Lecture notes for Math 309

Tuomas Tajakka

July 28, 2020

1 Lecture 1

1.1 Why study linear systems of ODE's?

- In many natural systems, there are several quantities depending on time, and various time derivatives of these quantities are (often linearly) related to each other by laws of nature.
- Higher order ODE's can be reduced to systems of first order ODE's!

Example 1.1.1. Consider the equation of a **damped harmonic oscillator**:

$$mu''(t) + cu'(t) + ku(t) = f(t),$$

where $u(t)$ is the displacement of the oscillator from its equilibrium at time t, k is the spring constant, c is the damping constant, m is the mass of the oscillator, and $f(t)$ is an external force depending on t. We can turn this into a system of first order linear equations by introducing a new variable $x = u'$. This way we obtain a system of linear first order ODE's:

$$\begin{aligned} u' &= x \\ x' &= -\frac{k}{m}u - \frac{c}{m}x + \frac{1}{m}f(t) \end{aligned}$$

A general system of first order ODE's is a set of equations

$$\begin{aligned} x_1'(t) &= F_1(t, x_1, \dots, x_n) \\ &\vdots \\ x_n'(t) &= F_n(t, x_1, \dots, x_n). \end{aligned}$$

Any ODE can be reduced to a system like this: the equation

$$u^{(n)}(t) = F(t, u, u', \dots, u^{(n-1)})$$

becomes

$$\begin{aligned} x_1' &= x_2 \\ &\vdots \\ x_{n-1}' &= x_n \\ x_n' &= F(t, x_1, \ldots, x_n), \end{aligned}$$

when we make the substitution $x_1 = u, x_2 = u', \ldots, x_n = u^{(n-1)}$. If all the function $F_1, \ldots, F_n$ are **linear** in the variables $x_1, \ldots, x_n$, that is,

$$x_i'(t) = a_{i1}(t)x_1(t) + \ldots + a_{in}(t)x_n(t) + b_i(t),$$

then we can write the system as a matrix equation

$$\begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{11}(t) & \cdots & a_{1n}(t) \end{pmatrix} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} + \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix},$$

or $\vec{x}'(t) = A(t)\vec{x}(t) + \vec{b}(t)$, or simply $\vec{x}' = A\vec{x} + \vec{b}$. We can use matrices and linear algebra to study such equations!

1.2 Linear systems of equations

A **linear system of equations** is a set of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m &= b_n, \end{aligned}$$

where $x_1, \ldots, x_m$ are the **variables** and $a_{11}, \ldots, a_{nm}, b_1, \ldots, b_n$ are real or complex numbers. We can write this as a matrix equation $A\vec{x} = \vec{b}$, where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

A **solution** of the system is a vector $\vec{v}$ such that $A\vec{v} = \vec{b}$. The system is called **homogeneous** if $\vec{b} = 0$, otherwise it is **inhomogeneous**. To solve a system of linear equations, we use Gaussian elimination. We first form the augmented system $[A \,|\, \vec{b}]$, and then use **row operations** to put $[A \,|\, \vec{b}]$ into **reduced row echelon form**.

Example 1.2.1. We can write the system of equations

$$\begin{aligned} 2y - 8z &= 4 \\ x - 2y + z &= 0 \end{aligned}$$

in a matrix form as

$$\begin{pmatrix} 0 & 2 & -8 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}.$$

To solve the system, we perform the row operations

$$\left(\begin{array}{ccc|c} 0 & 2 & -8 & 4 \\ 1 & -2 & 1 & 0 \end{array}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 4 \end{array}\right) \xrightarrow{R_1 \to R_1 + R_2} \left(\begin{array}{ccc|c} 1 & 0 & -7 & 4 \\ 0 & 2 & -8 & 4 \end{array}\right)$$
$$\xrightarrow{R_2 \to \frac{1}{2} R_2} \left(\begin{array}{ccc|c} 1 & 0 & -7 & 4 \\ 0 & 1 & -4 & 2 \end{array}\right)$$

and read off the solution $(x, y, z) = (4 + 7t, 2 + 4t, t)$. Here t can be any real or complex number.

The homogeneous equation always has the **trivial solution** $\vec{x} = 0$, but the inhomogeneous equation may or may not have solutions. If after some row operations one of the rows becomes zero, then the inhomogeneous equation has no solutions, unless the right side of the augmented system also has a zero on the same row.

Theorem 1.2.2 (Principle of superposition)**.** *If $\vec{v}_1$ and $\vec{v}_2$ are solutions of the homogeneous system $A\vec{x} = 0$, then so is any linear combination $a_1\vec{v}_1 + a_2\vec{v}_2$.*

Proof. Since $A\vec{v}_1 = A\vec{v}_2 = 0$, we have

$$A(a_1\vec{v}_1 + a_2\vec{v}_2) = a_1 A\vec{v}_1 + a_2 A\vec{v}_2 = 0.$$

□

2 Lecture 2

2.1 Matrix algebra

- Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ of the same size $m \times n$ can be added elementwise: $A + B = (a_{ij} + b_{ij})$.
- Any matrix A can by multiplied elementwise by any scalar c, which can be real or complex: $cA = (ca_{ij})$.
- If the number r of columns of A is equal to the number of rows of B, we can form the product AB by "taking dot products" of the rows of A with the columns of B: the ijth entry of AB is $a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ir}b_{rj}$.

These operations satisfy many of the familiar rules of algebra:

$$(A+B)+C = A+(B+C), \quad A+B = B+A, \quad c(A+B) = cA+cB, \quad (c+d)A = cA+dA$$

$$c(AB) = (cA)B = A(cB), \quad (AB)C = A(BC), \quad A(B+C) = AB+AC, \quad (A+B)C = AC+BC.$$

However, it is **not** always true that $AB = BA$, even if both products exist and have the same size. If this equation *does* hold for some particular matrices A and B, then we say that A and B **commute**. Necessarily A and B must be square matrices if the commute. Commuting matrices turn out to be important when we discuss the matrix exponential.

- Any matrix can be transposed by interchanging its rows and columns: $A^T = (a_{ji})$.
- Any matrix can be conjugated by taking complex conjugates of its entries: $\overline{A} = (\overline{a_{ij}})$. If the entries of A are all real, conjugation doesn't change A.
- The conjugate transpose $A^* = \overline{A}^T$ is called the **Hermitian conjugate** or **adjoint** of A.

The transpose and conjugate transpose satisfy the rules $(AB)^T = B^T A^T$ and $(AB)^* = B^* A^*$ respectively.

Example 2.1.1. If $A = \left(\begin{smallmatrix} i & -1 \\ 2+i & \pi \end{smallmatrix}\right)$ then $A^* = \left(\begin{smallmatrix} i & 2+i \\ -1 & \pi \end{smallmatrix}\right)$ and $A^* = \left(\begin{smallmatrix} -i & 2-i \\ -1 & \pi \end{smallmatrix}\right)$

2.2 Linear independence and bases

The vector space $\mathbb{R}^n$ consists of all column vectors of length n with real entries. A **linear combination** of vectors is a sum $c_1\vec{v}_1 + \ldots + c_k\vec{v}_k$. The vectors $\vec{v}_1, \ldots, \vec{v}_k$ are **linearly dependent** if there exists a linear combination

$$c_1\vec{v}_1 + \ldots + c_k\vec{v}_k = 0,$$

where at least one of the $c_1, \ldots, c_k$ is nonzero. If the only such linear combination is $c_1 = \ldots = c_k = 0$, i.e. the **trivial** linear combination, then the vectors are **linearly independent**.

Example 2.2.1. Are the vectors $(1,4,7)^T, (2,5,8)^T, (3,6,9)^T$ linearly independent? No, since

$$-\begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} - \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A set of vectors $\vec{v}_1, \ldots, \vec{v}_k$ **spans** $\mathbb{R}^n$, if any vector $\vec{v}$ in $\mathbb{R}^n$ can be expressed as a linear combination

$$\vec{v} = c_1\vec{v}_1 + \ldots + c_k\vec{v}_k.$$

A **basis** of $\mathbb{R}^n$ is any set of vectors that is both linearly independent and spans $\mathbb{R}^n$. Any basis of $\mathbb{R}^n$ has exactly n vectors, and if a set of n vectors either spans and is linearly independent, then it is automatically a basis. The set $\vec{e}_1 = (1,0,\ldots,0)^T, \vec{e}_2 = (0,1,\ldots,0)^T, \vec{e}_n = \ldots, (0,0,\ldots,1)^T$ is called the **standard basis** of $\mathbb{R}^n$.

Note 2.2.2. Most of our vectors will have real entries, but occasionally we need to consider complex vectors as well. Everything that we have said about linear independence and bases holds for complex vectors when we allow complex coefficients.

2.3 Linear maps and matrices

A **map** or a **function** $f : \mathbb{R}^m \to \mathbb{R}^n$ is a rule that associates to each vector in $\mathbb{R}^m$ a unique vector in $\mathbb{R}^n$. Here $\mathbb{R}^m$ is the **domain** or source of f, and $\mathbb{R}^n$ is the **codomain** or target of f.

Example 2.3.1. Define $f : \mathbb{R}^2 \to \mathbb{R}^3$ by the rule

$$f\begin{pmatrix}\theta\\ \phi\end{pmatrix} = \begin{pmatrix}\cos\theta\sin\phi\\ \sin\theta\sin\phi\\ \cos\phi\end{pmatrix}.$$

Geometrically this map takes $\mathbb{R}^2$ onto the unit sphere of $\mathbb{R}^3$. You might recognize θ and ϕ as spherical coordinates.

A map f is **linear** if it satisfies the rules $f(\vec{v}+\vec{w}) = f(\vec{v}) + f(\vec{w})$ and $f(c\vec{v}) = cf(\vec{v})$ for all vectors $\vec{v}, \vec{w}$ and real numbers c, or equivalently the single rule $f(c\vec{v}+d\vec{w}) = cf(\vec{v}) + df(\vec{w})$. Linear maps preserve the vector space structure.

Example 2.3.2. The map

$$f\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}x+7y-11z\\ -2x+3z\end{pmatrix}$$

is linear, unlike either of

$$g\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}x-y\\ x^2+y\end{pmatrix}, \qquad h\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}x+2y-2\\ -x+3y\end{pmatrix}.$$

If A is an $m\times n$ matrix, we can define a linear map $f_A : \mathbb{R}^m \to \mathbb{R}^n$ by $f_A(\vec{v}) = A\vec{v}$. Conversely, any linear map $f : \mathbb{R}^m \to \mathbb{R}^n$ gives rise to an $m \times n$ matrix A_f given by

$$A_f = [f(\vec{e}_1)\ f(\vec{e}_2)\ \ldots\ f(\vec{e}_m)],$$

which satisfies $f_{A_f} = f$. This shows that studying linear maps is (more or less) the same as studying matrices.

3 Lecture 3

3.1 Invertible matrices and determinant

A matrix is a **square matrix** if it has the same number of rows and columns. We can view $n\times n$ square matrices as maps $\mathbb{R}^n \to \mathbb{R}^n$. The $n \times n$ **identity matrix** is the matrix

$$I = I_n = \begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}.$$

It satisfies $IA = A$ for all $n \times m$ matrices A, and $BI = B$ for all $m \times n$ matrices B. The associated linear map is the identity map: $f_I(\vec{v}) = \vec{v}$ for all vectors $\vec{v}$. A square matrix A is **invertible** or **nonsingular** if there exists another square matrix B of the same size as A, such that $AB = I$ and $BA = I$. (In fact either of these equations implies the other.) This matrix B is unique if it exists, and is denoted A^{-1} and called the **inverse** of A. If no such B exists, then A is **singular**. If both A and B are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$.

If A is an invertible matrix, and we know the inverse A^{-1}, the equation $A\vec{x} = \vec{b}$ is easy to solve: just multiply both sides with the inverse to get $\vec{x} = A^{-1}\vec{b}$. This is the *unique* solution of the inhomogeneous equation. In particular, the *only* solution of the homogeneous equation is the trivial solution $\vec{x} = 0$.

The **determinant** is the unique map from the set of $n \times n$ matrices to real (or complex) numbers that is

- multilinear: $\det([\vec{v}_1 \cdots a\vec{v}_i + b\vec{w}_i \cdots \vec{v}_n]) = a\det([\vec{v}_1 \cdots \vec{v}_i \cdots \vec{v}_n]) + b\det[(\vec{v}_1 \cdots \vec{w}_i \cdots \vec{v}_n)]$ for every i,
- alternating: $\det([\vec{v}_1 \cdots \vec{v}_i\ \vec{v}_{i+1} \cdots \vec{v}_n]) = -\det([\vec{v}_1 \cdots \vec{v}_{i+1}\ \vec{v}_i \cdots \vec{v}_n])$ for every i,
- normalized: $\det(I) = 1$.

The determinant can be computed using the minor expansion. The main utility of the determinant is to tell whether a square matrix a invertible or not:

Theorem 3.1.1. *A square matrix A is invertible if and only if* $\det(A)$ *is nonzero.*

The usual way of finding the inverse of A is to form the augmented system $[A|I]$ and perform row operations to turn the right hand side into the identity matrix; the right hand side will then be the inverse: $[I|A^{-1}]$. If this is impossible, then A is singular. Here is a quick way to find the inverse of a 2 matrix:

$$\text{if} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{then} \quad A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In words, "interchange diagonals, negate off-diagonals, divide by determinant".

Test for being a basis
We can use these ideas to tell if given vectors $\vec{v}_1, \dots, \vec{v}_n$ are a basis for $\mathbb{R}^n$. Any linear combination $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$ can be written as a matrix product $A\vec{c}$, where

$$A = [\vec{v}_1\ \dots\ \vec{v}_n], \qquad \vec{c} = (c_1\ \dots\ c_n)^T.$$

Any equation of linear dependence is then a solution $\vec{c}$ of the equation $A\vec{x} = 0$. The vectors are linearly independent if and only the zero vector $\vec{c} = 0$ is the only solution of this equation, if and only if A is invertible. In light of the previous theorem,

Corollary 3.1.2. *The vectors $\vec{v}_1, \dots, \vec{v}_n$ form a basis of $\mathbb{R}^n$ if and only if the matrix $A = [\vec{v}_1 \cdots \vec{v}_n]$ has nonzero determinant.*

Example 3.1.3. Are the vectors $(1,4,7)^T, (2,5,8)^T, (3,6,9)^T$ linearly independent? No, since

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 \cdot (5 \cdot 9 - 6 \cdot 8) - 2 \cdot (4 \cdot 9 - 6 \cdot 7) + 3 \cdot (4 \cdot 8 - 5 \cdot 7) = 0.$$

3.2 Eigenvalue, eigenvector, and eigenspace

Let A be an $n \times n$ matrix. An **eigenvector** of A with **eigenvalue** λ is a *nonzero* vector $\vec{v}$ such that

$$A\vec{v} = \lambda\vec{v}.$$

This is called the **eigenvalue equation**.

