MATH21B – FINAL REVIEW
SPRING 2018, HARVARD UNIVERSITY

Lecture 1 & 2: solving systems of linear equations

· We can write a system of linear equations as an augmented matrix $[A \mid \vec{b}]$.

· In *Gauss-Jordan elimination* we simplify the (augmented) matrix using the following three operations (the three S's)
  (I) Swap two rows.
  (II) Scale a row by a non-zero number.
  (III) Substract a multiple of a row from another.

· During Gauss-Jordan elimination, our goal is to put the matrix in *reduced row echelon form*:
  – In a non-zero row, the first non-zero entry is 1.
  – If a column has a leading 1, all the other entries in the column are 0.
  – In a row with a leading 1, every row above has a leading 1 to the left.

$$\begin{bmatrix} \boxed{1} & 2 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

↑ free variable

· The *rank* of A is the number of leading 1's in $\text{rref}(A)$.

Lecture 3: the number of solutions

· We can read off the number of solutions to a system of m linear equations $A\vec{x} = \vec{b}$ in n unknowns from $\text{rref}([A \mid \vec{b}])$:
  – If $\text{rank}(A) = \text{rank}(A \mid \vec{b}) = n$ then there is exactly one solution.
  – If $\text{rank}(A) < \text{rank}(A \mid \vec{b})$ then there are no solutions.
  – If $\text{rank}(A) = \text{rank}(A \mid \vec{b}) < n$ then there are infinitely many solutions.

Lecture 4: linear transformations

· A transformation T is *linear* if $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ and $T(\lambda\vec{x}) = \lambda T(\vec{x})$ (which implies $T(\vec{0}) = \vec{0}$, a helpful check to see whether a transformation is linear). It is then implemented by a matrix A, that is $T(\vec{x}) = A\vec{x}$.

· To find the matrix of a linear transformation, take the ith column to be the image of the ith standard vector.

· An $n \times n$-matrix A is invertible if and only if its rank is n. Then we can find the matrix for the inverse of the linear transformation $T(\vec{x}) = A\vec{x}$ by putting the (super-)augmented matrix $[A \mid \text{id}_n]$ in reduced row echelon form.

Lecture 5: transformations in geometry

- To see what a linear transformation does look at the columns.
- Particular useful linear transformations to remember are: rotations, dilations, reflections, and projections. Of these the first three are invertible, but projections are not.

Lecture 6: matrix algebra

- You can add matrices and multiply them with a number:
$$\lambda \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}.$$
- You can multiply an $(m \times n)$-matrix A with an $(n \times k)$-matrix B:
$$AB = \begin{bmatrix} | & \cdots & | \\ A(\vec{b}_1) & \cdots & A(\vec{b}_k) \\ | & \cdots & | \end{bmatrix}$$
where $\vec{b}_i$ denotes the ith column of B. Matrix multiplication is associative, $A(BC) = (AB)C$. It is not commutative, in general $AB \neq BA$ (though this can occasionally happen for particular A and B).
- If A is the $(m \times n)$-matrix for $T\colon \mathbb{R}^n \to \mathbb{R}^m$, and B is the $(n \times k)$-matrix for $S\colon \mathbb{R}^k \to \mathbb{R}^n$, then AB is the $(m \times k)$-matrix for $T \circ S\colon \mathbb{R}^k \to \mathbb{R}^m$.

Lecture 7: images and kernels

- The kernel of an $(m \times n)$-matrix A is the set of vectors $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \vec{0}$:
$$\ker(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}.$$
This tells you about the uniqueness of solutions.
- The image of an $(m \times n)$-matrix A is the set of vectors $\vec{y}$ in $\mathbb{R}^m$ that can be obtained as $A\vec{x}$ for some $\vec{x} \in \mathbb{R}^n$:
$$\operatorname{im}(A) = \{A\vec{x} \in \mathbb{R}^m \mid \vec{x} \in \mathbb{R}^n\}$$
This tells you about the existence of solutions.
- The kernel and image of an $(m \times n)$-matrix A are subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively; this means they are closed under addition of vectors and multiplication of vectors by a scalar.
- A subspace $V \subset \mathbb{R}^n$ is said to be spanned by a collection of $\vec{x}_1, \vec{x}_2, \ldots$ of vectors if every vector in V can be written as a linear combination of $\vec{x}_1, \vec{x}_2, \ldots$.
- The image $\operatorname{im}(A)$ is spanned by the columns of A corresponding to those columns of $\operatorname{rref}(A)$ containing a leading 1. To obtain spanning vectors for the kernel $\ker(A)$, parametrize the solutions of $A\vec{x} = \vec{0}$ using $\operatorname{rref}(A)$.

