Extra credit assignment: harder problems.

You can hand in any number of these problems. Each one adds 0.5% to your term mark.

1. **Vandermonde Determinant.** The goal of this problem is to compute the determinant of the matrix A defined by $a_{ij} = x_i^j$, for $i, j = 0, \dots, n$, where $x_0, \dots, x_n$ are variables (so it is an $(n+1) \times (n+1)$ -matrix).

   We are going to prove that $\det(A) = \prod_{0 \le i < j \le n} (x_i - x_j)$.

   (a) Compute the 2×2 and 3×3 Vandermonde determinants.

   Hint. We actually did this in lecture.

   (b) Use column operations to make the first row have the form $1\,0\,\dots 0$. Record the resulting matrix.

   (c) Now use induction to prove the result.

   Remark. Many other proofs exist. One of my favourite ones uses the properties of polynomials: fix the values of all the variables except for x_0, and think of x_0 as a variable. Now if you plug any of the fixed values x_i for x_0, the determinant clearly becomes 0. Then (by the properties of polynomials that you will study in Math 323) the expression $x_0 - x_i$ has to divide the determinant (viewed as a polynomial in the x_i). Since you could swap rows, this applies to every expression $x_i - x_j$. Now just comparing the degrees and leading coefficients of these polynomials, we obtain the result.

2. **Operator calculus.** Let $A : V \to V$ be a linear operator on a vector space V over a field F (we are not assuming that V is finite-dimensional in this problem). We define the powers $A^n : V \to V$ as the composition of A with itself n times: $A^n(v) = A(A(..(Av))..)$. Then given a polynomial $p(x) = a_n x^n + \dots + a_0$, where $a_i \in F$, we can define $p(A) : V \to V$ to be the linear operator $p(A) = a_n A^n + \dots + a_0 \mathrm{Id}$, where $\mathrm{Id} : V \to V$ is the identity. Suppose that v is an eigenvector for A with eigenvalue λ, i.e, $Av = \lambda v$, and $v \ne 0$.

   (a) Prove that v is an eigenvector for A^n with eigenvalue λ^n.

   (b) Prove that v is an eigenvalue for $p(A)$ with eigenvalue $p(\lambda)$.

(c) Now let V be finite-dimensional. **Fact:** (You can use this fact without proof). One can define *matrix norm* and the notion of convergence for series of $n \times n$-matrices. If a series of matrices $\sum_{k=1}^{\infty} A_k$ converges, you can work with it according to the same rules as you would with absolutely convergent series of numbers, e.g. multiply every term by scalar, add two such series, etc. For a matrix A, define

$$e^A := \mathrm{Id} + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \ldots.$$

(Note that we just plugged in the matrix A into the usual Taylor series for the exponential function). You can assume without proof that the series converges.

Prove that if A is diagonalizable and $A = CDC^{-1}$, where D is the diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A on the diagonal, then $e^A = Ce^DC^{-1}$, and e^D is the diagonal matrix with $e^{\lambda_1}, \ldots, e^{\lambda_n}$ on the diagonal.

(d) Now let A be a nilpotent matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Compute e^A.

Remark. In the same way, any function that is defined by a convergent Taylor series can be applied easily to a diagonalizable matrix, and less easily, to a general matrix.

3. **Systems of linear differential equations.** In this question, we consider the simplest systems of differential equations with constant coefficients. We typically denote an unknown function by x, and the variable by t, so for example, the equation $x' = x$ has a general solution $x = ce^t$, where c is a constant.

(a) Write the system of differential equations

$$\begin{aligned} x_1' &= x_1 + 4x_2 \\ x_2' &= x_1 + x_2 \end{aligned}$$

in the matrix form.

(b) Suppose we have a system of differential equations written in the form $x' = Ax$, where $x = (x_1, \ldots, x_n)$ is an n-tuple of functions of t, and A is a matrix of scalars. Suppose A diagonalizes (as in the previous problem), and $A = CDC^{-1}$, where D is a diagonal matrix (with the eigenvalues of A on the diagonal). Prove that a general solution to this system has the form $x(t) = Ce^{Dt}C^{-1}B$, where we still think of x as an n-tuple of functions, the exponential is the matrix exponential as in the previous problem, and B is an n-vector of constants.