Example 3.2.1. The vector $(1,-1)^T$ is an eigenvector of $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ with eigenvalue -1.

We can rewrite the eigenvalue equation as

$$(A - \lambda I)\vec{v} = 0.$$

Since $\vec{v} \neq 0$, the matrix $A - \lambda I$ must be singular, and conversely if $A - \lambda I$ is singular, then such nonzero $\vec{v}$ exists. This tells us that the eigenvalues are precisely the zeros of the **characteristic polynomial**

$$p_A(x) = \det(A - xI).$$

Example 3.2.2. Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. The characteristic polynomial of A is

$$p_A(x) = \det \begin{pmatrix} 1-x & 2 \\ 2 & 1-x \end{pmatrix} = (1-x)^2 - 4 = x^2 - 2x - 3 = (x+1)(x-3).$$

To find the eigenvectors associated to an eigenvalue λ, we solve the equation $(A - \lambda I)\vec{x} = 0$. The **eigenspace** $E_\lambda(A)$ of A for λ is the null space of the matrix $A - \lambda I$. It is a vector subspace of $\mathbb{R}^n$. The *nonzero* elements of $E_\lambda(A)$ are the eigenvectors of associated to λ.

Example 3.2.3. Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. The eigenvalues of A are 3 and -1, and we found that $(1,-1)^T$ is in $E_{-1}(A)$. In fact, $E_{-1}(A)$ is one-dimensional, and $(1,-1)^T$ is a basis for it. Let's find $E_3(A)$. We have

$$A - 3I = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}.$$

To find the null space, we row reduce to $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, and see that the null space is

$$E_3(A) = \{(t,t)^T \mid t \in \mathbb{R}\},$$

i.e. the one-dimensional subspace of $\mathbb{R}^2$ spanned by $(1,1)^T$.

Even if the matrix A has real entries, the eigenvalues can be complex numbers. The characteristic polynomial of a real matrix A has real coefficients, so its roots are either real or appear as pairs of **complex conjugates**. The eigenvectors associated to such pairs of eigenvalues are also complex conjugates of each other. In this case (and why not more generally) we should view the eigenspaces as complex vector subspaces of $\mathbb{C}^n$ rather than real subspaces of $\mathbb{R}^n$.

Example 3.2.4. If $A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$, then

$$p_A(x) = (2-x)^2 + 1 = x^2 - 4x + 5$$

has roots $\lambda_1 = 2+i$ and $\lambda_2 = 2-i$. The eigenspace $E_{2+i}(A)$ is spanned by $(i,1)^T$, and $E_{2-i}(A)$ is spanned by $(-i,1)^T$.

4 Lecture 4

4.1 Algebraic and geometric multiplicity

The characteristic polynomial $p_A(x) = \det(A - xI)$ can be factored over the complex numbers as

$$p_A(x) = (-1)^n (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}.$$

The numbers $\lambda_1, \dots, \lambda_r$ are the distinct eigenvalues of A, and the exponent m_i is the **algebraic multiplicity** of λ_i, denoted $m_a(\lambda_i)$. Since $p_A(x)$ has degree n, we have $m_1 + \cdots + m_r = n$. For each eigenvalue λ_i, the matrix $A - \lambda_i I$ is singular, so it has nontrivial null space. The dimension of the null space is the **geometric multiplicity** of λ_i, denoted $m_g(\lambda_i)$. It can be shown that the two multiplicities satisfy the inequalities

$$1 \le m_g(\lambda_i) \le m_a(\lambda_i) \le n.$$

Example 4.1.1. Let $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. The characteristic polynomial is

$$p_A(x) = (2-x)(2-x)(-1-x) = -(x-2)^2(x+1).$$

The eigenvalues are 2 and -1 with algebraic multiplicities $m_a(2) = 2, m_a(-1) = 1$. To find the geometric multiplicities, we solve the equation $(A - \lambda I)x = 0$ for $\lambda = 2, -1$. If $\lambda = 2$, this gives

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,$$

whose solutions space is one-dimensional, spanned by $(1,0,0)^T$. Thus, $m_g(2) = 1 < 2 = m_a(2)$. For $\lambda = -1$ we get the equation

$$\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,$$

whose solution is also one-dimensional, spanned by $(0,0,1)^T$. Thus, $m_g(-1) = 1$, as expected.

4.2 Diagonalization

A **diagonal matrix** is a square matrix whose only nonzero entries lie on the main diagonal. A diagonal matrix is sometimes denoted by $\mathrm{diag}(a_1, a_2, \ldots, a_n)$, where $a_1, \ldots, a_n$ are the diagonal elements listed from top left to bottom right.

Example 4.2.1. The matrices

$$\begin{pmatrix} -1 & 0 \\ 0 & 2i+3 \end{pmatrix} = \mathrm{diag}(-1, 2i+3) \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & 9001 \end{pmatrix} = \mathrm{diag}(0, \pi, 9001)$$

are diagonal.

Two square matrices A and B are said to be **similar** if there exists an invertible matrix S such that $B = SAS^{-1}$. A matrix that is similar to a diagonal matrix is called **diagonalizable**.

Suppose that A is a similar to a diagonal matrix $D = \mathrm{diag}(a_1, \ldots, a_n)$, so that we have an equation $A = SDS^{-1}$ for some matrix S. Multiplying both sides on the right by S gives $AS = SD$. We let the column vectors of S be $\vec{v}_1, \ldots, \vec{v}_n$, and write this equation more explicitly as

$$A\,[\vec{v}_1\ \vec{v}_2\ \cdots\ \vec{v}_n] = [\vec{v}_1\ \vec{v}_2\ \cdots\ \vec{v}_n] \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},$$

which turns into

$$[A\vec{v}_1\ A\vec{v}_2\ \cdots\ A\vec{v}_n] = [a_1\vec{v}_1\ a_2\vec{v}_2\ \cdots\ a_n\vec{v}_n],$$

or $A\vec{v}_1 = a_1\vec{v}_1,\ \ldots,\ A\vec{v}_n = a_n\vec{v}_n$. These are nothing but eigenvalue equations for the matrix A. This implies that the diagonal elements of D are the eigenvalues of A, and the columns of S are corresponding eigenvectors. Moreover, since S is invertible, eigenvectors of A form a *basis* of $\mathbb{R}^n$ (or $\mathbb{C}^n$). Conversely, if the eigenvalues of A form a basis, we can form the diagonal matrix D by placing the eigenvalues of A on the diagonal of D, and the matrix S by placing the corresponding eigenvectors as the columns of S *in the same order*. Thus, we have

Theorem 4.2.2. *A square matrix A is diagonalizable if and only if there exists a basis consisting of eigenvectors of A. Such a basis is called an* **eigenbasis**.

In other words, to diagonalize A, we must find n linearly independent eigenvectors A. When is this possible? It can be shown that eigenvectors corresponding to distinct eigenvalues are linearly independent. Thus, the number of linearly independent eigenvectors is the sum of the geometric multiplicities. But since $m_g(\lambda_i) \leq m_a(\lambda_i)$ for each eigenvalue λ_i, and $m_a(\lambda_1) + \cdots + m_a(\lambda_r) = n$, we have

Corollary 4.2.3. *A square matrix A is diagonalizable if and only the geometric and algebraic multiplicities of its eigenvalues agree: $m_g(\lambda) = m_a(\lambda)$ for every eigenvalue λ of A.*

Moreover, since $m_g(\lambda) \geq 1$ always, we have

Corollary 4.2.4. *If the algebraic multiplicity of every eigenvalue of A is one, then A is diagonalizable.*

Example 4.2.5. Let $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. We found that $\lambda = 2$ is an eigenvalue of A with algebraic multiplicity 2 but geometric multiplicity 1. Thus, A is *not* diagonalizable.

Example 4.2.6. Let $A = \begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix}$. Let's try to diagonalize it. We first find the characteristic polynomial:

$$p_A(x) = (-2-x)(1-x) - 2(-1) = x^2 + x = x(x+1).$$

Thus, the eigenvalues are 0 and -1, each with algebraic multiplicity 1, so A is diagonalizable. The vectors $(2,1)^T$ and $(1,1)^T$ are eigenvectors of A with eigenvalues -1 and 0 respectively, and they form an eigenbasis. Thus,

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

5 Lecture 5

5.1 Jordan normal form

As we saw in the previous lecture, a matrix A is diagonalizable if and only if it admits an eigenbasis if and only if the algebraic and geometric multiplicities of its eigenvalues agree. We also saw that not all matrices are diagonalizable. (Nondiagonalizable matrices are sometimes called **defective**.) How close to a diagonal matrix can we get?

The **Jordan block** of size m with eigenvalue λ is the matrix

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda \end{pmatrix},$$

i.e. it has λ's on the main diagonal and 1's on the superdiagonal.

Example 5.1.1. The following are examples of Jordan blocks:

$$J_3(7) = \begin{pmatrix} 7 & 1 & 0 \\ 0 & 7 & 1 \\ 0 & 0 & 7 \end{pmatrix}, \quad J_2(\pi) = \begin{pmatrix} \pi & 1 \\ 0 & \pi \end{pmatrix}, \quad J_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_1(13) = (13).$$

A matrix B is in **Jordan normal form** or is a **Jordan matrix** if it is of the form

$$B = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_l}(\lambda_l) \end{pmatrix}.$$

Example 5.1.2. The following matrices are in Jordan normal form:

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(What are the λ_i and m_i in each case?)

Theorem 5.1.3. *Every square matrix is similar to a matrix in Jordan normal form. The Jordan normal form of a matrix is unique up to reordering of the Jordan blocks.*

5.2 Generalized eigenvectors

Let A be a square matrix. To diagonalize A, we would need to find a basis of eigenvectors of A. Such a basis does not always exists, but something close to it does. If λ is an eigenvalue of A, then a **generalized eigenvector** of rank m corresponding to λ is a vector $\vec{v}$ such that $(A-\lambda I)^m\vec{v} = 0$ but $(A-\lambda I)^{m-1}\vec{v} \neq 0$. We see that a generalized eigenvector of rank 1 is just an ordinary eigenvector.

Example 5.2.1. Let $A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. The vector $\vec{v} = (1,0)^T$ is an honest eigenvector of A with eigenvalue $\lambda = 1$, since $A\vec{v} = \vec{v}$. The vector $\vec{w} = (2,1)^T$ is a generalized eigenvector of rank 2 corresponding to λ, since

$$(A-\lambda I)\vec{w} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad (A-\lambda I)^2\vec{w} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In fact $(A-\lambda I)^2 = 0$, so *every* vector is a generalized eigenvector.

A **Jordan chain** of generalized eigenvectors is a sequence of vectors $\vec{v}_1, \ldots, \vec{v}_m$ such that $\vec{v}_1$ is an eigenvector of A with eigenvalue λ, and $(A-\lambda I)\vec{v}_i = \vec{v}_{i-1}$ for $i = 2, \ldots, m$. Equivalently, $A\vec{v}_i = \lambda\vec{v}_i + \vec{v}_{i-1}$.

Example 5.2.2. Let $A = \frac{1}{2}\begin{pmatrix} 3 & 1 & -1 \\ -2 & 6 & 0 \\ 1 & -1 & 3 \end{pmatrix}$. The vectors $\vec{v_1} = (1,1,0)^T, \vec{v_2} = (0,1,-1)^T, \vec{v_3} = (1,2,1)^T$ form a Jordan chain of length 3 corresponding to eigenvalue $\lambda = 2$, since

$$A\vec{v}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = 2\vec{v}_1, \quad A\vec{v}_2 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = \vec{v}_1 + 2\vec{v}_2, \quad A\vec{v}_3 = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} = \vec{v}_2 + 2\vec{v}_3.$$

The key to finding the Jordan normal form of a matrix is the following

Theorem 5.2.3. *Every square matrix A admits a basis of generalized eigenvectors, and the basis can be chosen to consist entirely of Jordan chains.*

Recall that the dimension of the eigenspace $E_\lambda(A)$ is the geometric multiplicity of λ. It turns out that the dimension of the space of generalized eigenvectors corresponding to λ is *always* the algebraic multiplicity. In this way the generalized eigenvectors "fix" the possible lack of genuine eigenvectors.

5.3 Finding the Jordan normal form

Let A be a square matrix, and let $\vec{v}_1, \ldots, \vec{v}_m$ be a Jordan chain of generalized eigenvectors corresponding to eigenvalue λ. Form a matrix T by placing the vectors $\vec{v}_i$ as its columns:

$$T = [\vec{v}_1 \; \vec{v}_2 \; \cdots \; \vec{v}_m].$$

What happens when we multiply T by A from the left?

$$\begin{aligned} AT &= A\,[\vec{v}_1 \; \vec{v}_2 \; \cdots \; \vec{v}_m] = [A\vec{v}_1 \; A\vec{v}_2 \; \cdots \; A\vec{v}_m] = [\lambda\vec{v}_1 \; \lambda\vec{v}_2 + \vec{v}_1 \; \cdots \; \lambda\vec{v}_m + \vec{v}_{m-1}] \\ &= [\vec{v}_1 \; \vec{v}_2 \; \cdots \; \vec{v}_m] \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} = TJ_m(\lambda). \end{aligned}$$

Now choose a basis of generalized eigenvectors of A consisting of Jordan chains, so that each chain is grouped together starting with an eigenvector, and form the square matrix S by placing the chosen basis as the columns of S *in the given order*. The above argument generalizes to show that $AS = SJ$, where

$$J = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_l}(\lambda_l) \end{pmatrix}$$

is in Jordan normal form. Since the columns of S form a basis, S is invertible, so multiplying by S^{-1} gives $A = SJS^{-1}$. Thus, J is the Jordan normal form of A.

We summarize the recipe to find the Jordan normal form of an $n \times n$ matrix A as follows:

1. Find the eigenvalues of A.
2. For each eigenvalue λ, find a basis of the eigenspace $E_\lambda(A)$.
3. For each vector $\vec{v}$ with eigenvalue λ found in the previous step, form a Jordan chain as follows. Let $\vec{v}_1 = \vec{v}$, and solve $(A - \lambda I)\vec{x} = \vec{v}_1$ to find $\vec{v}_2$. Then solve $(A - \lambda I)\vec{x} = \vec{v}_2$ to find $\vec{v}_3$, and so on.
4. The previous step results in a number of vectors equal to the algebraic multiplicity of λ for each λ, altogether n linearly independent vectors. Let these vectors be the columns of a matrix S, where each Jordan chain is grouped together starting with the eigenvector. Now $A = SJS^{-1}$, where J is in Jordan normal form.