Lecture 8: bases

- A collection $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ of vectors in $\mathbb{R}^n$ is *linearly independent* if $a_1\vec{v}_1 + \ldots + a_k\vec{v}_k = 0$ implies that $a_1 = \ldots = a_k$; this means that there are no non-zero relations between them. If it is not linearly independent it said to be *linearly dependent*; in that case we can express at least one of the $\vec{v}_i$'s in terms of the others.

- To find a linear independent subset among a collection of vectors, put them as the columns of a matrix A and solve $A\vec{x} = \vec{0}$ by Gauss-Jordan elimination. If there is a unique solution (i.e. there is no free variable), then they are linearly independent. In there is not a unique solution, the vectors corresponding to the columns containing a leading one will be linearly independent.
- A linearly independent spanning set for a subspace $V \subset \mathbb{R}^n$ is called a *basis*. Every vector in V is a unique linear combination of vectors in the spanning set.
- The spanning sets of $\ker(A)$ and $\operatorname{im}(A)$ read off from the $\operatorname{rref}(A)$ will be linearly independent, and thus give a basis for $\ker(A)$ and $\operatorname{im}(A)$ respectively.

Lecture 9: dimension

- The dimension of a linear subspace $V \subset \mathbb{R}^n$ is the number of elements in a basis of V.
- The dimension of the image of A is equal to the rank of A. The dimension of the kernel of A is called the *nullity* of A.
- The rank-nullity theorem says that the $\operatorname{rank}(A) + \operatorname{nullity}(A)$ equals the number of columns on A. To see this, recall that $\operatorname{rank}(A)$ equals the number of columns of $\operatorname{rref}(A)$ containing a leading one, and $\operatorname{nullity}(A)$ equals the number of free variables (which correspond to those columns of $\operatorname{rref}(A)$ not containing a leading one).

Lecture 10: coordinates

- Switching to a different basis is useful when solving problems.
- When going from the standard $(\vec{e}_1, \ldots, \vec{e}_n)$ of $\mathbb{R}^n$ to $(\vec{a}_1, \ldots, \vec{a}_n)$ form the $(n \times n)$-matrix S with these vectors as columns. The components of a vector $\vec{v}$ with respect to the new basis are $S^{-1}\vec{v}$. A matrix A expressed in terms of the new basis is given by $B = S^{-1}AS$.
- If $B = S^{-1}AS$ then we say that B is *similar* to A.

Lecture 11: linear spaces

- To give the structure of a *linear space* on a set V is to give ways to add elements and scale them, together with a zero vector $\vec{0}$ (this is somewhat imprecise, see Definition 4.1.1 of Bretscher for details).
- Most of our linear spaces V arise as linear subspaces and then we need to check that (i) $\vec{0}$ is in V, (ii) V is closed under addition, (iii) V is closed under scalar multiplication.
- The following are common examples of linear spaces:
  - The set $C(\mathbb{R})$ of continuous real-valued functions $f\colon \mathbb{R} \to \mathbb{R}$. The addition and scaling is pointwise, and $\vec{0}$ is the function that is 0 everywhere.
  - The set $P_n(\mathbb{R})$ of polynomials with real coefficients of degree $\leq n$. This is a linear subspace of $C(\mathbb{R})$.
  - The set $M_{m,n}(\mathbb{R})$ of $(m \times n)$-matrices with real entries. The addition and scaling are entry-wise, and $\vec{0}$ is the matrix with all entries 0.

