix Q is said to be *orthogonal* if $QQ^T = Q^TQ = I$.

6.2) Fix an orthonormal basis of $\mathbb{R}^n$ (e.g. the standard basis). Show that, for $v \in \mathbb{R}^n$, thought of as a column vector, one has $\|v\|^2 = v^Tv$.

6.3) Let Q be an orthogonal matrix. Show that $\|Qv\|^2 = \|v\|^2$. (Q is an "isometry".)

6.4) Show that if $M = -M^T$ (i.e. M is *skew-symmetric*) then $\exp(M)$ is orthogonal.

6.5) Show that if M is skew-symmetric then $\exp(M)$ has determinant 1.

6.6) Give an example of an orthogonal matrix with determinant -1.

Let $A^\dagger := \overline{A^T}$ denote the (complex) conjugate transpose of A.[1] The matrix A is *Hermitian* if $A = A^\dagger$. The matrix U is *unitary* if $UU^\dagger = U^\dagger U = I$.

6.7) Show that if A is Hermitian, then $\exp(iA)$ is unitary.

6.8) Calculate $\left(I + \varepsilon\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right)\left(I + \varepsilon\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right)^T$.

6.9) If you want the expression you just got to be equal to I to "first order" in ε (i.e. imagining that $\varepsilon^2 = 0$ while $\varepsilon \neq 0$), what conditions must you impose on a, b, c, d?

6.10) For arbitrary dimension, what condition must you impose on the matrix X to ensure that $(I + \varepsilon X)(I + \varepsilon X)^T = I + O(\varepsilon^2)$? Here $O(\varepsilon^2)$ denotes a matrix whose entries are all bounded by a constant times ε^2 as $\varepsilon \to 0$.[2]

6.11) Consider side by side your work in parts 4, 9, and 10 of this problem. Do you see how these are three sides of the same coin?

6.12) What questions would I have asked to prompt you to investigate unitary matrices, in a way parallel to what you just did for orthogonal matrices?

7. Let U(1) denote the set of 1×1 unitary matrices (which we will think of as interchangible with certain complex numbers). This set is in fact a *group*: any two elements can be multiplied to give another element, and every element is invertible.

7.1) Show that $\cos\theta + i\sin\theta \in \mathrm{U}(1)$ for all $\theta \in \mathbb{R}$.

7.2) Show that every $z \in \mathrm{U}(1)$ can be written as $\cos\theta + i\sin\theta \in \mathrm{U}(1)$ for some $\theta \in \mathbb{R}$. In conjunction with the

[1] $A^\dagger$ is pronounced "A dagger". The notation A^* for the conjugate transpose is also common, but beware that physicists often write A^* to mean the conjugate and not the conjugate transpose.

[2] This is *big O notation*.

previous question, you now have a complete description of $\mathrm{U}(1)$.

7.3) Note that $I = 1 \in \mathrm{U}(1)$. What condition must you impose on $h \in \mathrm{Mat}_{1\times 1}(\mathbb{C})$ to ensure that $1 + \varepsilon h \in \mathrm{U}(1)$ to first order in ε?

The condition you found in the previous problem is linear. Thus, the "Lie algebra" $\mathfrak{u}(1)$ of matrices which satisfy this condition is a vector space (of matrices, i.e. the "vectors" in this abstract vector space are matrices). Consider $\mathfrak{u}(1)$ as a vector space over $\mathbb{R}$, not over $\mathbb{C}$. This means that you are allowed to take $\mathbb{R}$-linear combinations of matrices in the vector space $\mathfrak{u}(1)$, but not $\mathbb{C}$-linear combinations.

One can pick a basis for this $\mathbb{R}$-vector space $\mathfrak{u}(1)$, i.e. a finite list of matrices such that every matrix in the Lie algebra $\mathfrak{u}(1)$ can be written uniquely as an $\mathbb{R}$-linear combination of matrices in your list of basis matrices. In this case, you should find that the Lie algebra $\mathfrak{u}(1)$ is 1-dimensional over $\mathbb{R}$, i.e. it is a 1-dimensional subspace of the space $\mathrm{Mat}_{1\times 1}(\mathbb{C}) \cong \{a + bi : a, b \in \mathbb{R}\}$. Note that $\mathrm{Mat}_{1\times 1}(\mathbb{C})$ is 2-dimensional as an $\mathbb{R}$-vector space: the two element set $\{1, i\}$ is a basis since $a + bi = a \cdot 1 + b \cdot i$, and there is no real number r such that $r \cdot 1 = i$.