By the theorem above, this recipe will always work. Notice that for each eigenvalue λ, the number of Jordan chains corresponding to λ equals the geometric *multiplicity* $m_g(\lambda)$, and the number of vectors in all these chains equals the *algebraic* multiplicity $m_a(\lambda)$.

6 Lecture 6

6.1 Examples of finding Jordan normal form

Example 6.1.1. In the previous example we found that the matrix $A = \frac{1}{2}\begin{pmatrix} 3 & 1 & -1 \\ -2 & 6 & 0 \\ 1 & -1 & 3 \end{pmatrix}$ has Jordan chain $\vec{v_1} = (1,1,0)^T, \vec{v_2} = (0,1,-1)^T, \vec{v_3} = (1,2,1)^T$ corresponding to eigenvalue $\lambda = 2$. This implies that $m_a(2) = 3$ and $m_g(\lambda) = 1$. Thus, we have the decomposition $A = SJS^{-1}$, where

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & -1 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad S^{-1} = \frac{1}{2}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

The matrix J is the Jordan normal form of A.

Example 6.1.2. Let's find the Jordan decomposition of $A = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ 1 & -1 & 1 \end{pmatrix}$. The characteristic polynomial is

$$p_A(x) = (2-x)\left((3-x)(1-x)+1\right) = (2-x)^3,$$

so A has a single eigenvalue $\lambda = 2$ with algebraic multiplicity 3. To find a basis for the eigenspace $E_2(A)$, we solve the linear system $(A - 2I)\vec{x} = 0$, which in this case reads

$$\begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This simplifies to the single equation $x_1 - x_2 - x_3 = 0$. Thus, the nullspace is 2-dimensional, so the geometric multiplicity of $\lambda = 2$ is 2, which tells us that A is not diagonalizable. A basis for the eigenspace is given for example by $\vec{v}_1 = (1,0,1)^T, \vec{v}_2 = (0,-1,1)^T$. We are one vector short of a basis of generalized eigenvectors, so we must find a nonzero solution to either $(A-2I)\vec{x} = \vec{v}_1$ or $(A-2I)\vec{x} = \vec{v}_2$. The first equation won't give nonzero solutions as you can check, but the second one has for example the solution $\vec{v}_3 = (0,0,-1)^T$. Thus, the Jordan decomposition for A is

$$\begin{aligned}
\begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ 1 & -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & -1 & -1 \end{pmatrix},
\end{aligned}$$

and the Jordan normal form of A is $J = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

6.2 Matrix valued functions

A **matrix function** or a matrix-valued function is a rule that assigns to each number on an interval (α, β) a matrix of a fixed size $n \times m$. A special case is a **vector function**. We write such functions as

$$A(t) = \begin{pmatrix} a_{11}(t) & \cdots & a_{1m}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nm}(t) \end{pmatrix}, \quad \vec{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

The function $a_{ij}(t)$ are called the **component functions** of $A(t)$. We can view such a matrix function as a map $\mathbb{R} \to \mathbb{R}^{nm}$. We say that $A(t)$ is **continuous** if all the component functions are continuous in the usual sense, and similarly $A(t)$ is **differentiable** if all the components are differentiable. We define the **derivative** and the **definite integral** of $A(t)$ component wise:

$$\frac{dA}{dt} = \left(\frac{da_{ij}}{dt} \right), \quad \int_a^b A(t)dt = \left(\int_a^b a_{ij}(t)dt \right).$$

Example 6.2.1. If $A(t) = \begin{pmatrix} \sin t & t \\ 1 & \cos t \end{pmatrix}$, then

$$A'(t) = \begin{pmatrix} \cos t & 1 \\ 0 & -\sin t \end{pmatrix}, \quad \int_0^\pi A(t)dt = \begin{pmatrix} 2 & \frac{\pi^2}{2} \\ \pi & 0 \end{pmatrix}$$

Both derivative and integral are *linear*:

$$\frac{d}{dt}(c_1A(t)+c_2B(t)) = c_1\frac{d}{dt}A(t)+c_2\frac{d}{dt}B(t), \quad \int_a^b (c_1A(t)+c_2B(t))dt = c_1\int_a^b A(t)dt+c_2\int_a^b B(t)dt.$$

Moreover, the derivative satisfies a generalized **product rule**

$$\frac{d}{dt}(A(t)B(t)) = A(t)\frac{dB(t)}{dt} + \frac{dA(t)}{dt}B(t),$$

which as a special case gives $\frac{d}{dt}(CA) = C\frac{d}{dt}A$ if C is a constant matrix.

7 Lecture 7

7.1 Systems of first order ordinary linear equations

Recall from the first lecture that a **system of first order ordinary linear equations** is a collection of equations of the form

$$\begin{aligned} x_1'(t) &= a_{11}(t)x_1(t) + \ldots + a_{1n}(t)x_n(t) + b_1(t) \\ &\vdots \\ x_n'(t) &= a_{n1}(t)x_1(t) + \ldots + a_{nn}(t)x_n(t) + b_n(t), \end{aligned}$$

where $a_{11}(t), \ldots, a_{nn}(t), b_1(t), \ldots, b_n(t)$ are given functions. We write this in matrix form as

$$\begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{11}(t) & \cdots & a_{1n}(t) \end{pmatrix} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} + \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix},$$

or $\vec{x}'(t) = A(t)\vec{x}(t) + \vec{b}(t)$, or simply $\vec{x}' = A\vec{x} + \vec{b}$. A **solution** of the system on some interval $\alpha < t < \beta$ is any vector function $\vec{x}(t) = (x_1(t), \ldots, x_n(t))^T$ that is defined and continuously differentiable on $\alpha < t < \beta$ and satisfies the equation. An **initial value problem** is a linear system of ODE's $\vec{x}' = A\vec{x} + \vec{b}$ together with **initial data** $\vec{x}(t_0) = \vec{x}_0$ which the solution is required to satisfy, where $\alpha < t_0 < \beta$ and $\vec{x}_0$ are given.

Theorem 7.1.1 (Existence and uniqueness)**.** *If the functions $A(t)$ and $\vec{b}(t)$ are continuous on the interval $\alpha < t < \beta$, then there exists a unique solution to the initial value problem $\vec{x}'(t) = A(t)\vec{x}(t) + \vec{b}(t), \vec{x}(t_0) = \vec{x}_0$ for any $\alpha < t_0 < \beta$ and any vector $\vec{x}_0$, and this solution exists on the whole interval $\alpha < t < \beta$.*

7.2 The homogeneous equation

As with single homogeneous linear ODE's, we have the principle of superposition.

Theorem 7.2.1 (Principle of superposition)**.** *If $\vec{x}_1(t), \ldots, \vec{x}_r(t)$ are solutions to the homogeneous system $\vec{x}'(t) = A(t)\vec{x}(t)$, then for any scalars $c_1, \ldots, c_r$, the linear combination $c_1\vec{x}_1(t) + \ldots + c_r\vec{x}_r(t)$ is also a solution.*

This follows easily from the fact that both matrix multiplication and differentiation are linear operations. Another way of phrasing the principle of superposition is that the set of solutions of the homogeneous equation form a vector space. By the existence and uniqueness theorem, any solution is characterized by its value at a fixed time t_0, and this value can be any vector in $\mathbb{R}^n$. This implies

Corollary 7.2.2. *The space of solutions of the homogeneous equation is n-dimensional.*

A basis $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ for this solution space is called a **fundamental set of solutions**, and the collection of all linear combinations $c_1\vec{x}_1(t) + \ldots + c_n\vec{x}_n(t)$ the **general solution**.

Given n solutions $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ of $A\vec{x} = 0$, when are they a fundamental set of solutions? We can form the matrix

$$X(t) = [\vec{x}_1(t) \;\ldots\; \vec{x}_n(t)]$$

whose columns are the given solutions. Its determinant is called the **Wronskian** and denoted $W[\vec{x}_1, \ldots, \vec{x}_n]$. Note that the Wronskian is a function of t.

Theorem 7.2.3. *If $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ are solutions of $A(t)\vec{x}(t) = 0$, the Wronskian $W[\vec{x}_1, \ldots, \vec{x}_n]$ is either never zero, or identically zero. If it is nonzero, then the solutions $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ are a fundamental set of solutions.*

Proof. Assume that

$$W[\vec{x}_1, \ldots, \vec{x}_n](t_0) = \det[\vec{x}_1(t_0), \ldots, \vec{x}_n(t_0)] = 0$$

for some value $t_0 \in (\alpha, \beta)$. This means that the vectors $\vec{x}_1(t_0), \ldots, \vec{x}_n(t_0)$ are linearly dependent, so we can express one of them as a linear combination of the others, so by reordering the solutions, we may assume

$$\vec{x}_1(t_0) = c_2\vec{x}_2(t_0) + \cdots + c_n\vec{x}_n(t_0).$$

Define a new vector function

$$\vec{y}(t) = c_2\vec{x}_2(t) + \cdots + c_n\vec{x}_n(t).$$

Since $\vec{y}(t)$ is a linear combination of solutions of the equation $\vec{x}' = A\vec{x}$, it is a solution by the Superposition Principle. Moreover, by our assumption, the two solutions $\vec{x}_1(t)$ and $\vec{y}(t)$ agree when $t = t_0$. By the uniqueness part of the Existence and Uniqueness theorem, they must be the same solution, that is, $\vec{x}_1(t) = \vec{y}(t)$ for all $t \in (\alpha, \beta)$. In other words,

$$\vec{x}_1(t) = c_2\vec{x}_2(t) + \cdots + c_n\vec{x}_n(t),$$

so the vectors $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ are linearly dependent for all t. Thus, $\det[\vec{x}_1(t), \ldots, \vec{x}_n(t)] = 0$ for all t, that is, the Wronskian is identically zero. We have proved that the Wronskian is either always nonzero or identically zero.

Assume now that $W[\vec{x}_1, \ldots, \vec{x}_n]$ is nonzero. Let

$$c_1\vec{x}_1 + \cdots + c_n\vec{x}_n = 0$$

be an equation of linear independence. Let $t_0 \in (\alpha, \beta)$ be any number. Evaluating at t_0 gives

$$c_1\vec{x}_1(t_0) + \cdots + c_n\vec{x}_n(t_0) = 0$$

By assumption we have

$$W[\vec{x}_1, \ldots, \vec{x}_n](t_0) = \det[\vec{x}_1(t_0), \ldots, \vec{x}_n(t_0)] \neq 0,$$

so the vectors $\vec{x}_1(t_0), \ldots, \vec{x}_n(t_0)$ are linearly independent. Thus, $c_1 = \ldots = c_n = 0$, hence the solutions $\vec{x}_1, \ldots, \vec{x}_n$ are linearly independent. Since the solution space is n-dimensional, these solutions must form a basis, which we also call a fundamental set of solutions. $\square$

Example 7.2.4. Consider the system of equations

$$\begin{cases} x_1' &= x_1 - 2e^{-3t}x_2 \\ x_2' &= 2x_2 \end{cases} \quad \Leftrightarrow \quad \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & -2e^{-3t} \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The coefficient matrix $A = \begin{pmatrix} 1 & -2e^{-3t} \\ 0 & 2 \end{pmatrix}$ is continuous for $-\infty < t < \infty$, so for any initial values, the solution exists for all $-\infty < t < \infty$. Some solutions are

$$\vec{x}_1(t) = \begin{pmatrix} e^t \\ 0 \end{pmatrix}, \quad \vec{x}_2(t) = \begin{pmatrix} e^{-t} \\ e^{2t} \end{pmatrix}.$$

The Wronskian of these solutions is

$$W[\vec{x}_1(t), \vec{x}_2(t)] = \det \begin{pmatrix} e^t & e^{-t} \\ 0 & e^{2t} \end{pmatrix} = e^{3t},$$

which is never zero. Thus, these two solutions are a fundamental set of solutions, and the general solution is

$$\vec{x}(t) = c_1 \begin{pmatrix} e^t \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} e^{-t} \\ e^{2t} \end{pmatrix}.$$

8 Lecture 8

8.1 System of linear ODE's with constant coefficients

Consider the homogeneous equation $\vec{x}'(t) = A\vec{x}(t)$, where A is a *constant* matrix. Our goal is to find the general solution of for this system. The equation is analogous to the single ODE

$$y^{(n)} = a_0 y + a_1 y' + \cdots + a_{n-1} y^{(n-1)}.$$

Recall how this was solved: we guess that the solution has the form $y(t) = e^{rt}$, substitute this function into the equation, and obtain a polynomial equation for the parameter r.

Let's attempt something similar. Assume that $\vec{v}e^{rt}$ is a solution to the equation $\vec{x}' = A\vec{x}$, where $\vec{v}$ is a nonzero constant vector. Substituting this function into the equation gives

$$r\vec{v}e^{rt} = A\vec{v}e^{rt}.$$

Since e^{rt} is never zero, we can divide by it and obtain $r\vec{v} = A\vec{v}$, and rewrite this as $(A - rI)\vec{v} = 0$. This is nothing but an eigenvalue equation for A. Thus, we see that functions of the form $\vec{v}e^{\lambda t}$ are solutions of $\vec{x}' = A\vec{x}$, where $\vec{v}$ is an eigenvector of A and λ is the corresponding eigenvalue.

Example 8.1.1. Consider the system

$$\vec{x}' = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \vec{x}.$$

Substituting the function $(x_1, x_2)^T e^{rt}$ into the equation results in the eigenvalue equation

$$\begin{pmatrix} 1-r & 1 \\ 4 & 1-r \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

which has nonzero solutions if and only if $(1-r)^2 - 4 = 0$ if and only if $r = 3, -1$. An eigenvector with eigenvalue 3 is $(1, 2)^T$, and an eigenvector with eigenvalue -1 is $(1, -2)^T$. Thus, we obtain solutions

$$\vec{x}_1(t) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{3t}, \quad \vec{x}_2(t) = \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{-t}.$$

The Wronskian of these solutions is

$$W[\vec{x}_1, \vec{x}_2] = \det \begin{pmatrix} e^{3t} & e^{-t} \\ 2e^{3t} & -2e^{-t} \end{pmatrix} = -4e^{2t},$$

which is never zero. Thus, these solutions form a fundamental system of solutions, and the general solution is

$$\vec{x}(t) = c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{3t} + c_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{-t}.$$

8.2 Fundamental matrix

Assume that $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ are a fundamental set of solutions to the homogeneous equation $\vec{x}'(t) = A(t)\vec{x}(t)$ on some interval (α, β). The matrix that has these solutions as its columns,

$$\Psi(t) = [\vec{x}_1(t), \ldots, \vec{x}_n(t)],$$

is called a **fundamental matrix** of the equation. Note that since there are many fundamental sets of solutions, there are also many fundamental matrices, but by the existence and uniqueness theorem, any fundamental matrix is uniquely determined by its value $\Psi(t_0)$ at any fixed $t_0 \in (\alpha, \beta)$, and this value can be any invertible matrix.