Lecture 12: orthogonality

- Two vectors $\vec{x}$, $\vec{y}$ are *orthogonal* if $\vec{x}\cdot\vec{y}=0$. The *orthogonal complement* $V^{\perp}$ of a linear subspace V of $\mathbb{R}^n$ is given by $\{\vec{x} \mid \vec{x}\cdot\vec{v}=0 \text{ for all } \vec{v}\in V\}$. It is a linear subspace and its dimension $n-\dim(V)$.
- The *length* of a vector $\vec{x}$ is $||\vec{x}||=\sqrt{\vec{x}\cdot\vec{x}}$. Pythagoras says that $||\vec{x}+\vec{y}||^2=||\vec{x}||^2+||\vec{y}||^2$ if $\vec{x}$ and $\vec{y}$ are orthogonal. There are some standard inequalities:

$$\text{Cauchy-Schwartz inequality } |\vec{x}\cdot\vec{y}|\leq ||\vec{x}||||\vec{y}|| \text{ (with equality if } \vec{x} \text{ and } \vec{y} \text{ are parallel)},$$

$$\text{Triangle inequality: } ||\vec{x}+\vec{y}||\leq ||\vec{x}||+||\vec{y}||.$$

- The *angle* α between two vectors is given by $\cos(\alpha)=(\vec{x}\cdot\vec{y})/(||\vec{x}||||\vec{y}||)$. Thus orthogonal vectors have angle $90°$ (or $270°$) between them.
- A collection of vectors is *orthonormal* if they are are pairwise orthogonal and have length 1. They are automatically linearly independent.
- If $\vec{a}_1,\ldots,\vec{a}_k$ in $\mathbb{R}^n$ are orthonormal, then the *orthogonal projection* onto their span is given by QQ^T, where Q is the matrix with columns given by the $\vec{a}_i$'s and Q^T is its *transpose*, obtained by switching the rows and columns (flipping the matrix).

Lecture 13: Gram-Schmidt and QR

- The Gram-Schmidt procedure produces an orthonormal basis $\{\vec{w}_1,\ldots,\vec{w}_k\}$ from a basis $\{\vec{v}_1,\ldots,\vec{v}_k\}$. It is given by iteratively scaling and making the dot products 0:

$$\begin{aligned}
\vec{w}_1 &= \frac{1}{||\vec{v}_1||}\vec{v}_1\\
\vec{w}_2 &= \frac{1}{||\vec{v}_2-(\vec{v}_2\cdot\vec{w}_1)\vec{w}_1||}(\vec{v}_2-(\vec{v}_2\cdot\vec{w}_1)\vec{w}_1)\\
&\vdots\\
\vec{w}_k &= \frac{1}{||\vec{v}_k-(\sum_{i=1}^{k-1}(\vec{v}_k\cdot\vec{w}_i)\vec{w}_i)||}\left(\vec{v}_k-(\sum_{i=1}^{k-1}(\vec{v}_k\cdot\vec{w}_i)\vec{w}_i)\right)
\end{aligned}$$

- If A is the matrix with columns given by the $\vec{v}_i$'s, then we have $A=QR$ with Q a matrix with orthonormal basis for columns, given by the $\vec{w}_i$'s,

$$Q=\begin{bmatrix}\vdots & \vdots & \cdots\\ \vec{w}_1 & \vec{w}_2 & \cdots\\ \vdots & \vdots & \cdots\end{bmatrix}$$

and R an upper triangular matrix given by expressing the $\vec{v}_i$'s in terms of the $\vec{w}_i$'s

$$\vec{v}_i=r_{ii}\vec{w}_i+\sum_{j<i}a_{ji}\vec{w}_j$$

and taking the matrix

$$R=\begin{bmatrix}r_{11} & a_{12} & a_{13} & \cdots\\ 0 & r_{22} & a_{23} & \cdots\\ 0 & 0 & r_{33} & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{bmatrix}.$$

This is called the QR-decomposition of A, and it is useful for computations.

Lecture 14: orthogonal transformations

- The transpose A^T is obtained by switching columns and rows. It satisfies $(A+B)^T = A^T + B^T$ and $(\lambda A)^T = \lambda A^T$. Slightly harder to see is that it satifies $(AB)^T = B^T A^T$ (so *switches* order of multiplication!), and $(A^{-1})^T = (A^T)^{-1}$).
- An $(n \times n)$-matrix is *orthogonal* if $A^T A = \mathrm{id}_n$. This is equivalent to the columns of A being an orthonormal basis of $\mathbb{R}^n$. In particular, in a QR-decomposition of an $(n \times n)$-matrix the Q is orthogonal.
- The inverse of an orthogonal matrix A is A^T. The product of two orthogonal matrices is orthogonal.
- Orthogonal matrices preserve angles and length. Examples are rotations and reflections.