Let $g \in \mathfrak{u}(1)$ span the 1-dimensional $\mathbb{R}$-vector space $\mathfrak{u}(1)$, i.e. take a basis $\{g\}$ of one element. You can choose g to be any nonzero element of $\mathfrak{u}(1)$.

7.4) Write down the solution to the differential equation $x' = gx$ with initial condition $x(0) = 1$, using the matrix exponential (which, you will note, for 1×1 matrices is the same as the usual exponential).

You showed in a previous problem that if A is Hermitian then $\exp(iA)$ is unitary. Note that tg/i is Hermitian for all $t \in \mathbb{R}$. Therefore the solution $x(t)$ is in $\mathrm{U}(1)$ for all $t \in \mathbb{R}$. Recall from part 2 earlier that every $z \in \mathrm{U}(1)$ can be written in the form $\cos\theta + i\sin\theta$ for $\theta \in \mathbb{R}$. Therefore, $x(t) = \cos\theta(t) + i\sin\theta(t)$.

7.5) What is the value of $\theta(0)$?

7.6) Evaluate $\frac{d}{dt}\cos\theta(t) + i\sin\theta(t)\big|_{t=0}$ using the chain rule.

7.7) Give an expression for $\frac{d}{dt}\cos\theta(t) + i\sin\theta(t)\big|_{t=0}$ by inspecting the differential equation $x' = gx$.

7.8) Oberve that, over the course of this problem, you have shown how Euler's formula is a result of exponentiating (i.e. $it \mapsto \exp(it)$) the "local symmetry" of the Lie algebra $\mathfrak{u}(1)$ inherent in the differential equation $x' = ix$ to the "global symmetry" of the Lie group $\mathrm{U}(1)$.

8. Let $\mathrm{SO}(2)$ denote the set of 2×2 orthogonal matrices with determinant 1. This set is in fact a *group*: any two elements can be multiplied to give another element, and every element is invertible.

8.1) Show that $\left(\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right) \in \mathrm{SO}(2)$ for all $\theta \in \mathbb{R}$.

8.2) Show that every matrix in $\mathrm{SO}(2)$ can be written as $\left(\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ for some $\theta \in \mathbb{R}$. In conjunction with the previous question, you now have a complete description of $\mathrm{SO}(2)$.

8.3) Note that $I \in \mathrm{SO}(2)$. What conditions must you impose on $h \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ to ensure that $I + \varepsilon h \in \mathrm{SO}(2)$ to first order in ε?

8.4) Note that the conditions you found in the previous problem are linear. Thus, the "Lie algebra" $\mathfrak{so}(2)$ of matrices which satisfy these conditions is a vector space. Give a basis for this vector space, i.e. a finite list of matrices such that every matrix which satisfies your conditions can be written uniquely as a linear combination of matrices in your list.

8.5) You should find that your Lie algebra from the previous problem is 1-dimensional. Let g be a matrix which spans this space (here this is just picking any nonzero elemeent). Solve the initial value problem $x' = gx$ and $x(0) = (1, 0)^T$. Give your solution in terms of real numbers only.

8.6) Recall that every matrix in $\mathrm{SO}(2)$ can be written as $\left(\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right)$. Evaluate $\frac{d}{d\theta}\left(\begin{smallmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ at $\theta = 0$.

Define $\varphi : \mathbb{C} \to \mathrm{Mat}_2(\mathbb{R})$ by $\varphi(a + bi) = \left(\begin{smallmatrix}a & -b \\ b & a\end{smallmatrix}\right)$, where $a, b \in \mathbb{R}$.

8.7) Show that $\varphi(zw) = \varphi(z)\varphi(w)$ for all $z, w \in \mathbb{C}$. The function φ is a *homomorphism*: it preserves structure, in the sense that it doesn't matter if you multiply before or after applying φ.

8.8) Show that $|z| = \det\varphi(z)$ for all $z \in \mathbb{C}$.

8.9) What is the preimage of SO(2) under φ? I.e. what is the set of complex numbers $\{z \in \mathbb{C} : \varphi(z) \in \mathrm{SO}(2)\}$?