Since matrix differentiation is componentwise, and the columns of $\Psi(t)$ are solutions, we have

$$\frac{d}{dt}\Psi(t) = [\vec{x}_1'(t), \ldots, \vec{x}_n'(t)] = [A(t)\vec{x}_1(t), \ldots, A(t)\vec{x}_n(t)] = A(t)[\vec{x}_1(t), \ldots, \vec{x}_n(t)] = A(t)\Psi(t).$$

Thus, the fundamental matrix satisfies the equation $\frac{d}{dt}\Psi(t) = A(t)\Psi(t)$, and conversely any matrix function whose columns are linearly independent and that satisfies this equation is a fundamental matrix, and its columns form a fundamental set of solutions. Also, since $\vec{x}_1(t), \ldots, \vec{x}_n(t)$ are a fundamental solution, the matrix $\Psi(t)$ is invertible for all t.

The fundamental matrix $\Psi(t)$ itself is *not a solution* of the differential equation, for the simple reason that a solution is a vector function, not a matrix function. Rather, $\Psi(t)$ encodes the data of all solutions in a single matrix function as follows. Recall that any solution can be written as a linear combination

$$\vec{x}(t) = c_1\vec{x}(t) + \cdots + c_n\vec{x}(t).$$

This linear combination can be viewed as the matrix product

$$\vec{x}(t) = [\vec{x}_1(t), \ldots, \vec{x}_n(t)] \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \Psi(t) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \Psi(t)\vec{c}.$$

Thus, finding the general solution of the equation is essentially the same thing as finding a fundamental matrix.

To solve an initial value problem $\vec{x}' = A\vec{x}, \vec{x}(t_0) = \vec{x}_0$, we notice that if $\Psi(t_0)\vec{c} = \vec{x}_0$, then since $\Psi(t_0)$ is invertible, we must have $\vec{c} = \Psi(t_0)^{-1}\vec{x}_0$. Thus, the unique solution to the initial value problem is

$$\vec{x}(t) = \Psi(t)\vec{c} = \Psi(t)\Psi(t_0)^{-1}\vec{x}_0.$$

A particularly nice fundamental matrix is the one such that $\Psi(t_0) = I$. We denote this fundamental matrix by $\Phi(t)$. Since now also $\Psi(t_0)^{-1} = I$, the above solution of the initial value problem simplifies to

$$\vec{x}(t) = \Phi(t)\vec{x}_0.$$

Example 8.2.1. For the system

$$\vec{x}' = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \vec{x}$$

from the previous example, one possible fundamental matrix is

$$\Psi(t) = \begin{pmatrix} e^{3t} & e^{-t} \\ 2e^{3t} & -2e^{-t} \end{pmatrix}.$$

If $t_0 = 0$, then this is not the fundamental matrix $\Phi(t)$, since $\Psi(0) = \left(\begin{smallmatrix} 1 & 1 \\ 2 & -2 \end{smallmatrix}\right)$. To find $\Phi(t)$, we would solve for the coefficients c_1, c_2 in the two systems of linear equations

$$c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The answer is a little complicated:

$$\Phi(t) = \begin{pmatrix} \frac{1}{2}\left(e^{3t} + e^{-t}\right) & \frac{1}{4}\left(e^{3t} - e^{-t}\right) \\ e^{3t} - e^{-t} & \frac{1}{2}\left(e^{3t} + e^{-t}\right) \end{pmatrix}.$$

9 Lecture 9

9.1 Complex solutions

Sometimes a system $\vec{x}' = A\vec{x}$ with A consisting of real valued functions may give rise to complex solutions. Even in that case we really care about real solutions. The following result allows us to find them.

Theorem 9.1.1. *If $\vec{x}(t) = \vec{u}(t) + i\vec{v}(t)$ is a solution to the homogeneous system $\vec{x}' = A\vec{x}$, where the components of A are real valued, then the real part $\vec{u}(t)$ and the imaginary part $\vec{v}(t)$ are also solutions.*

Proof. Write the equation as $\vec{x}' - A\vec{x} = 0$. Since A is real, we can split the equation into real and imaginary parts as follows:

$$\vec{u}' - A\vec{u} + i(\vec{v}' - A\vec{v}) = 0.$$

A complex number is zero if and only if its real and imaginary parts are zero, so this implies

$$\vec{u}' - A\vec{u} = 0, \quad \vec{v}' - A\vec{v} = 0.$$

Hence $\vec{u}$ and $\vec{v}$ are solutions. $\square$

Example 9.1.2. Consider the differential equation $\vec{x}\,' = A\vec{x}$, where $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. The characteristic polynomial of A is $p_A(x) = x^2 + 1$, so the eigenvalues of A are $\lambda_1 = i$ and $\lambda_2 = -i$. An eigenvector corresponding to $\lambda_1 = i$ is $\vec{v} = \begin{pmatrix} i \\ 1 \end{pmatrix}$, so a solution to the equation is

$$\vec{x}(t) = \begin{pmatrix} i \\ 1 \end{pmatrix} e^{it}.$$

Using Euler's formula, we find the real and imaginary parts of this solutions:

$$\vec{x}(t) = \begin{pmatrix} i \\ 1 \end{pmatrix} (\cos t + i \sin t) = \vec{x}(t) = \begin{pmatrix} i\cos t - \sin t \\ \cos t + i \sin t \end{pmatrix} = \vec{x}(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} + i\vec{x}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$

Thus, we conclude that

$$\operatorname{Re}\vec{x}(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \quad \text{and} \quad \operatorname{Im}\vec{x}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

are solutions as well. The Wronskian of these two solutions is

$$W[\vec{x}_1, \vec{x}_2] = \det \begin{pmatrix} -\sin t & \cos t \\ \cos t & \sin t \end{pmatrix} = -\sin^2 t - \cos^2 t = -1,$$

which is never zero, so these solutions for a fundamental set.

9.2 Matrix exponential

The method of forming solutions from the eigenvalues and corresponding eigenvectors of the coefficient matrix is useful when the coefficient matrix A admits an eigenbasis, since in that case we obtain a fundamental set of solutions. However, not every matrix admits an eigenbasis. Instead, we now describe a method of finding the fundamental matrix $\Phi(t)$ directly from the coefficient matrix A.

The "obvious" solution to $y' = ay$ is $y(t) = e^{at}$. This leads us to dream of a matrix function that looks like $\vec{x}(t) = e^{At}$, but what would it mean to raise e to a matrix power?

Recall that one way of defining the function e^x is using the Taylor series

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n,$$

which converges for all (complex) values of x. Now if A is a square matrix, then at least the series

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

makes sense formally, since A can be raised to any positive power. It turns out that this series converges for any A. We call it the **matrix exponential**.

It is not always true that $e^A e^B = e^{A+B}$, but this equation holds whenever A and B *commute*, that is, $AB = BA$. This follows since for commuting matrices A, B it holds that

$$(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k},$$

so using the Euler summation formula,

$$\begin{aligned} e^A e^B &= \left(\sum_{n=0}^{\infty} \frac{1}{n!} A^n\right)\left(\sum_{n=0}^{\infty} \frac{1}{n!} B^n\right) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{1}{k!}\frac{1}{(n-k)!} A^k B^{n-k} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = \sum_{n=0}^{\infty} \frac{1}{n!} (A+B)^n = e^{A+B}. \end{aligned}$$

Now consider the differential equation $\vec{x}' = A\vec{x}$, where A is a constant matrix. We claim that the function

$$\Phi(t) = e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$$

is a fundamental matrix with $\Phi(0) = I$. The second part of the claim is clear, since

$$\Phi(0) = \sum_{n=0}^{\infty} \frac{0^n}{n!} A^n = A^0 = I.$$

For the first part, we differentiate term by term:

$$\frac{d}{dt}\Phi(t) = \frac{d}{dt}e^{tA} = \sum_{n=0}^{\infty} \frac{d}{dt}\left(\frac{t^n}{n!}A^n\right) = \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!}A^n$$
$$= \sum_{n=0}^{\infty} \frac{t^n}{n!}A^{n+1} = A\sum_{n=0}^{\infty} \frac{t^n}{n!}A^n = A\Phi(t).$$

This shows that the columns of $\Phi(t)$ are solutions of the differential equation. To check that they are a fundamental set of solutions, we must show that they are linearly independent. For this, we notice that tA commutes with $-tA$, so

$$e^{tA}e^{-tA} = e^{tA+(-tA)} = e^0 = I,$$

so the matrix e^{tA} is invertible for all t, hence its columns are linearly independent.

We have now reduced finding the solution of $\vec{x}' = A\vec{x}$ to the problem of computing the matrix exponential e^{tA}.

Example 9.2.1. Let $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$. We find that

$$A^2 = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}, \quad A^3 = \begin{pmatrix} a^3 & 0 \\ 0 & b^3 \end{pmatrix}, \quad \ldots, \quad A^n = \begin{pmatrix} a^n & 0 \\ 0 & b^n \end{pmatrix},$$

so

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!}A^n = \sum_{n=0}^{\infty} \frac{1}{n!}\begin{pmatrix} a^n & 0 \\ 0 & b^n \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{1}{n!}a^n & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{1}{n!}b^n \end{pmatrix} = \begin{pmatrix} e^a & 0 \\ 0 & e^b \end{pmatrix}.$$

Similarly,

$$e^{tA} = \begin{pmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{pmatrix},$$

so we deduce that the differential equation $\vec{x}' = A\vec{x}$ has general solution

$$\vec{x}(t) = c_1\begin{pmatrix} e^{at} \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 0 \\ e^{bt} \end{pmatrix}.$$

This agrees with our earlier method since A has eigenvalues a and b with eigenvectors $(1,0)^T$ and $(0,1)^T$ respectively.

Example 9.2.2. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. We find that

$$A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I, \quad A^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -A, \quad A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I,$$

and the powers of A cycle through these four matrices. Thus,

$$A^{2n} = \begin{pmatrix} (-1)^n & 0 \\ 0 & (-1)^n \end{pmatrix}, \quad A^{2n+1} = \begin{pmatrix} 0 & -(-1)^n \\ (-1)^n & 0 \end{pmatrix},$$

and

$$
\begin{aligned}
e^{tA} &= \sum_{n=0}^{\infty} \frac{1}{n!}(tA)^n = \sum_{n=0}^{\infty} \begin{pmatrix} \frac{(-1)^n}{n!} t^{2n} & 0 \\ 0 & \frac{(-1)^n}{n!} t^{2n} \end{pmatrix} + \sum_{n=0}^{\infty} \begin{pmatrix} 0 & \frac{-(-1)^n}{n!} t^{2n+1} \\ \frac{(-1)^n}{n!} t^{2n} & 0 \end{pmatrix} \\
&= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^{2n} & -\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^{2n+1} \\ \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^{2n+1} & \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^{2n} \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.
\end{aligned}
$$

We deduce that the general solution of $\vec{x}' = A\vec{x}$ is

$$
\vec{x}(t) = c_1 \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} + c_2 \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}.
$$

Example 9.2.3. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We find that

$$
A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \quad \dots, \quad A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix},
$$

so

$$
\begin{aligned}
e^{tA} &= \sum_{n=0}^{\infty} \frac{1}{n!}(tA)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} t^n & nt^n \\ 0 & t^n \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{1}{n!} t^n & \sum_{n=1}^{\infty} \frac{1}{(n-1)!} t^n \\ 0 & \sum_{n=0}^{\infty} \frac{1}{n!} t^n \end{pmatrix} \\
&= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{1}{n!} t^n & t \sum_{n=0}^{\infty} \frac{1}{n!} t^n \\ 0 & \sum_{n=0}^{\infty} \frac{1}{n!} t^n \end{pmatrix} = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix}.
\end{aligned}
$$

Thus, we see that the equation $\vec{x}' = A\vec{x}$ is

$$
\vec{x}(t) = c_1 \begin{pmatrix} e^t \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} te^t \\ e^t \end{pmatrix}
$$

10 Lecture 10

10.1 Finding the matrix exponential

The key to computing the matrix exponential is Jordan normal form. Let A be a square matrix, and let $A = SJS^{-1}$ be a Jordan decomposition for it (where the columns of S form chains of generalized eigenvectors of A and J is the Jordan normal form of A). Now

$$
A^n = (SJS^{-1})^n = SJS^{-1} \cdot SJS^{-1} \cdot \dots \cdot SJS^{-1} = SJ^nS^{-1},
$$

so

$$
e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = \sum_{n=0}^{\infty} \frac{1}{n!} SJ^nS^{-1} = S \left(\sum_{n=0}^{\infty} \frac{1}{n!} J^n \right) S^{-1} = Se^JS^{-1}.
$$

Thus, we need to understand e^J where J is in Jordan normal form.

Recall that a matrix in Jordan normal form looks like

$$J = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_l}(\lambda_l) \end{pmatrix}.$$

When raising J to a positive power, the different Jordan blocks do not interfere with each other, so we have

$$J^n = \begin{pmatrix} J_{m_1}(\lambda_1)^n & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2)^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{m_l}(\lambda_l)^n \end{pmatrix},$$

and consequently

$$e^J = \begin{pmatrix} e^{J_{m_1}(\lambda_1)} & 0 & \cdots & 0 \\ 0 & e^{J_{m_2}(\lambda_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{J_{m_l}(\lambda_l)}. \end{pmatrix}$$

Thus, we are reduced to understanding e^J for a single Jordan block

$$J = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

We can write

$$J = \Lambda + N = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} + \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

where Λ is a diagonal matrix with the same entry λ along the whole diagonal, and N has ones on the superdiagonal and zeros everywhere else. The matrix $\Lambda = \lambda I$ is a **scalar matrix**, a scalar multiple of the identity matrix. Scalar matrices commute with every other matrix, so in particular $\Lambda N = N\Lambda$. Thus, by the properties of matrix exponential, we have

$$e^J = e^{\Lambda + N} = e^{\Lambda} e^N.$$

The matrix Λ is easy to exponentiate:

$$\Lambda^n = \operatorname{diag}(\lambda^n, \ldots, \lambda^n), \quad \text{so} \quad e^{\Lambda} = \operatorname{diag}(e^{\lambda}, \ldots, e^{\lambda}).$$

In fact the exponential of a diagonal matrix D is always the diagonal matrix whose diagonal entries are the exponentials of the entries of D, in the same order.