Lecture 15: least squares and data fitting

- The *least squares solution* to a system of linear equations $A\vec{x} = \vec{b}$ is a solution $\vec{x}^*$ of $A^T A \vec{x}^* = A^T \vec{b}$. These always exist and are unique if $\ker(A) = \{0\}$, in which case $A^T A$ is invertible and then given by $\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$.
- If $\ker(A) = \{0\}$ then $A^T A$ is invertible and the matrix for orthogonal projection onto $\mathrm{im}(A)$ is given by $A(A^T A)^{-1} A^T$. This simplifies to the usual formula when the columns of A are orthonormal.

Lecture 16: determinants I

- Leibniz's definition of the *determinants* $\det(A)$ of an $(n \times n)$-matrix is given by
$$\det(A) = \sum_{\pi} \mathrm{sign}(\pi) a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)}$$
where $a_{i\pi(i)}$ denotes the entry in the ith row and $\pi(i)$th column.
- You can compute it recursively by *Laplace expansion*. Let A_{ij} be the matrix obtained by A by deleting the ith row and jth column. Then we have
$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j})$$
where a_{1j} are the entries of A on the first row. You can pick any other row or column to expand, as long as you introduce the appropriate sign.
- If a matrix is block sum of two smaller matrices:
$$M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$
then $\det(M) = \det(A)\det(B)$.
- You can obtain the determinant using Gauss-Jordan elimination. If A is a matrix and during Gauss-Jordan elimination you scale by factors $\lambda_1, \ldots, \lambda_k$ and swap rows s times, then
$$\det(A) = (-1)^s \frac{1}{\lambda_1 \cdots \lambda_k} \det(\mathrm{rref}(A)).$$
In particular $\det(A) \neq 0$ if and only if A is invertible.

Lecture 17: determinants II

- Geometrically, $|\det(A)|$ is the volume of the image of the unit cube under the linear transformation corresponding to A. If the sign $\det(A)$ is positive A preserves the orientation, if it is negative A flips the orientation.
- Here are the important properties of the determinant: (i) $\det(AB) = \det(A)\det(B)$, (ii) $\det(A^T) = \det(A)$, (iii) $\det(A^{-1}) = 1/\det(A)$, (iv) $\det(SAS^{-1}) = \det(A)$, (v) $\det(A^n) = \det(A)^n$.
- Let A_{ij} be the matrix obtained by A by deleting the ith row and jth column. Then we have
$$(A^{-1})_{ij} = (-1)^{i+j}\det(A_{ji})/\det(A).$$
- Questions you might ask yourself before computing a determinant $\det(A)$?
  - Is A upper or lower triangular?
  - Is A a block sum of two smaller matrices?
  - Is A non-invertible (so that $\det(A) = 0$)?
  - Is A a product of matrices?
  - Are there only a few non-zero patterns?
  - Is there a row or column with lots of zeroes to start Laplace expansion?
  - Can I swap or add rows or columns to simplify A?
  - Can I quickly row reduce to the upper-triangular case?
  - (Later: can we compute the eigenvalues of A?)

Lecture 18: eigenvectors, eigenvalues, and discrete dynamical systems

- If $\vec{v}$ is a non-zero vector such that $A\vec{v} = \lambda\vec{v}$, then $\vec{v}$ is called an *eigenvector* of A and λ is its *eigenvalue*. The eigenvalues are the roots of the characteristic polynomial $\det(A - \lambda \mathrm{id}_n)$, whose constant coefficient is $\det(A)$. The eigenvectors for eigenvalue λ_i are obtained by finding $\ker(A - \lambda_i \mathrm{id}_n)$.
- If $\vec{v} \mapsto A\vec{v}$ describes the change in state after one time step of a discrete dynamical, then we can find the trajectories $A^s\vec{v}$ by decomposing $\vec{v}$ into eigenvectors; if $\vec{v} = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n$ with $\vec{v}_i$ an eigenvector of A with eigenvalue λ_i, then
$$A^s\vec{v} = c_1\lambda_1^s\vec{v}_1 + \ldots + c_n\lambda_n^s\vec{v}_n.$$