Remark: This is not surprising, it basically says that the solution has an expected form $x = e^{At}B$, analogously to the 1-dimensional case.

(c) Using the previous part, solve the system from (a).

(d) Now suppose you have a higher order homogenous differential equation with constant coefficients, namely, an equation of the form

$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0x = 0.$$

Then it can be converted into a system of linear equations by the following trick: let $x_1 = x$, $x_2 = x'$, $x_3 = x''$, ... $x_n = x^{(n-1)}$. Rewrite the given equation as a linear system as above using this notation, and write down the matrix for this system.

(e) Now recall the method of solving higher-order differential equations using the characteristic equation: take a dummy variable r, and replace every derivative with the corresponding power of r, obtaining the polynomial equation $r^n + a_{n-1}r^{n-1} + \cdots + a_0 = 0$ (this is called the *characteristic equation* of the given DE). Let $\lambda_1, \ldots, \lambda_n$ be the roots of the characteristic equation (assume they are distinct). Prove that λ_i are the eigenvalues of the matrix you obtained in the previous part of the problem. (In other words, the characteristic equation is the characteristic polynomial of that matrix). Derive the general form of the solution to this differential equation.

Hint: remember from homework, we had a "Frobenius companion matrix" for a given characteristic polynomial).

(f) Finally, an example of the situation when there is a repeated eigenvalue. Consider the system

$$\begin{aligned} x_1' &= \lambda x_1 + x_2 \\ x_2' &= \lambda x_2 \end{aligned}$$

Show that $(te^{\lambda t}, e^{\lambda t})$ is a solution. Derive a general form of the solution.

Remark: the matrix of this system is a 2×2-Jordan block. We will not go in-depth of this, but it shows concretely that the Jordan form of the matrix of the system of the differential equations plays a key role in determining how many associated vectors of the form $t^k e^{\lambda t}$ to consider.

4. **Linear recurrences.** Let V be the complex vector space of all sequences $\bar{x} = (x_0, x_1, \ldots, x_n, \ldots)$ with $x_i \in \mathbb{C}$ (it is infinite-dimensional).

We say that a sequence $\bar{x} \in V$ satisfies a *linear relation of degree* k if there exist coefficients $c_0, \ldots, c_{k-1} \in \mathbb{C}$ with $c_0 \neq 0$ such that $x_{n+k} = \sum_{i=0}^{k-1} c_i x_{n+i}$ for all $n \geq 0$. The goal of this problem is to explore how to find all the sequences satisfying a given linear recurrence relation of degree k. We define the *characteristic polynomial* of a linear recurrence relation by

$$p(t) = t^k - \sum_{i=0}^{k-1} c_i t^i.$$

(a) Write down in this form the linear relation defining the Fibonacci sequence. What is its degree?

(b) Explain why we require $c_0 \neq 0$.

(c) Let $L : V \to V$ be the *left shift operator*: $L(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$. Prove that a sequence $\bar{x} \in V$ satisfies a linear recurrence relation if and only if it lies in the kernel of the linear operator $p(L) : V \to V$ (see the above problem for the meaning of $p(L)$).

(d) Prove that a sequence satisfying a linear recurrence of degree k is determined by the k initial values $x_0, \ldots, x_{k-1}$. Conclude that $\ker p(L)$ has dimension k.

(e) Assuming that $p(t)$ has k distinct roots $\lambda_1, \ldots, \lambda_n$, find a basis for $\ker p(L)$.

(f) Let $(F_0, F_1, \ldots, F_{k-1})$ be any numbers. Show that the system of k equations $\sum_{i=0}^{k-1} A_i \lambda_i^j = F_j$ $(1 \leq j \leq k)$ in the unknowns A_i has a unique solution.