Exponentiation of the matrix N is a little more complicated, but the key property here is that N is a **nilpotent** matrix, meaning that $N^k = 0$ for some positive k. In fact, if N is $m \times m$, then $N^m = 0$. More concretely,

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \ldots, \quad N^{m-1} = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad N^m = 0,$$

as you can verify. Thus,

$$\begin{aligned} e^{tN} &= \sum_{n=0}^{\infty} \frac{t^n}{n!} N^n = I + tN + t^2 N^2 + \ldots + t^{m-1} N^{m-1} \\ &= \begin{pmatrix} 1 & t & \frac{1}{2}t^2 & \cdots & \frac{1}{(m-1)!} t^{m-1} \\ 0 & 1 & t & \cdots & \frac{1}{(m-2)!} t^{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}. \end{aligned}$$

Now we simply put these ingredients together. If A is a square matrix, the we compute e^A using the following steps:

1. Find the Jordan decomposition $A = SJS^{-1}$. Note that we also need S^{-1}, not just S.
2. For each Jordan block J' in J, write $J' = \Lambda + N$, where Λ is a scalar matrix and N is a nilpotent matrix with ones along the superdiagonal and zeros elsewhere.
3. Compute e^Λ and e^N using the above formulas.
4. Compute $e^{J'} = e^\Lambda e^N$. Form e^J by placing the blocks $e^{J'}$ on the diagonal in the appropriate order.
5. Compute $e^A = Se^J S^{-1}$.

Example 10.1.1. In Lecture 6 we found the Jordan decomposition

$$\begin{aligned} A = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 1 \\ 1 & -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & -1 & -1 \end{pmatrix}. \end{aligned}$$

Using the above recipe, we find that

$$e^{tA} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & -1 & -1 \end{pmatrix}.$$

10.2 More on fundamental matrix

We saw that the differential equation $\vec{x}\,' = A\vec{x}$, where A is a constant matrix, has fundamental matrix

$$\Phi(t) = e^{tA} = Se^{tJ}S^{-1},$$

where $A = SJS^{-1}$ is the Jordan decomposition of A. This fundamental matrix has the property that $\Phi(0) = I$, which makes it very convenient from a theoretical point of view, and also for solving initial value problems. However, computing S^{-1} can be tedious, and the form of $\Phi(t)$ can be quite complicated.

Another convenient fundamental matrix is given by

$$\Psi(t) = \Phi(t)S = Se^{tJ}.$$

If $S = (s_{ij})$ and $\Phi(t) = [\vec{x}_1(t), \ldots, \vec{x}_n(t)]$, then the ith column of $\Psi(t)$ is

$$\vec{y}_1(t) = s_{1i}\vec{x}_1(t) + \ldots + s_{ni}\vec{x}_n(t),$$

which is a linear combination of solutions to the differential equation, hence also a solution, so the columns of $\Psi(t)$ are solutions. Moreover, $\Psi(0) = S$ is invertible, hence the Wronskian of the columns is nonzero at $t = 0$, hence for all t, so the columns form a fundamental set of solutions.

If A is diagonalizable, then the fundamental matrix $\Psi(t)$ described above is just the same fundamental matrix as which we would have found using the initial method for solving the system $\vec{x}\,' = A\vec{x}$: find a basis $\vec{v}_1, \ldots, \vec{v}_n$ consisting of eigenvectors of A, and form solutions

$$\vec{x}_1(t) = \vec{v}_1 e^{\lambda_1 t}, \ldots, \vec{x}_n(t) = \vec{v}_n e^{\lambda_n t},$$

where λ_i is the eigenvalue associated to $\vec{v}_i$.

The usefulness of this fundamental matrix $\Psi(t)$ comes when we study the geometry of the solutions. The solutions $\vec{x}(t) = \vec{v}e^{\lambda t}$ are the simplest solutions to the equation, and more complicated solutions are just linear combinations of solutions of this form (at least in the diagonalizable case).

10.3 Direction field and phase plane

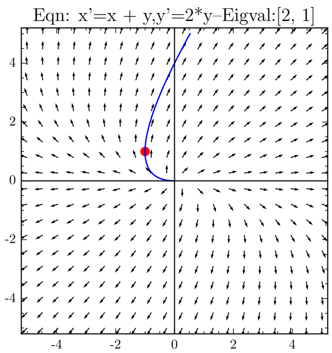
Now that we have learned methods for solving the differential equation $\vec{x}\,' = A\vec{x}$ with constant A, our next task is to understand the solutions geometrically and classify them. We will only do this in the 2-dimensional case.

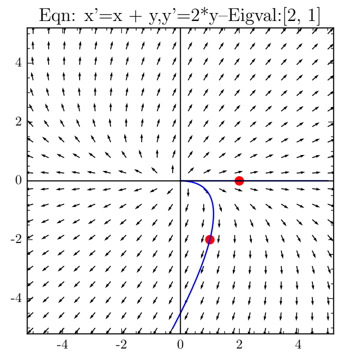
Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be a constant 2×2 matrix, and consider the differential equation $\vec{x}\,' = A\vec{x}$, where $\vec{x} = (x_1, x_2)^T$. The **phase plane** is simply the x_1x_2-plane. A **phase portrait** is a picture of the phase plane together with the trajectories of some representative solutions. To gain some geometric idea of what the solution look like without actually solving the system, we can draw a **direction field** as follows. For each point $(x_1, x_2)^T$ in the phase plane, draw a small arrow in the direction of

$$A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}.$$

We can view the solutions as parametric curves in the x_1x_2-plane, and the arrows in the direction field are the tangent directions of the solutions curves. We can think of the matrix A defining an ocean current, depicted by the direction field, and solutions as natural paths along which a small boat would float.

The result is called a direction field. The following direction field is associated to the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$.





10.4 Classifying solutions in 2D: two real eigenvalues

In the 2-dimensional case, we can classify and understand the solutions of $\vec{x}\,' = A\vec{x}$ based on the eigenvalues of the coefficient matrix A. Recall that the zero function $\vec{x}(t) = 0$ is a solution to the homogeneous equation. For this reason, the origin is called a **critical** or **equilibrium point** of the equation, since a solution starting there stays there for all time. We assume that the matrix A is nonsingular, which implies that the origin is the *only* equilibrium point.

Assume that A has two distinct real eigenvalues $\lambda_1 < \lambda_2$. Then A is diagonalizable since both eigenvalues must be simple. Thus, we find a fundamental set of solutions and the general solution by

$$x_1(t) = \vec{v}_1 e^{\lambda_1 t}, \quad x_2(t) = \vec{v}_2 e^{\lambda_2 t}, \qquad \vec{x}(t) = c_1\vec{v}_1 e^{\lambda_1 t} + c_2\vec{v}_2 e^{\lambda_2 t}.$$

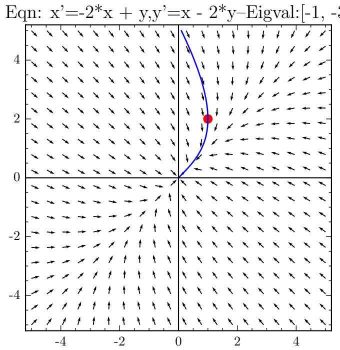
How can we understand the solutions? What do they look like? What happens when $t \to \infty$ or $t \to -\infty$? We divide the discussion into cases based on the signs of the eigenvalues.

Case 1: $0 < \lambda_1 < \lambda_2$
In this case we see that as $t \to \infty$, any nonzero solution grows unbounded, and as $t \to -\infty$, any solution tends to the origin. If $\vec{x}(0)$ lies on the line through origin with direction $\vec{v}_1$, then $c_2 = 0$, and the solution will stay on that line. Similarly if $c_1 = 0$, then $\vec{x}(t)$ stays on the line with direction $\vec{v}_2$.

If $c_1, c_2 \neq 0$, then since $e^{\lambda_2 t}$ grows faster than $e^{\lambda_1 t}$ as $t \to \infty$, the tangent direction of the solution approaches $\vec{v}_2$ as $t \to \infty$, and since $e^{\lambda_2 t}$ shrinks faster than $e^{\lambda_1 t}$ as as $t \to -\infty$, the solution becomes tangent to $\vec{v}_1$ as $t \to -\infty$. The equilibrium point at the origin is called a **node**, and is **unstable** since solutions become unbounded as $t \to \infty$.

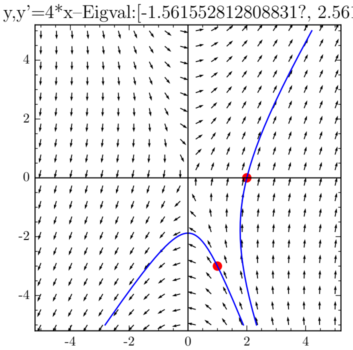
Case 2: $\lambda_1 < \lambda_2 < 0$
In this case we see that any solution tends to the origin as $t \to \infty$, and the solutions look exactly like in the previous case, except that the roles of λ_1 and λ_2 have changed, and the solutions flow in the opposite direction: as $t \to \infty$, they become tangent to $\vec{v}_2$, and as $t \to \infty$, the become tangent to $\vec{v}_1$. The equilibrium is a node, but this time it is **asymptotically stable**, since solutions tend exponentially toward it as $t \to \infty$.

Eqn: x’=-2*x + y,y’=x - 2*y–Eigval:[-1, -3]



Case 3: $\lambda_1 < 0 < \lambda_2$.
If $\vec{x}(0)$ lies on the line through origin with direction $\vec{v}_1$, then $c_2 = 0$, and the solution will stay on that line and tend to the origin as $t \to \infty$. If $\vec{x}(0)$ lies on the line through origin with direction $\vec{v}_2$, then $c_1 = 0$, and the solution will stay on that line and become unbounded as $t \to \infty$. Other solutions become unbounded and tangent to $\vec{v}_2$ since $e^{\lambda_1 t} \to 0$ and $e^{\lambda_2 t} \to \infty$ as $t \to \infty$. The equilibrium at the origin is called a **saddle point**, and it is **unstable** since most solutions grow exponentially as $t \to \infty$.

Eqn: x’=x + y,y’=4*x–Eigval:[-1.561552812808831?, 2.561552812808830?]



Example 10.4.1. The equilibrium of the system

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is an unstable node, since the eigenvalues of the coefficient matrix are $\lambda_1 = 1, \lambda_2 = 2$ with eigenvectors $\vec{v}_1 = (1,0)^T$ and $\vec{v}_2 = (1,1)^T$ respectively.

Example 10.4.2. The equilibrium of the system

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} -3 & -1 \\ -1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is an asymptotically stable node, since the eigenvalues of the coefficient matrix are $\lambda_1 = -4, \lambda_2 = -2$ with eigenvectors $\vec{v}_1 = (1,1)^T$ and $\vec{v}_2 = (1,-1)^T$ respectively.

Example 10.4.3. The equilibrium of the system

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is a saddle point, since the eigenvalues of the coefficient matrix are $\lambda_1 = 2, \lambda_2 = -1$ with eigenvectors $\vec{v}_1 = (1,1)^T$ and $\vec{v}_2 = (1,2)^T$ respectively.

11 Lecture 11

11.1 Complex solutions in 2D

We saw before that if a real matrix A has complex eigenvalues and eigenvectors, then we get real solutions to the system $\vec{x}' = A\vec{x}$ by taking real and imaginary parts of a complex solution. Let's now look more closely at the two-dimensional case. If A has complex eigenvalues, we get two real solutions by taking the real and imaginary parts of a single complex solution. We would like to know that these two solutions form a fundamental set of solutions. We use another lemma to prove this.

Lemma 11.1.1. *If A is a real matrix with complex eigenvalue λ and corresponding eigenvector $\vec{v}$, then also the complex conjugate $\bar{\lambda}$ is an eigenvalue of A with eigenvector $\overline{\vec{v}}$.*

Proof. Since $\vec{v}$ is an eigenvector of A with eigenvalue λ, we have $\lambda\vec{v} = A\vec{v}$. Taking complex conjugates and using the fact that $\bar{A} = A$ as A is real, we find

$$\bar{\lambda}\overline{\vec{v}} = \overline{\lambda\vec{v}} = \overline{A\vec{v}} = \bar{A}\overline{\vec{v}} = A\overline{\vec{v}}.$$

$\square$

Theorem 11.1.2. *Assume that A is a 2×2 real matrix with a complex eigenvalue $\lambda = a + ib$, with $b \neq 0$, and eigenvector $\vec{v}$ associated to λ. The real and imaginary parts of $\vec{x}(t) = \vec{v}e^{\lambda t}$ form a fundamental set of solutions to the equation $\vec{x}\,' = A\vec{x}$.*

Proof. By the previous lemma, since A is real, $\bar{\lambda} = a - ib$ is also an eigenvalue of A. Since $b \neq 0$, we have $\lambda \neq \bar{\lambda}$, and so $\vec{v}$ and $\overline{\vec{v}}$ correspond to distinct eigenvalues, so they are linearly independent. Define

$$\vec{w}_1 = \operatorname{Re} \vec{v} = \frac{1}{2}(\vec{v} + \overline{\vec{v}}), \quad \vec{w}_2 = \operatorname{Im} \vec{v} = \frac{1}{2i}(\vec{v} - \overline{\vec{v}}).$$

The vectors $\vec{w}_1$ and $\vec{w}_2$ are linearly independent, since if $c_1\vec{w}_1 + c_2\vec{w}_2 = 0$, then

$$\left(\frac{1}{2}c_1 + \frac{1}{2i}c_2\right)\vec{v} + \left(\frac{1}{2}c_1 - \frac{1}{2i}c_2\right)\overline{\vec{v}} = 0,$$

and since $\vec{v}$ and $\overline{\vec{v}}$ are linearly independent, this implies that

$$\frac{1}{2}c_1 + \frac{1}{2i}c_2 = 0, \quad \frac{1}{2}c_1 - \frac{1}{2i}c_2 = 0.$$

This pair of equations only has the trivial solution since the coefficient matrix has determinant $-\frac{1}{2i}$.