Lecture 19: eigenspaces

- The coefficient of λ^{n-1} in the characteristic polynomial is $(-1)^{n-1}\mathrm{tr}(A)$, where the *trace* $\mathrm{tr}(A)$ is the sum of the diagonal entries. The algebraic multiplicity of an eigenvalue is its multiplicity as a root of the characteristic polynomials and the trace $\mathrm{tr}(A)$ is also equal to the sum of eigenvalues counted with algebraic multiplicity.
- If λ is an eigenvalue of A, then $E_\lambda = \ker(A - \lambda \mathrm{id}_n)$ is the λ-*eigenspace* of A. Its dimension is called the *geometric multiplicity* of λ. We have that $1 \leq$ geometric multiplicity $\leq$ algebraic multiplicitivity.
- If an $(n \times n)$-matrix A has n distinct eigenvalues, then there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A, which is called an *eigenbasis*.

Lecture 20: similarity and diagonalization

- Recall that $(n \times n)$-matrices A and B are *similar* if there is an invertible matrix S such that $B = S^{-1}AS$ (i.e. they differ by a change of basis). Similar matrices have the same determinant, same trace, same eigenvalues, same characteristic polynomial, same algebraic multiplicities, and same geometric multiplicities. The converse is not true.
- An $(n \times n)$-matrix A is *diagonalizable* if it is similar to a diagonal matrix B. Then in $B = S^{-1}AS$, the columns of S are an eigenbasis for A, and we diagonalize a matrix by finding an eigenbasis. Thus if A has n different eigenvalues it is diagonalizable.

Lecture 21: complex eigenvalues

- Complex numbers $z \in \mathbb{C}$ are numbers of the form $z = a + bi$ where $i = \sqrt{-1}$. You can picture these as points in the plane. Euler's formula says that $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, which gives $e^{\pi i} + 1 = 0$ for $\theta = \pi$. The absolute value $|a + bi|$ is given by $\sqrt{a^2 + b^2}$.
- Every polynomial of degree n has n complex roots (counted with algebraic multiplicity). This implies that every matrix has a complex eigenvector, though it is still might not have an eigenbasis.

Lecture 22: stability

- A discrete dynamical system $\vec{x}(t+1) = A\vec{x}(t)$ is *asymptotically stable* if $\vec{x}(t) \to \vec{0}$ as $t \to \infty$. This is true if and only if all (complex) eigenvalues λ_i of A satisfy $|\lambda_i| < 1$.

Lecture 23: symmetric matrices

- If an $(n \times n)$-matrix A satisfies $A^T = A$ it is called *symmetric*, and if it satisfies $A^T = -A$ it is called *anti-symmetric*.
- Symmetric matrices have real eigenvalues and eigenvectors for different eigenvalues are orthogonal.
- The *spectral theorem* says that a symmetric matrix can always be diagonalized and has an orthonormal eigenbasis. Thus there exists an orthogonal matrix S such that $S^{-1}AS$ is diagonal with diagonal entries the real eigenvalues of A.

Lecture 24: differential equations I

- An equation of the form $\frac{dx(t)}{dt} = f(x(t))$ is called a *(ordinary) differential equation*. You can solve them when given an *initial condition* $x(0)$. They can often be solved by *separation of variables*.
- A linear differential equation is one of the form $\frac{dx(t)}{dt} = cx(t)$ for some constant $c \in \mathbb{R}$ (or $c \in \mathbb{C}$). The solution with initial condition $x(0)$ is given by $x(t) = e^{ct}x(0)$.
- A system of linear differential equation is one of the form $\frac{d\vec{x}(t)}{dt} = A\vec{x}(t)$, with A an $(n \times n)$-matrix and $\vec{x}(t) = [x_1(t), \ldots, x_n(t)]^T$. To solve $\frac{d\vec{x}(t)}{dt} = A\vec{x}(t)$ with initial condition $\vec{x}(0)$, diagonalize A and write $\vec{x}(0)$ as a linear combination $\vec{x}(0) = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n$ of eigenvectors $\vec{v}_i$ with eigenvalues λ_i. Then the solution is $\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + \cdots + c_n e^{\lambda_n t}\vec{v}_n$.