You can now note that φ is injective (i.e. different inputs yield different outputs), so it is an "isomorphism" between SO(2) and its preimage U(1). This means that they are basically the same thing and interchangible in all situations.

9. In this problem we will give an example of how the perspective on Euler's formula outlined in problem 6 can be generalized to other Lie groups. The following example concerns a case that is particularly important in quantum mechanics.

Let SU(2) denote the group of 2×2 unitary matrices with determinant 1. Define

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

These are called the *Pauli matrices*. Note that they are Hermitian.

9.1) Show that the Pauli matrices satisfy the following relations.

$$\begin{aligned} \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 &= I \\ \sigma_1\sigma_2 = -\sigma_2\sigma_1 &= i\sigma_3 \\ \sigma_2\sigma_3 = -\sigma_3\sigma_2 &= i\sigma_1 \\ \sigma_3\sigma_1 = -\sigma_1\sigma_3 &= i\sigma_2 \end{aligned}$$

Fix $v \in \mathbb{R}^3$ with $\|v\| = 1$. Define $v \cdot \vec{\sigma} := v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3$.

9.2) Show that $(v \cdot \vec{\sigma})^{2k} = I$ for all $k \in \mathbb{Z}_{\geq 0}$, and that $(v \cdot \vec{\sigma})^{2k+1} = v \cdot \vec{\sigma}$.

9.3) Show that $\{i\sigma_1, i\sigma_2, i\sigma_3\}$ forms an $\mathbb{R}$-basis for the Lie algebra $\mathfrak{su}(2)$. (Reference previous problems to see how to do this.)

9.4) Using the series definition of exp, show that, for any $\theta \in \mathbb{C}$,

$$\exp(i\theta(v \cdot \vec{\sigma})) = I\cos\theta + i(v \cdot \vec{\sigma})\sin\theta.$$

9.5) Explain why, if $\theta \in \mathbb{R}$, then $I\cos\theta + i(v \cdot \vec{\sigma})\sin\theta \in \mathrm{SU}(2)$.

10. The *Schrödinger equation* is

$$i\hbar\frac{\partial}{\partial t}\psi(t) = H\psi(t),$$

where the *Hamiltonian* H is a Hermitian matrix, $\psi(t) \in \mathbb{C}^n$ is a vector of functions from $\mathbb{R}$ to $\mathbb{C}$, and $\hbar \approx 1.05457 \cdot 10^{-34}\,\frac{\mathrm{kg}\cdot\mathrm{m}^2}{\mathrm{s}}$ is Planck's constant divided by 2π. The Schrödinger equation is the fundamental equation of motion in quantum mechanics, playing the same role as $F = ma$ does in classical mechanics. Basically all of non-relativistic quantum mechanics is solving the Schrödinger equation.

10.1) Write down the solution to the Schrödinger equation with initial condition $\psi = \psi(0)$ at $t = 0$ in terms of the matrix exponential.

Every Hermitian matrix is diagonalizable and has real eigenvalues. Let $\psi_1, \ldots, \psi_n$ be an eigenbasis with corresponding eigenvalues $E_1, \ldots, E_n$. (These eigenvalues are the energies of the corresponding eigenvectors a.k.a. "eigenstates".)

10.2) Write down the general solution to the Schrödinger equation as a linear combination of the eigenstates ψ_k.

Let v_0, v_1, v_2, v_3 be real numbers, thought of as having units of energy, and define $v := (v_1, v_2, v_3) \in \mathbb{R}^3$ (note that v_0 is omitted here). For the remainder of this problem, set $H := v_0 I + v \cdot \vec{\sigma}$, where $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli matrices defined in a previous problem.

10.3) Show that $H = v_0 I + v \cdot \vec{\sigma}$ is the general form of a 2×2 Hermitian matrix.

10.4) Using the result of a previous problem, show that

$$e^{-iHt/\hbar} = e^{-iv_0 t/\hbar}\left(I\cos\left(\tfrac{\|v\|}{\hbar}t\right) - i\left(\tfrac{v}{\|v\|}\cdot\vec{\sigma}\right)\sin\left(\tfrac{\|v\|}{\hbar}t\right)\right).$$

10.5) What Lie group is $e^{-iHt/\hbar}$ an element of?

10.6) Show that the eigenvalues of H are $v_0 \pm \|v\|$. This can be shown with minimal calculation by reflecting on the previous parts of this problem.