Now consider the real part $\vec{u}(t)$ and the imaginary part $\vec{v}(t)$ of the solution $\vec{x}(t) = \vec{v}e^{\lambda t}$. When $t = 0$, we have $\vec{x}(0) = \vec{v} = \vec{w}_1 + i\vec{w}_2$, hence $\vec{u}(0) = \vec{w}_1$ and $\vec{v}(0) = \vec{w}_2$. Thus, the Wronskian $W[\vec{u}, \vec{v}]$ is nonzero when $t = 0$, hence nonzero always, so $\vec{u}(t)$ and $\vec{v}(t)$ for a fundamental set of solutions. □

11.2 Classifying solutions in 2D: eigenvalues are complex conjugates

Consider the differential equation $\vec{x}\,' = A\vec{x}$, where A is a real 2×2 matrix. When A has two distinct real nonzero eigenvalues, we classified the equilibrium point at $\vec{x} = 0$ as either an unstable node, asymptotically stable node, or saddle point, depending on whether the eigenvalues of A are both positive, both negative, or one positive and one negative, respectively.

Assume now that A has two complex conjugate eigenvalues. Let one of them be $\lambda = a + ib$, with an associated eigenvector $\vec{v} = (\alpha_1 + i\beta_1, \alpha_2 + i\beta_2)^T$. From the previous section we know that the real and imaginary parts of $\vec{x}(t) = \vec{v}e^{\lambda t}$ form a fundamental set of solutions. Let's compute these two solutions:

$$\begin{aligned}\vec{x}(t) = \vec{v}e^{(a+ib)t} &= \begin{pmatrix} \alpha_1 + i\beta_1 \\ \alpha_2 + i\beta_2 \end{pmatrix} e^{at}(\cos bt + i \sin bt) \\ &= \begin{pmatrix} \alpha_1 \cos bt - \beta_1 \sin bt \\ \alpha_2 \cos bt - \beta_2 \sin bt \end{pmatrix} e^{at} + i \begin{pmatrix} \beta_1 \cos bt + \alpha_1 \sin bt \\ \beta_2 \cos bt + \alpha_2 \sin bt \end{pmatrix} e^{at},\end{aligned}$$

so

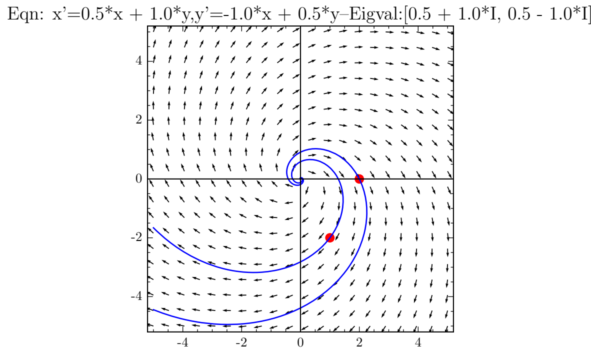
$$\vec{x}_1(t) = \begin{pmatrix} \alpha_1 \cos bt - \beta_1 \sin bt \\ \alpha_2 \cos bt - \beta_2 \sin bt \end{pmatrix} e^{at} \quad \text{and} \quad \vec{x}_2(t) = \begin{pmatrix} \beta_1 \cos bt + \alpha_1 \sin bt \\ \beta_2 \cos bt + \alpha_2 \sin bt \end{pmatrix} e^{at}.$$

The vectors

$$\begin{pmatrix} \alpha_1 \cos bt - \beta_1 \sin bt \\ \alpha_2 \cos bt - \beta_2 \sin bt \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \beta_1 \cos bt + \alpha_1 \sin bt \\ \beta_2 \cos bt + \alpha_2 \sin bt \end{pmatrix}$$

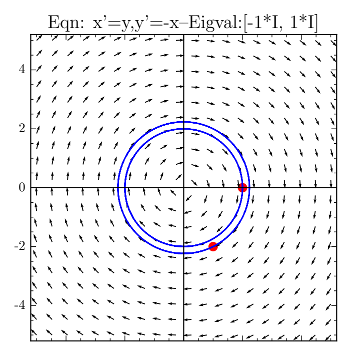
rotate around the origin in closed ellipses, so the long-term behavior of the solutions depends only on the real part of λ. Again, we have three cases.

Case 1: $\operatorname{Re}\lambda > 0$
In this case the solutions spiral outward from the origin. The equilibrium point at the origin is called an **unstable spiral point**.



Case 2: $\operatorname{Re}\lambda < 0$
Now the solutions spiral inward towards the origin. The phase portrait looks the same as in the previous case, except the solutions move in the opposite direction. The equilibrium point at the origin is called an **asymptotically stable spiral point**.

Case 3: $\operatorname{Re}\lambda = 0$.
In this case the solution do not spiral inward or outward, but rotate around the origin in ellipses. The equilibrium point at the origin is called a **center**. It is **stable** since solutions starting near the origin stay near the origin, but it *not asymptotically stable*, since solutions do not tend toward the origin.



11.3 Classifying solutions in 2D: only one eigenvalue

What if A has only one eigenvalue λ. The algebraic multiplicity of λ is necessarily 2, but the geometric multiplicity can be either 1 or 2.

Case 1: $m_g(\lambda) = 2$
In this case A has two linearly independent eigenvectors with eigenvalue λ, so the eigenspace must be the whole plane. Thus, the general solution is

$$\vec{x}(t) = (c_1\vec{e}_1 + c_2\vec{e}_2)e^{\lambda t}.$$

Each solution stays on a ray emanating from the origin. The equilibrium point at the origin is called a **proper node** or **star point**. If $\lambda > 0$, the solutions tend to infinity, and the equilibrium is unstable. If $\lambda < 0$, the solutions tend to the origin, and the equilibrium is asymptotically stable.

Case 2: $m_g(\lambda) = 1$
In this case, A has only one linearly independent eigenvector $\vec{v}$, and a generalized eigenvector $\vec{w}$ with the property that $A\vec{w} = \lambda\vec{w} + \vec{v}$. The general solution is

$$\vec{x}(t) = c_1\vec{v}e^{\lambda t} + c_2(\vec{w} + t\vec{v})e^{\lambda t}.$$

Since the $t\vec{v}e^{\lambda t}$ dominates when $t \to \pm\infty$, the solutions become tangent to $\vec{v}$ when $t \to \infty$ and $t \to -\infty$. The equilibrium point at the origin is called an **improper node**. If $\lambda > 0$, the equilibrium is unstable, and if $\lambda < 0$, it is asymptotically stable.

12 Lecture 12

12.1 Inhomogeneous equation

We have learned to solve the homogeneous equation $\vec{x}\,' = A\vec{x}$ in the case where A is a constant matrix, and studied the solutions in the two-dimensional case. Now we turn our attention to solving the inhomogeneous equation

$$\vec{x}\,' = A\vec{x} + \vec{b}.$$

There are several methods for solving the inhomogeneous equation, but we will discuss only one of them, namely variation of parameters. Although in our examples, the coefficient matrix A will be constant, we outline the method of variation of parameters for time-dependent A.

Assume that $A(t)$ and $\vec{b}(t)$ are continuous on some interval (α, β). From the existence and uniqueness theorem we know that for any initial value $\vec{x}(0) = \vec{x}_0$, there exists a unique solution to the equation $\vec{x}\,' = A(t)\vec{x} + \vec{b}(t)$ satisfying the initial condition. Our strategy for solving this initial value problem is to find the general solution to the inhomogeneous equation, and then match the undetermined coefficient so that the solution satisfies the initial condition. To find the general solution, we have a useful lemma, whose proof is a homework problem.

Lemma 12.1.1. *The general solution of the inhomogeneous equation* $\vec{x}\,' = A\vec{x} + \vec{b}$ *is the general solution of the homogeneous equation* $\vec{x}\,' = A\vec{x}$ *plus any particular solution of the inhomogeneous equation.*

In other words, we must first find the general solution of the homogeneous equation, and then find *any* solution of the inhomogeneous equation. We already know how to find the general solution of the homogeneous equation in the case of a constant matrix A, so let's concentrate on finding a single solution to the inhomogeneous one.

12.2 Variation of parameters

Recall that the fundamental matrix $\Psi(t)$ is a matrix whose columns form a fundamental set of solutions of the equation $\vec{x}\,' = A\vec{x}$, and the general solution is given by $\vec{x}(t) = \Psi(t)\vec{c}$, where $\vec{c}$ is an undetermined constant vector. The idea of variation of parameters is to let this vector $\vec{c}$ depend on t and seek a solution of this form. For clarity, we rename this vector as $\vec{u}(t)$. So we attempt to find a solution of the form

$$\vec{x}(t) = \Psi(t)\vec{u}(t).$$

We substitute this into the equation $\vec{x}\,' = A(t)\vec{x} + \vec{b}(t)$ to obtain

$$\Psi'(t)\vec{u}(t) + \Psi(t)\vec{u}\,'(t) = A(t)\Psi(t)\vec{u}(t) + \vec{b}(t).$$

Since $\Psi(t)$ is a fundamental matrix of the homogeneous equation, it satisfies $\Psi'(t) = A(t)\Psi(t)$. Thus, the left terms in the above equation cancel, and we have

$$\Psi(t)\vec{u}\,'(t) = \vec{b}(t).$$

Since $\Psi(t)$ is invertible for all t, we can write

$$\vec{u}\,'(t) = \Psi^{-1}(t)\vec{b}(t),$$

and (at least in principle) integrate to get

$$\vec{u}(t) = \int \Psi^{-1}(t)\vec{b}(t)dt.$$

The particular solution to the inhomogeneous is now

$$\vec{x}(t) = \Psi(t)\vec{u}(t) = \Psi(t)\int \Psi^{-1}(t)\vec{b}(t)dt,$$

and by the lemma above, the general solution to is

$$\vec{x}(t) = \Psi(t)\vec{c} + \Psi(t)\int \Psi^{-1}(t)\vec{b}(t)dt,$$

where $\vec{c}$ is an undetermined constant vector, whose components we must find to solve the initial value problem.

Although this form is easy to write down for theoretical purposes, in practice it is usually easiest to start with the equation $\Psi(t)\vec{u}\,'(t) = \vec{b}(t)$ and find by hand a function $\vec{u}(t)$ satisfying this equation, using row reduction.

12.3 Examples

Example 12.3.1. Consider the equation

$$\vec{x}\,'(t) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \vec{x}(t) + \begin{pmatrix} 2e^{-t} \\ 3t \end{pmatrix}.$$

We start by finding a fundamental solution for the homogeneous equation. The eigenvalues of the coefficient matrix A are $\lambda_1 = -3, \lambda_2 = -1$ (so the equilibrium point is an asymptotically stable node, by the way), with eigenvectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Thus, we get a fundamental set of solutions, fundamental matrix, and general solution as

$$\vec{x}_1(t) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-3t}, \quad \vec{x}_2(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t}.$$

$$\Psi(t) = \begin{pmatrix} e^{-3t} & e^{-t} \\ -e^{-3t} & e^{-t} \end{pmatrix}, \quad \vec{x}(t) = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-3t} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t}.$$

We must now solve the system of equations

$$\begin{pmatrix} e^{-3t} & e^{-t} \\ -e^{-3t} & e^{-t} \end{pmatrix} \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} 2e^{-t} \\ 3t \end{pmatrix}.$$

Row reduction gives

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} e^{2t} - \frac{3}{2}te^{3t} \\ 1 + \frac{3}{2}te^{t} \end{pmatrix},$$

so we get equations

$$u_1' = e^{2t} - \frac{3}{2}te^{3t}, \quad u_2' = 1 + \frac{3}{2}te^{t},$$

which by integration (by parts) yields

$$u_1 = \frac{1}{2}e^{2t} - \frac{1}{2}te^{3t} + \frac{1}{6}e^{3t}, \quad u_2 = t + \frac{3}{2}te^{t} - \frac{3}{2}e^{t}.$$

Thus, a particular solutions of the inhomogeneous equation is

$$\begin{aligned}
\Psi(t)\vec{u} &= \begin{pmatrix} e^{-3t} & e^{-t} \\ -e^{-3t} & e^{-t} \end{pmatrix} \begin{pmatrix} \frac{1}{2}e^{2t} - \frac{1}{2}te^{3t} + \frac{1}{6}e^{3t} \\ t + \frac{3}{2}te^{t} - \frac{3}{2}e^{t} \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ 1 \end{pmatrix} te^{t} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-t} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} t - \frac{1}{3} \begin{pmatrix} 4 \\ 5 \end{pmatrix},
\end{aligned}$$

and the general solution is

$$\vec{x}(t) = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-3t} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} te^{t} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-t} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} t - \frac{1}{3} \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

To find the solution that satisfies $\vec{x}(0) = (1,1)^T$, we set

$$\vec{x}(0) = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which gives the system of equations

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \frac{13}{6} \\ \frac{19}{6} \end{pmatrix},$$

which as solution $c_1 = -\frac{1}{2}, c_2 = \frac{8}{3}$. Thus, the particular solution with initial value $(1,1)^T$ is

$$\vec{x}(t) = -\frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-3t} + \frac{8}{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} te^{t} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-t} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} t - \frac{1}{3} \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

Example 12.3.2.

$$\vec{x}\,' = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix} \vec{x} + \begin{pmatrix} -\cos t \\ \sin t \end{pmatrix}.$$

13 Lecture 13

13.1 Review of first and second order linear ODE's

Consider the linear inhomogeneous ordinary differential equation

$$y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = f(x)$$

and the corresponding homogeneous equation

$$y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0,$$

where $a_0, \ldots, a_{n-1}$ are real numbers and $f(x)$ is a continuous function. As with the linear system of first order equations $\vec{x}\,' = A\vec{x} + \vec{b}$, there is an existence and uniqueness theorem for the inhomogeneous equation. In this case it guarantees that for any point x_0 and any initial values

$$y(x_0) = c_1, \quad y'(x_0) = c_2, \quad \ldots, \quad y^{(n-1)}(x_0) = c_n,$$

there exists a unique solution satisfying these initial condition. Also the principle of superposition holds for the homogeneous equation, which implies (together with the existence and uniqueness theorem) that the solutions of the homogeneous equation form an n-dimensional vector space, and also that the general solution of the inhomogeneous equation equals the general solution of the homogeneous equation plus any particular solution of the inhomogeneous one.

To solve the homogeneous equation, we attempt to find a solution in the form $y(x) = e^{rx}$. Substituting this into the homogeneous equation and dividing by e^{rx} (which is never zero) lead to the polynomial equation

$$r^n + a_{n-1} r^{n-1} + \cdots + a_1 r + a_0 = 0,$$

called the **characteristic equation**. For each root $r_1, \ldots, r_s$ of this polynomial, we get a solution $e^{r_i x}$. If a root is repeated m_i times, we get additional solutions $xe^{r_i x}, \ldots, x^{m_i - 1}e^{r_i x}$. If some of the roots appear in complex conjugate pairs, then we obtain real solutions by taking the real and imaginary parts of $x^k e^{r_i x}$. This results in n linearly independent solutions.

For us, the most important examples are of order two.