- The system of linear differential equation $\frac{d\vec{x}(t)}{dt} = A\vec{x}(t)$ is *asymptotically stable* if $\vec{x}(t) \to \vec{0}$ as $t \to \infty$. This holds if and only if $\mathrm{Re}(\lambda_j) < 0$ for all eigenvalues λ_j.
- A *phase portrait* of a system of linear differential equations $\frac{d\vec{x}(t)}{dt} = A\vec{x}(t)$ for $n = 2$ is given by drawing a vector $A\vec{x}$ at $\vec{x} \in \mathbb{R}^2$. The solutions are curves tangent to these vectors.

Lecture 25: differential equations II

- The *harmonic oscillator* is the name for the differential equation $\frac{d^2x(t)}{dt^2} = -k^2x(t)$. It has solutions $x(t) = a\cos(kt) + b\cos(kt)$, with a, b depending on the initial condition, and is solved by adding a dummy variable to make it a system of two linear differential equations.
- Recall the system of linear differential equation $\frac{d\vec{x}(t)}{dt} = A\vec{x}(t)$ is *asymptotically stable* if $\vec{x}(t) \to \vec{0}$ as $t \to \infty$. This holds if and only if $\mathrm{Re}(\lambda_j) < 0$ for all eigenvalues λ_j. For $n = 2$, this can be rephrased as $\det(A) > 0$ and $\mathrm{tr}(A) < 0$.

Lecture 26: non-linear dynamical systems

- Analyzing a system of differential equations

$$\frac{dx}{dt} = f(x, y)$$
$$\frac{dy}{dt} = g(x, y)$$

is done by following the steps:

(1) Find the *nullclines* $\{(x, y) \mid f(x, y) = 0\}$ and $\{(x, y) \mid g(x, y) = 0\}$.
(2) The horizontal and vertical nullclines intersect in the *equilibria*, those points (x, y) where $f(x, y) = 0$ and $g(x, y) = 0$.
(3) Start drawing a phase portrait by finding the signs of $f(x, y)$ and $g(x, y)$ in the regions bounded by the nullclines.
(4) By linearizing the system near an equilibrium (a, b) we get a system of linear differential equations (the matrix is called the *Jacobian*)

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} \frac{df(a,b)}{dx} & \frac{df(a,b)}{dy} \\ \frac{dg(a,b)}{dx} & \frac{dg(a,b)}{dy} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

This describes the behavior near (a, b). In particular, we use it to determine the stability of the equilibria.
(5) Finish drawing the phase portrait by understanding some representative trajectories.

Lecture 27: differential operators

- We are interested in the linear spaces C^∞ of smooth real-valued functions on $\mathbb{R}$, and C^∞_{per} of smooth periodic real-valued functions on $\mathbb{R}$ with period 2π. These contain interesting subspaces: $\mathcal{P} \subset C^\infty$ given by the polynomials (this is the span of $1, x, x^2, \ldots$) and $\mathcal{T} \subset C^\infty_{\mathrm{per}}$ of trigonometric polynomials (this is the span of $1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots$).
- On both C^∞ and C^∞_{per}, the map D sending f to $\frac{df}{dt}$ is a linear transformation: $D0 = 0$, $D(\lambda f) = \lambda D(f)$ and $D(f + g) = D(f) + D(g)$ follow directly from the rules of differentiation. A combination of powers of D, such as $D^5 - D + 1$, is called *differential operator*.

- The eigenvalues of the differential operator D on C^∞ are all $\lambda \in \mathbb{R}$ with eigenvector given by $e^{\lambda x}$. The differential operator D on C^∞_{per} only has complex eigenvalues in for $n \in \mathbb{Z}$ with eigenvector given by e^{inx}.

Lecture 28: the ODE cookbook

- When solving $p(D)f = g$ for $p(D)$ a polynomial in the differential operator D, you first find a single *particular solution* x_p satisfying $p(D)x_p = g$. Then you find the *homogeneous solutions* x_h of the differential equation $p(D)x = 0$. All solutions are given by $x_p + x_h$ where x_p is the single particular solution and x_h is a homogenous solution.
- In general, there is an operator method to find both the particular and homogeneous solutions. Because it is a lot of work, one usually follows the "cookbook."
- If p is of degree d, the space of homogenous solutions is d-dimensional. For example, for $p(D) = D^2 + bD + c = (D - \lambda_1)(D - \lambda_2)$ with λ_1, λ_2 real there are two cases: (i) if $\lambda_1 \neq \lambda_2$ we have $C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$, (ii) if $\lambda_1 = \lambda_2$ we have $C_1 e^{\lambda_1 t} + C_2 t e^{\lambda_2 t}$.