(g) Show that for any recurrence relation of degree k, any initial k-tuple of values extends to a unique solution of the recurrence relation.

(h) Find a non-recursive formula for the n-th Fibonacci number.

Remark. In fact, the problem on linear recurrences can be viewed as a "discrete version" of the previous problem on differential equations.

5. **Fourier series.** In this problem we use the standard inner product $(f, g) = \int_{-\pi}^{\pi} f(x)g(x)dx$ on the space of continuous functions on $(-\pi, \pi)$.

(a) Show that $\left\{\frac{1}{\sqrt{2\pi}}\right\} \cup \left\{\frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(nx)\right\}_{n=1}^{\infty}$ is an orthonormal system there.

(b) Let a_0, a_n, b_n be the coefficient of $f(x) = 2\pi\,|x| - x^2$ with respect to $\frac{1}{\sqrt{2\pi}}$, $\frac{1}{\sqrt{\pi}}\cos(nx)$, $\frac{1}{\sqrt{\pi}}\sin(nx)$. Find these.

(c) Show that for any x, the series $\frac{1}{\sqrt{2\pi}}a_0+\frac{1}{\sqrt{\pi}}\sum_{n=1}^{\infty}\left(a_n\cos(nx)+b_n\sin(nx)\right)$ is absolutely convergent.

Facts: 1. The system above is *complete*, in that the only function orthogonal to the span is the zero function. If we denote the partial sums $\left(S_N f\right)(x) = a_0\frac{1}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}}\sum_{n=1}^{N}\left(a_n\cos(nx)+b_n\sin(nx)\right)$, this shows $S_N f \xrightarrow[N\to\infty]{} f$ "on average" in the sense that $\|f - S_N f\|^2_{L^2(-\pi,\pi)} = \int_{-\pi}^{\pi}\left|f(x)-\left(S_N f\right)(x)\right|^2\,dx \xrightarrow[N\to\infty]{} 0$ (in fact, this holds for any f such that $\int_{-\pi}^{+\pi}|f(x)|^2\,dx<\infty$).

2. For any $x\in(-\pi,\pi)$ if the sequence of real numbers $\{\left(S_N f\right)(x)\}_{N=1}^{\infty}$ converges, and if f is continuous at x, then limit of the sequence is $f(x)$.

(d) Conclude that $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$, a discovery of Euler's.

6. **Applications of Cauchy-Schwarz inequality.**

(a) Prove that if the series $\sum_{n=1}^{\infty}|a_n|^2$ and $\sum_{n=1}^{\infty}|b_n|^2$ converge, then the series $\sum_{n=1}^{\infty}a_n b_n$ converges absolutely.

(b) If $a_1+a_2+\cdots+a_n=n$ show that $a_1^4+\cdots+a_n^4\geq n$.

Hint: Apply Cauchy-Schwarz twice.

7. **Linear dual and the adjoint map.** Recall that if V is a linear space over $\mathbb{R}$, its *linear dual* V^* is the space of linear functionals $f:V\to\mathbb{R}$. We denote by $\langle\ ,\ \rangle$ the map

$$\langle\ ,\ \rangle : V^*\times V\to\mathbb{R},\quad \langle f,v\rangle = f(v).$$

(a) Prove that any inner product $(\ ,\ )$ on V defines an isomorphism between V and V^* by the formula $v\mapsto(w\mapsto(w,v))$.

(b) Let V and W be linear spaces, and let $A:V\to W$ be a linear map. Define the *adjoint* of A to be the map $A^{ad}:W^*\to V^*$ defined by the relation

$$\langle A^{ad}(f),v\rangle = \langle f,A(v)\rangle\quad\text{for all } v\in V, f\in W^*.$$

Prove that if $V=W$ is a Euclidean space, then A is self-adjoint if and only if $A^{ad}=A$ if we use the identification of V and V^* as in Part (a).