Example 13.1.1. The general solution of $y'' + y' - 6y = 0$ is $y(x) = C_1 e^{2x} + C_2 e^{-3x}$ since the characteristic polynomial is $r^2 + r - 6 = (r-2)(r+3)$.

Example 13.1.2. The general solution of $y'' + 2y' + y = 0$ is $y(x) = C_1 e^{-x} + C_2 x e^{-x}$ since the characteristic polynomial is $r^2 + 2r + 1 = (r+1)^2$.

Example 13.1.3. The equation $y'' + 4y' + 5y = 0$ has characteristic polynomial $r^2 + 4r + 5 = (r+2)^2 + 1$, which has roots $-2 \pm i$. Thus, a solution is $y(x) = e^{(-2+i)x} = e^{-2x}(\cos x + i \sin x)$. This has real part $y_1(x) = e^{-2x} \cos x$ and imaginary part $e^{-2x} \sin x$, so the general solution is $y(x) = C_1 e^{-2x} \cos x + C_2 e^{-2x} \sin x$.

To solve the inhomogeneous equation, we typically guess the general form of the particular solution based on the inhomogeneous part $f(x)$.

Example 13.1.4. Consider $y'' + y' - 6y = \cos x$. We attempt to find a particular solution in the form $y(x) = a \cos x + b \sin x$. Substituting this into the differential equation gives

$$(b - 7a) \cos x + (-b - 7a) \sin x = \cos x.$$

For this to hold for all x, we must have $b - 7a = 1$ and $-b - 7a = 0$, which has the solution $a = -\frac{7}{50}, b = \frac{1}{50}$. Thus, the general solution is

$$y(x) = C_1 e^{2x} + C_2 e^{-3x} - \frac{7}{50} \cos x + \frac{1}{50} \sin x.$$

Example 13.1.5. Consider $y'' + y' - 6y = e^{2x}$. This time the obvious try Ce^{2x} is already a solution to the homogeneous equation, so it cannot work. Instead, we try $y(x) = Cxe^{2x}$. Substituting this into the differential equation gives

$$5Ce^{2x} = e^{2x},$$

so we must have $C = \frac{1}{5}$. Thus, the general solution is

$$y(x) = C_1 e^{2x} + C_2 e^{-3x} + \frac{1}{5} x e^{2x}.$$

13.2 Two-point boundary value problem

A **two-point boundary value problem** is a linear second order ODE together with boundary data at two distinct points x_1 and x_2. The boundary data could be either the values of the function at x_1 and x_2, or the value of the derivative at these points, or the value of the function at one point and the derivative at the other. A typical two-point boundary value problem looks like

$$\begin{cases} y'' + p(x)y' + q(x)y = g(x) \\ y(x_1) = y_1,\ y(x_2) = y_2. \end{cases}$$

If the function $g(x)$ is identically zero and also $y_1 = y_2 = 0$, the problem is called **homogeneous**, otherwise it is **inhomogeneous**. The homogeneous equation always has the solution $y(x) = 0$, but it may have more solutions, even infinitely many. The inhomogeneous equation may have a single solution, or multiple solutions, or no solutions at all.

Example 13.2.1. Consider the homogeneous boundary value problem

$$y'' + 2y = 0, \quad y(0) = y(\pi) = 0.$$

The general solution of the differential equation is

$$y(x) = C_1 \cos(\sqrt{2}x) + C_2 \sin(\sqrt{2}x).$$

The first boundary condition implies that $C_1 = 0$, and the then second leads to $C_2 \sin(\sqrt{2}\pi) = 0$. Since $\sqrt{2}\pi$ is not an integer multiple of π, we must have $C_2 = 0$. Thus, the initial value problem has a unique solution.

Example 13.2.2. This time, consider the inhomogeneous boundary value problem

$$y'' + 2y = 0, \quad y(0) = 1, \quad y(\pi) = 0.$$

The first boundary condition implies that $C_1 = 1$, and then the second leads to

$$\cos(\sqrt{2}\pi) + C_2 \sin(\sqrt{2}\pi) = 0 \quad \Rightarrow \quad C_2 = -\cot(\sqrt{2}\pi).$$

Again, there is a unique solution.

14 Lecture 14

14.1 More examples of boundary value problems

Example 14.1.1. Consider the homogeneous boundary value problem

$$y'' + y = 0, \quad y(0) = y(\pi) = 0.$$

The general solution of the differential equation is

$$y(x) = C_1 \cos(x) + C_2 \sin(x).$$

The first boundary condition implies $C_1 = 0$, so we have $y(x) = C_2 \sin(x)$. This function satisfies both boundary conditions for any constant of C_2, so this time there are infinitely many solutions.

Example 14.1.2. Consider the inhomogeneous boundary value problem

$$y'' + y = 0, \quad y(0) = 1, \quad y(\pi) = a,$$

where a is a given real number. The first boundary condition implies that $C_1 = 1$, and the second one that $C_1 \cos(\pi) = -C_1 = a$. These two conditions are incompatible unless $a = -1$. Thus, if $a \neq -1$, the boundary value problem has no solutions. If $a = -1$ however, the function $y(x) = \cos(x) + C_2 \sin(x)$ is a solution for any C_2, so we get infinitely many solutions.

14.2 Eigenvalue problems

We saw that the homogeneous boundary value problem can have either only the trivial solution $y(x) = 0$, or also nontrivial solutions. These cases seemed to correspond to the inhomogeneous equation having a unique solution, or having either no solutions or many solutions. The situation is analogous to the linear equations

$$A\vec{x} = 0, \quad A\vec{x} = \vec{b},$$

with A an $n \times n$ matrix, where we know that if the homogeneous equation has only the trivial solution, then the inhomogeneous equation has a unique solution, or else if the homogeneous equation has many solutions, then the inhomogeneous equation either has no solutions or many solutions.

In fact, if A is a square matrix, we know that the equation

$$(A - \lambda I)\vec{x} = 0$$

has nontrivial solutions only for finitely many values of λ, and those values are called the eigenvalues of A. If we now look back at the examples of boundary value problems, we saw that

$$y'' + 2y = 0, \quad y(0) = y(\pi) = 0$$

has only the trivial solution, whereas

$$y'' + y = 0, \quad y(0) = y(\pi) = 0$$

has infinitely many solutions. Maybe there is something special in the coefficient 1 in front of y in the second equation.

Consider the equation

$$y'' + \lambda y = 0, \quad y(0) = y(\pi) = 0,$$

where λ is a real number. We say that the values of λ for which this equation has nontrivial solutions are called **eigenvalues**, and the corresponding nonzero solutions **eigenfunctions**. Let's find all the eigenvalues of this problem.

The simplest case is $\lambda = 0$, where the general solution is $y(x) = Ax + B$. The first boundary condition forces $B = 0$, and consequently the second forces $A = 0$. Thus, there are no nontrivial solutions, so $\lambda = 0$ is *not* an eigenvalue.

Next, assume that $\lambda < 0$. We write $\lambda = -\mu^2$ with $\mu > 0$ to simplify notation. The general solution is

$$y(x) = C_1 e^{\mu x} + C_2 e^{-\mu x}.$$

The two boundary conditions imply that

$$y(0) = C_1 + C_2 = 0, \quad y(\pi) = C_1 e^{\mu\pi} + C_2 e^{-\mu\pi} = 0.$$

This is a pair of linear equations for C_1 and C_2, and in matrix form reads

$$\begin{pmatrix} 1 & 1 \\ e^{\mu\pi} & e^{-\mu\pi} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = 0.$$

The coefficient matrix has determinant $e^{-\mu\pi} - e^{\mu\pi} = e^{\mu\pi}(e^{-2\mu\pi} - 1) \neq 0$, since $\mu \neq 0$. Thus, the only possibility is $C_1 = C_2 = 0$, so the only solution is the trivial solution $y(x) = 0$. Thus, there are no negative eigenvalues.

Finally, assume that $\lambda > 0$. We write $\lambda = \mu^2$ to simplify notation. The general solution in this case is

$$y(x) = C_1 \cos(\mu x) + C_2 \sin(\mu x).$$

The boundary condition $y(0) = 0$ implies $C_1 = 0$, so the possible nontrivial solutions are of the form $y(x) = C_2 \sin(\mu x)$. The second boundary condition implies now that $C_2 \sin(\mu\pi) = 0$. Since we are looking for nonzero solutions, we need $\sin(\mu\pi) = 0$. Since the zeros of sine are precisely the integer multiples of π, we must require that $\mu = n$ is an integer. But any integer works, so we see that the positive eigenvalues are

$$\lambda_1 = 1, \quad \lambda_2 = 4, \quad \lambda_3 = 9, \quad \ldots, \quad \lambda_n = n^2, \quad \ldots$$

with corresponding eigenfunctions

$$y_1(x) = \sin(x), \quad y_2(x) = \sin(2x), \quad y_3(x) = \sin(3x), \quad \ldots, \quad y_n(x) = \sin(nx), \quad \ldots$$

and any nonzero multiples of them.

15 Lecture 15

15.1 Periodic functions

In the homework you show that the boundary value problem

$$y'' + \lambda y = 0, \quad y(0) = y(L) = 0$$

has eigenvalues $\frac{n^2\pi^2}{L^2}$ with corresponding eigenfunctions $\sin \frac{n\pi x}{L}$. When we solve partial differential equations, we are going to construct more complicated functions from these basic functions, and also from the functions $\cos \frac{n\pi x}{L}$. Let's study some properties of them.

Recall that $\sin x$ and $\cos x$ are periodic functions of period 2π. A function $f(x)$ is **periodic** with period T if $f(x+T) = f(x)$ for all x. The graph of a periodic function looks like the same pattern repeating every T units. Note that if $f(x)$ is periodic with period T, then it is periodic with period $2T, 3T, \ldots, nT, \ldots$. The smallest number T such that $f(x)$ is periodic with period T is called the **fundamental period**. The functions $\sin \frac{n\pi x}{L}$ and $\cos \frac{n\pi x}{L}$ are periodic with fundamental period $\frac{2L}{n}$, and in particular the functions

$$\sin \frac{\pi x}{L},\ \cos \frac{\pi x}{L},\ \sin \frac{2\pi x}{L},\ \cos \frac{2\pi x}{L},\ \ldots,\ \sin \frac{n\pi x}{L},\ \cos \frac{n\pi x}{L},\ \ldots$$

are all periodic with period $2L$.

Consider the set of all T-periodic functions, which we denote by $\mathcal{P}_T$. This set is in fact a *vector space*, which we check as follows. First, the zero function (like any constant function) is T-periodic. Second, if f and g are T-periodic, and if a and b are real numbers, then

$$(af + bg)(x + T) = af(x + T) + bg(x + T) = af(x) + bg(x) = (af + bg)(x),$$

which means that also $af + bg$ is T-periodic.

Example 15.1.1. The functions

$$\sin \frac{2\pi x}{L} - \cos \frac{5\pi x}{L} \quad \text{and} \quad \frac{1}{100} \cos \frac{100\pi x}{L} + \frac{1}{1000} \cos \frac{1000\pi x}{L} + \frac{1}{10000} \cos \frac{10000\pi x}{L}$$

are $2L$-periodic since they are linear combinations of $2L$-periodic functions.

A stark difference between the familiar vector space $\mathbb{R}^n$ and this new space $\mathcal{P}_T$ is that $\mathbb{R}^n$ is finite-dimensional, whereas $\mathcal{P}_T$ is infinite-dimensional, meaning that you can find an infinite collection of linearly independent functions. However, the space $\mathcal{P}_{2L}$ of $2L$+periodic functions resembles $\mathbb{R}^n$ in one crucial aspect: it is an *inner product space.*

15.2 Inner products

Recall that we can take the dot product of vectors in $\mathbb{R}^3$:

$$\vec{a} \cdot \vec{b} = (a_1, a_2, a_3)^T \cdot (b_1, b_2, b_3)^T = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

This can be readily generalized to $\mathbb{R}^n$:

$$\vec{x} \cdot \vec{y} = (x_1, \ldots, x_n)^T \cdot (y_1, \ldots, y_n)^T = x_1 y_1 + \ldots + x_n y_n.$$

The dot product is an example of an *inner product*: a function that takes in two vectors and spits out a real number, and satisfies the following properties:

1. Symmetry: $\vec{y} \cdot \vec{x} = \vec{x} \cdot \vec{y}$
2. Bilinearity: $(a_1 \vec{x}_1 + a_2 \vec{x}_2) \cdot y = a_1\, \vec{x}_1 \cdot \vec{y} + a_2\, \vec{x}_2 \cdot \vec{y}$ and $\vec{x} \cdot (b_1 \vec{y}_1 + b_2 \vec{y}_2) = b_1\, \vec{x} \cdot \vec{y}_1 + b_2\, \vec{x} \cdot \vec{y}_2$

3. Positive-definiteness: $\vec{x}\cdot\vec{x} \geq 0$ always, and $\vec{x}\cdot\vec{x} = 0$ if and only if $\vec{x} = 0$.

Using the dot product, we can define geometric notions such as length and angle. The **norm** of a vector $\vec{x}$ defined to be $|\vec{x}| = \sqrt{\vec{x}\cdot\vec{x}}$. Also, if $\vec{x}$ and $\vec{y}$ are nonzero vectors, we define the angle θ between them by the formula

$$\cos\theta = \frac{\vec{x}\cdot\vec{y}}{|\vec{x}||\vec{y}|}.$$

The vectors $\vec{x}$ and $\vec{y}$ are said to be **orthogonal** if $\vec{x}\cdot\vec{y} = 0$. If

$$\vec{e}_i = (0, \ldots, 1, \ldots, 0)^T$$

denotes the ith standard basic vector, then for any $\vec{x} = (x_1, \ldots, x_n)^T$, we have

$$\vec{x}\cdot\vec{e}_i = x_i,$$

so we can recover the coordinates of any vector by taking inner products with the vectors in the standard basis.