10.7) Nuclear magnetic resonance imaging (nMRI), a technique widely used in medicine, applies a magnetic field $\vec{B} = B\vec{n} \in \mathbb{R}^3$ to particles with magnetic moment $\vec{\mu} = \mu\vec{\sigma}$, where $\|\vec{n}\| = 1$ and $B, \mu \in \mathbb{R}$ with units of mass per (time^2 $\cdot$ current) and length^2 $\cdot$ current respectively. The magnetic field $\vec{B}$ separates particles into two eigenstates, a high-energy one and a low-energy one. The Hamiltonian of the resulting quantum-mechanical system is $H = -\vec{\mu}\cdot\vec{B}$. What is the energy difference between the two eigenstates?

Conserved quantities

11. For a matrix B and a vector $x(t)$, define $\langle B\rangle := x(t)^\dagger B x(t)$, where $\dagger$ denotes the conjugate transpose. Suppose x satisfies the differential equation $x' = Ax$.

11.1) Show that $\frac{d}{dt}\langle B\rangle = x(t)^\dagger(A^\dagger B + BA)x(t)$.

11.2) Suppose iA is Hermitian. Show that, if A and B commute, then the quantity $\langle B\rangle$ is constant as t varies.

In the Schrödinger equation $i\hbar\frac{\partial}{\partial t}\psi = H\psi$, the matrix H is always Hermitian. Thus, whenever $HB - BH = 0$, the expectation $\langle B\rangle$ for the quantum system is a conserved quantity as time elapses. A classical analogue of this fact exists, involving "Poisson brackets"; these are explored in a later problem.

11.3) Let $\sigma_1, \sigma_2, \sigma_3$ denote the Pauli matrices, defined previously. Take $A = i\sigma_j$. What is $\frac{d}{dt}\langle i\sigma_k\rangle$? Express your answer in terms of j and k, considering all nine possible combinations of values.

12. Let $\Omega = \left(\begin{smallmatrix} 0 & I_n \\ -I_n & 0\end{smallmatrix}\right)$ be the standard "symplectic form" on $\mathbb{R}^{2n}$. Let $\mathrm{Sp}(2n)$ be the Lie group of $2n \times 2n$ real matrices M satisfying $M^T\Omega M = \Omega$.

12.1) What condition must be imposed on matrices for them to be in the Lie algebra $\mathfrak{sp}(2n)$ of $\mathrm{Sp}(2n)$?

Let $A \in \mathfrak{sp}(4)$. Suppose $x(t) = (x_1, x_2, x_3, x_4)^T$ and $y(t) = (y_1, y_2, y_3, y_4)^T$ satisfy $x' = Ax$ and $y' = Ay$.

12.2) Show that the value $x_1y_3 + x_2y_4 - x_3y_1 - x_4y_2$ is independent of t.

13. *Hamiltonian mechanics* are a formulation of classical mechanics. For $i = 1, \ldots, n$, let q_i denote the positions of things in your system, and p_i the associated momenta. In Hamiltonian mechanics, every physical system is determined by a function $\mathcal{H} = \mathcal{H}(q, p)$, where $q = (q_1, \ldots, q_n)$ and $p = (p_1, \ldots, p_n)$. (The Hamiltonian $\mathcal{H}$ can also depend explicitly on time sometimes, but that's weird so let's not do that.) The equations of motion in Hamiltonian mechanics are

$$\dot{q}_i = \frac{\partial\mathcal{H}}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial\mathcal{H}}{\partial q_i}$$

for each $i = 1, \ldots, n$.

13.1) The total energy of the system is the real number $\mathcal{H}(q, p)$ evaluated at the point (q, p) in phase space. Using the chain rule, show that, if Hamilton's equations of motion are satisfied, then energy is conserved.

13.2) Note that the equations of motion in Hamiltonian mechanics preserve a symplectic form. Use the previous problem to prove *Liouville's theorem*, that phase space is incompressible (i.e. that regions of phase space do not change volume when evolved in time according to Hamiltonian dynamics).

Given two functions $F(q,p)$ and $G(q,p)$ on phase space, define the *Poisson bracket*

$$\{F, G\} := \sum_i \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}.$$

13.3) Show that $\dot{F} = \{F, H\}$. Explain the parallel with the quantum-mechanical version of this statement explored in problem 11.