Lecture 29: Fourier series I

- The linear space C^∞_{per} of 2π-periodic smooth functions has an *inner product*:

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx.$$

This has the same properties as the dot product on $\mathbb{R}^n$, and we call $||f|| = \sqrt{\langle f, f \rangle}$ the *length* of f, and θ determined by $\cos(\theta) = \frac{\langle f,g \rangle}{||f|| \cdot ||g||}$ the *angle* between f and g.
- The functions $\frac{1}{\sqrt{2}}$, $\cos(nx)$ and $\sin(nx)$ for $n \geq 1$ form an orthonormal set with respect to this inner product.
- Every smooth 2π-periodic function f can be written as an infinite linear combination of these:

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{n \geq 1} a_n \cos(nx) + b_n \sin(nx)$$

with $a_0 = \langle f, \frac{1}{\sqrt{2}} \rangle$, $a_n = \langle f, \cos(nx) \rangle$, $b_n = \langle f, \sin(nx) \rangle$. This is like "additive synthesis" of sounds.
- If f is *even*, $f(-x) = f(x)$, then only $\frac{1}{\sqrt{2}}$ and $\cos(nx)$ for $n \geq 1$ have non-zero coefficients. If f is *odd*, $f(-x) = -f(x)$, then only $\sin(nx)$ for $n \geq 1$ have non-zero coefficients.

Lecture 30: Fourier series II

- The Pythagorean theorem $||\vec{x}||^2 = \sum_{i=1}^n x_i^2$ extends to the length of 2π-periodic functions. This is called Parseval's identity and says that if

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{n \geq 1} a_n \cos(nx) + b_n \sin(nx)$$

(with $a_0 = \langle f, \frac{1}{\sqrt{2}} \rangle$, $a_n = \langle f, \cos(nx) \rangle$, $b_n = \langle f, \sin(nx) \rangle$), then we have that

$$a_0^2 + \sum_{n \geq 1} a_n^2 + b_n^2 = ||f||^2.$$

Lecture 31: partial differential equations I

- Partial differential equations are equations involving the partial derivatives of functions of multiple variables (in contrast with the ordinary differential equations for functions of a single variable studied before). We explain how to solve two important ones: the *heat equation* (with "thermal diffusivity" μ)
$$f_t(x,t) = \mu f_{xx}(x,t)$$
and next lecture, the *wave equation*.
- To solve the heat equation for $x \in [0,\pi]$, and $t \geq 0$, with initial condition smooth function $f(x,0) = g(x)$ for $x \in [0,\pi]$ and boundary condition $f(0,t) = f(\pi,t) = 0$. To do so, we write the initial condition as a Fourier series in sines and diagonalize:
$$g(x) = \sum_{n\geq 1} b_n \sin(nx) \qquad \text{with } b_n = \frac{2}{\pi}\int_0^\pi g(x)\sin(nx)dx,$$
$$f(x,t) = \sum_{n\geq 1} b_n \sin(nx) e^{-n^2\mu t}.$$

Lecture 32: partial differential equations II

- The *wave equation* is
$$f_{tt}(x,t) = c^2 f_{xx}(x,t).$$
with c is the "wave speed."
- To solve the wave equation for $x \in [0,\pi]$, and $t \geq 0$, with initial condition smooth function $f(x,0)$ for $x \in [0,\pi]$ with initial conditions smooth functions $f(x,0) = g(x)$ and $f_t(x,0) = h(x)$ for $x \in [0,\pi]$, and boundary condition $f(0,t) = f(\pi,t) = 0$. To do so, we write the initial conditions as
$$g(x) = \sum_{n\geq 1} b_n \sin(nx) \qquad \text{with } b_n = \frac{2}{\pi}\int_0^\pi g(x)\sin(nx)dx,$$
$$h(x) = \sum_{n\geq 1} c_n \sin(nx) \qquad \text{with } c_n = \frac{2}{\pi}\int_0^\pi h(x)\sin(nx)dx,$$
$$f(x,t) = \sum_{n\geq 1} b_n \sin(nx)\cos(nct) + \frac{c_n}{nc}\sin(nx)\sin(nct).$$