8. **Lorentz transformation.** In the course of his researches on electromagnetism, Henri Poincaré wrote down the following map $L_v:\mathbb{R}^2\to\mathbb{R}^2$ which he called the "Lorentz transformation":

$$L_v\begin{pmatrix}x\\ t\end{pmatrix} = \gamma_v\cdot\begin{pmatrix}x-vt\\ t-vx\end{pmatrix}.$$

Here v is a real parameter such that $|v| < 1$ and γ_v is also a number, defined by $\gamma_v = (1 - v^2)^{-1/2}$.

(a) Suppose $v = 0.6$ so that $\gamma_v = (1 - 0.6^2)^{-1/2} = 1.25$. Calculate $L_v \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, $L_v \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $L_v \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Check that $L_v \begin{pmatrix} 2 \\ 3 \end{pmatrix} = L_v \begin{pmatrix} 3 \\ 2 \end{pmatrix} + L_v \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

(b) Show that L_v is a linear transformation and write down its matrix (it should depend on v).

(c) ("Relativistic addition of velocities") Let $v, v' \in (-1, 1)$ be two parameters. Show that $L_v \circ L_{v'} = L_u$ for $u = \frac{v+v'}{1+vv'}$.

It is a fact that if $v, v' \in (-1, 1)$ then $\frac{v+v'}{1+vv'} \in (-1, 1)$ as well. *Hint*: Start by showing $\gamma_v \gamma_{v'} = \frac{\gamma_u}{1+vv'}$.

9. **The quaternions and $\mathrm{SO}_3(\mathbb{R})$.** Let $\mathbb{H}$ be the 4-dimensional vector space over $\mathbb{R}$ with the basis labelled $\mathbf{1}$, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, so that an element of $\mathbb{H}$ has the form $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Define *multiplication* on $\mathbb{H}$ by the rules:

$$\mathbf{1}x = x \text{ for every } x \in \mathbb{H}.$$
$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1,$$
$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

These rules extend *by distributivity* to give multiplication for any two elements of $\mathbb{H}$. The real vector space $\mathbb{H}$ with this multiplication structure is called the *division algebra of the real quaternions*. It was discovered by William Rowan Hamilton in 1843.

(a) Define the *quaternion norm* by $\|a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}\|^2 = a^2+b^2+c^2+d^2$. Prove that $\|xy\| = \|x\|\|y\|$ for any $x, y \in \mathbb{H}$.

(b) Prove that for every $x \in \mathbb{H}$ there exists $x^{-1} \in \mathbb{H}$ such that $xx^{-1} = x^{-1}x = \mathbf{1}$.

(c) Recall that when we view $\mathbb{C}$ as a 2-dimensional real vector space, multiplication by a complex number $a + bi$ in $\mathbb{C}$ corresponds to the matrix $\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$. Write the 4×4 - matrix that corresponds to the operation of multiplication by $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ on $\mathbb{H}$.

(d) Recall that computation with rotations on $\mathbb{R}^2$ can be made easy if we encode the rotations as multiplication by complex numbers of absolute value 1. In the same spirit, it turns out that rotations in $\mathbb{R}^3$ can be encoded using the quaternion multiplication. Namely, let $U = \{q \in \mathbb{H} : \|q\| = 1\}$. For every $q \in U$, we can define the operation of *conjugation by q* on $\mathbb{H}$ by $x \mapsto qxq^{-1}$. Let W be the 3-dimensional subspace of $\mathbb{H}$ consisting of the quaternions with zero real part, that is, of the form $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Prove that for every q, the conjugation by q is, in fact, a linear map from W to W. Prove further that if we identify W with $\mathbb{R}^3$ with the standard dot product, the conjugation by q is an isometry (i.e., corresponds to an orthogonal matrix in the standard basis). Prove further that the determinant of this matrix is 1.

Remark. In a fancy language, we just constructed a *group homomorphism* from the group of unit quaternions U to the group $\mathrm{SO}_3(\mathbb{R})$. This homomorphism turns out to be surjective, and every element of $\mathrm{SO}_3(\mathbb{R})$ has exactly two preimages in U.

Acknowledgment: I am grateful to Professor Silberman for Problems 5 and 8.