We can carry these notions to the space of $2L$-periodic functions by defining the following inner product:

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

This inner product satisfies the three formal properties listed above. (We have to be a little careful with the very last part: if $(f,f) = 0$, then $f(x) = 0$ *outside a set of measure zero*. We don't have to worry about the precise meaning of this. If f is continuous, then $f(x) = 0$ everywhere.) This inner product satisfies one more property that is very important for us: under this inner product, the set

$$\left\{\frac{1}{\sqrt{2}}, \sin\frac{\pi x}{L},\ \cos\frac{\pi x}{L},\ \sin\frac{2\pi x}{L},\ \cos\frac{2\pi x}{L},\ \ldots,\ \sin\frac{n\pi x}{L},\ \cos\frac{n\pi x}{L},\ \ldots\right\}$$

is an **orthonormal set**, meaning that if f and g are in this set, then $(f,g) = 0$ if $f \neq g$, and $(f,f) = 1$. This translates into the following **orthogonality relations**:

$$\frac{1}{L}\int_{-L}^{L}\sin\frac{m\pi x}{L}\sin\frac{n\pi x}{L}dx = \frac{1}{L}\int_{-L}^{L}\cos\frac{m\pi x}{L}\cos\frac{n\pi x}{L}dx = \begin{cases} 0, & m \neq n, \\ 1, & m = n. \end{cases}$$

$$\frac{1}{\sqrt{2}L}\int_{-L}^{L}\sin\frac{n\pi x}{L}dx = \frac{1}{\sqrt{2}L}\int_{-L}^{L}\cos\frac{m\pi x}{L}dx = \frac{1}{L}\int_{-L}^{L}\cos\frac{m\pi x}{L}\sin\frac{n\pi x}{L}dx = 0.$$

For example, when $m \neq n$, using the trigonometric formulas

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$$

we obtain

$$\begin{aligned}
\int_{-L}^{L} \sin\frac{m\pi x}{L}\sin\frac{n\pi x}{L}dx &= \frac{1}{2}\int_{-L}^{L}\left(\cos\frac{(m-n)\pi x}{L} - \cos\frac{(m+n)\pi x}{L}\right)dx \\
&= \frac{L}{2\pi}\Bigg|_{-L}^{L}\left(\frac{1}{m-n}\sin\frac{(m-n)\pi x}{L} - \frac{1}{m+n}\sin\frac{(m+n)\pi x}{L}\right) \\
&= 0.
\end{aligned}$$

Another example of an orthonormal set is the standard basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ of $\mathbb{R}^n$.

16 Lecture 16

16.1 Recap: inner product on the space of periodic functions

We defined an inner product on the space $\mathcal{P}_{2L}$ of $2L$-periodic functions by

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

Under this inner product, the set

$$\left\{\frac{1}{\sqrt{2}}, \sin\frac{\pi x}{L},\ \cos\frac{\pi x}{L},\ \sin\frac{2\pi x}{L},\ \cos\frac{2\pi x}{L},\ \ldots,\ \sin\frac{n\pi x}{L},\ \cos\frac{n\pi x}{L},\ \ldots\right\}$$

is an **orthonormal set**, meaning that if f and g are in this set, then $(f, g) = 0$ if $f \neq g$, and $(f, f) = 1$. This translates to the equations

$$\frac{1}{L}\int_{-L}^{L}\sin\frac{m\pi x}{L}\sin\frac{n\pi x}{L}dx \quad = \quad \frac{1}{L}\int_{-L}^{L}\cos\frac{m\pi x}{L}\cos\frac{n\pi x}{L}dx \quad = \quad \begin{cases} 0, & m \neq n, \\ 1, & m = n. \end{cases}$$

$$\frac{1}{\sqrt{2}L}\int_{-L}^{L}\sin\frac{n\pi x}{L}dx \quad = \quad \frac{1}{\sqrt{2}L}\int_{-L}^{L}\cos\frac{m\pi x}{L}dx \quad = \quad \frac{1}{L}\int_{-L}^{L}\cos\frac{m\pi x}{L}\sin\frac{n\pi x}{L}dx \quad = \quad 0,$$

where $m, n \geq 1$, and one more equation $\int_{-L}^{L}\frac{1}{2}dx = L$. One of these is proved in the book, and a couple others are left as homework. The idea is to use angle sum formulas to express the product of trigonometric functions as a sum of trigonometric functions, and then simply integrate. Another example of an orthonormal set is the standard basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ of $\mathbb{R}^n$.

One way of creating $2L$-periodic functions is to take a function defined on $[-L, L]$, and simply "copy-and-paste" the function so that it becomes defined on the whole real line. This will rarely result in a continuous function, but that doesn't really matter.

Example 16.1.1. We get a triangular wave by starting with the function $f(x) = |x|$ on, for example, the interval $[-2, 2]$ and copying this segment around so that $f(x + 4) = f(x)$ for all x.

16.2 Fourier series

The key idea in handling periodic functions is to try to express them as sums of sines and cosines. A **Fourier series** is a converging series of the form

$$\frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos\frac{m\pi x}{L} + \sum_{m=1}^{\infty} b_m \sin\frac{m\pi x}{L}.$$

"Converging" here means that for every x, the sequence of partial sums

$$S_N(x) = \frac{a_0}{2} + \sum_{m=1}^{N} a_m \cos\frac{m\pi x}{L} + \sum_{m=1}^{N} b_m \sin\frac{m\pi x}{L}$$

gets closer and closer to some number as $N \to \infty$. We see immediately that a Fourier series is a $2L$-periodic function, so the important question now is the following. Given a $2L$-periodic function f, is f expressible as Fourier series, and how to find the **Fourier-coefficients** a_m, b_m?

We will first concentrate on the second question. Assume that we are given a $2L$-periodic function f, and wish to write it as a Fourier series:

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos\frac{m\pi x}{L} + \sum_{m=1}^{\infty} b_m \sin\frac{m\pi x}{L}$$

The key is to use the orthogonality properties of sines and cosines discussed above. To find a_n, where $n \geq 1$, we multiply both sides by $\cos\frac{n\pi x}{L}$ and integrate from $-L$ to L:

$$\begin{aligned}
\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}dx &= \int_{-L}^{L}\left(\frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos\frac{m\pi x}{L} + \sum_{m=1}^{\infty} a_m \sin\frac{m\pi x}{L}\right)\cos\frac{n\pi x}{L}dx \\
&= \int_{-L}^{L}\frac{a_0}{2}\cos\frac{n\pi x}{L}dx + \sum_{m=1}^{\infty} a_m \int_{-L}^{L}\cos\frac{m\pi x}{L}\cos\frac{n\pi x}{L}dx \\
&\quad + \sum_{m=1}^{\infty} a_m \int_{-L}^{L}\sin\frac{m\pi x}{L}\cos\frac{n\pi x}{L}dx.
\end{aligned}$$

We won't worry about interchanging the sum and integral. Using now the orthogonality relations, the only nonzero term is the one having $\cos\frac{n\pi x}{L}$ twice. Thus,

$$\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}dx = La_n,$$

so that

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}dx.$$

Note that this is just the inner product of f with $\cos\frac{n\pi x}{L}$. Similar calculation gives

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}dx$$

and

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(x)dx.$$

We can thus use the same formula for a_0 and $a_n, n > 0$. Note that $a_0/2$ is just the average of the function f on the interval $[-L, L]$. The formulas for a_n and b_n are called the **Euler-Fourier formulas**. We can think of the Fourier coefficients as the "coordinates with respect to the basis" consisting of sines and cosines (and $\frac{1}{2}$).

Example 16.2.1. Let's find the Fourier coefficients of the function from the previous example:

$$f(x) = \begin{cases} -x, & -2 \leq x < 0 \\ x, & 0 \leq x \leq 2 \end{cases}, \quad f(x+4) = f(x).$$

Note that this function is 4-periodic. We have

$$\begin{aligned}
a_0 &= \frac{1}{2}\int_{-2}^{2} f(x)dx = \frac{1}{2}\int_{-2}^{0}(-x)dx + \frac{1}{2}\int_{0}^{2} xdx = \\
&= \frac{1}{2}\Big|_{-2}^{0} -\frac{1}{2}x^2 + \frac{1}{2}\Big|_{-2}^{0}\frac{1}{2}x^2 = 2, \\
a_n &= \frac{1}{2}\int_{-2}^{2} f(x)\cos\frac{n\pi x}{2}dx = \frac{1}{2}\int_{-2}^{0}\left(-x\cos\frac{n\pi x}{2}\right)dx + \frac{1}{2}\int_{0}^{2} x\cos\frac{n\pi x}{2}dx \\
&= \int_{0}^{2} x\cos\frac{n\pi x}{2}dx = \Big|_{0}^{2}\frac{2}{n\pi}x\sin\frac{n\pi x}{2} - \int_{0}^{2}\frac{2}{n\pi}\sin\frac{n\pi x}{2}dx = \Big|_{0}^{2}\frac{4}{(n\pi)^2}\cos\frac{n\pi x}{2} \\
&= \begin{cases} -\frac{8}{(n\pi)^2}, & n \text{ odd} \\ 0, & n \text{ even} \end{cases} \\
b_n &= \frac{1}{2}\int_{-2}^{2} f(x)\sin\frac{n\pi x}{2}dx = \frac{1}{2}\int_{-2}^{0}\left(-x\sin\frac{n\pi x}{2}\right)dx + \frac{1}{2}\int_{0}^{2} x\sin\frac{n\pi x}{2}dx = 0.
\end{aligned}$$

Thus,

$$f(x) = \frac{2}{2} - \sum_{n \text{ odd}} \frac{8}{n^2\pi^2}\cos\frac{n\pi x}{2} = 1 - \frac{8}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}\cos\frac{(2n-1)\pi x}{2}.$$

17 Lecture 17

17.1 Fourier convergence theorem

A function that can be represented as a Fourier series must be periodic, but not every periodic function has a converging Fourier series. Fortunately, most nice functions do.

A function f defined on an interval $[a, b]$ is **piecewise continuous** if there is a partition $a = x_0 < x_1 < \ldots < x_n = b$ of the interval, such that f is continuous on each subinterval (x_i, x_{i+1}), and the limits

$$f(x_i+) = \lim_{x \to x_i^+} f(x), \quad i = 0, \ldots, n-1, \quad f(x_i-) = \lim_{x \to x_i^-} f(x), \quad i = 1, \ldots, n$$

exist and are finite. It is not necessary for the function to be defined at the points x_i. In fact, if f is differentiable on the (open) subintervals but discontinuous at the points x_i, then the derivative is piecewise continuous but not defined at the points x_i.

Theorem 17.1.1. *If f is a $2L$-periodic function such that f and f' are piecewise continuous on $[-L, L]$, then the Fourier series of f converges to $f(x)$ at the points x where f is continuous, and to the value $(f(x+) + f(x-))/2$ at the points of discontinuity of f.*

Example 17.1.2. A "square wave" is given by the function

$$f(x) = \begin{cases} 0, & -L \le x < 0 \\ L, & 0 \le x < L \end{cases}$$

and $f(x + 2L) = f(x)$ for all x. The function f is $2L$-periodic and piecewise continuous with jump discontinuities at the points $x = kL, k = 0, \pm 1, \pm 2, \ldots$. The Fourier coefficients of f are

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}dx = \frac{1}{L}\int_{0}^{L} L\cos\frac{n\pi x}{L}dx = \Big|_0^L \sin\frac{n\pi x}{L}\frac{L}{n\pi} = 0,$$

if $n = 1, 2, \ldots$,

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(x)dx = \frac{1}{L}\int_{0}^{L} Ldx = L,$$

and

$$\begin{aligned} b_n &= \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}dx = \frac{1}{L}\int_{0}^{L} L\sin\frac{n\pi x}{L}dx = \Big|_0^L - \cos\frac{n\pi x}{L}\frac{L}{n\pi} \\ &= \frac{L}{n\pi}(1 - \cos n\pi) = \begin{cases} 0, & n \text{ even} \\ \frac{2L}{n\pi}, & n \text{ odd.} \end{cases} \end{aligned}$$

Thus,

$$f(x) = \frac{L}{2} + \frac{2L}{\pi}\sum_{n=1}^{\infty}\frac{1}{2n-1}\sin\frac{(2n-1)\pi x}{L}.$$

This series converges to the value of f when $x \ne kL, k = 0, \pm 1, \pm 2, \ldots$, and it converges to the value $\frac{L}{2}$ at the points of discontinuity $x = kL, k = 0, \pm 1, \pm 2, \ldots$. In particular, the Fourier series *does not* converges to the value of f at these points (although we could redefine f at these points to fix this).

17.2 Even and odd functions

There are two important classes of periodic functions for which the Fourier series simplifies considerably: even and odd functions.

A function f (not necessarily periodic) is **even** if $f(-x) = f(x)$ for all x, and it is **odd** if $f(-x) = -f(x)$. The graph of an even function is symmetric with respect to the y-axis, while the graph of an odd function is symmetric with respect to the origin.

Example 17.2.1. Examples of even functions are all constants, and

$$x^2,\ x^4,\ x^{2n},\ \frac{1}{1+x^2},\ e^{-x^2},\ \cosh x,\ |x|,\ \cos ax.$$

Some odd functions are

$$x,\ x^3,\ x^{2n+1},\ \arctan x,\ \sinh x,\ \sin ax.$$

The only function that is both even and odd is the zero function.

Even and odd functions interact as follows.

1. If f and g are even, then $f+g$ is even.
2. If f and g are odd, then $f+g$ is odd.
3. If f and g are even, then fg is even.
4. If f and g are odd, then fg is even.
5. If f is even and g is odd, then fg is odd.
6. If f is even (resp. odd), then $\frac{1}{f}$ is even (resp. odd).

The proofs are simple. For example, if f is even and g is odd, then

$$(fg)(-x) = f(-x)g(-x) = f(x)(-g(x)) = -f(x)g(x) = -(fg)(x),$$

so fg is odd.

The most useful property of even and odd functions for us is related to integration.

Lemma 17.2.2. *If f is an even (integrable) function, then*

$$\int_{-L}^{L} f(x)dx = 2\int_0^L f(x)dx.$$

If f is an odd function, then

$$\int_{-L}^{L} f(x)dx = 0.$$

Proof. Assume that f is even. We split the integral into to halves:

$$\int_{-L}^{L} f(x)dx = \int_{-L}^{0} f(x)dx + \int_0^L f(x)dx.$$

In the first integral on the right, we make a change of variables $x=-u$:

$$\int_{-L}^{0} f(x)dx = \int_{L}^{0} f(-u)d(-u).$$

Since $f(-u) = f(u)$, and $d(-u) = -du$, once we interchange the limits of integration, we have

$$\int_{-L}^{0} f(x)dx = \int_{0}^{L} f(u)du = \int_{0}^{L} f(x)dx.$$

Thus,

$$\int_{-L}^{L} f(x)dx = \int_{-L}^{0} f(x)dx + \int_{0}^{L} f(x)dx = \int_{0}^{L} f(x)dx + \int_{0}^{L} f(x)dx = 2\int_{0}^{L} f(x)dx.$$

The odd case is similar, except this time $f(-x) = -f(x)$, so we get one more minus sign, and the two halves cancel out giving zero. $\square$

18 Lecture 18

18.1 Cosine and sine series

Last time we defined even and odd functions, and found that if f is an even function, then

$$\int_{-L}^{L} f(x)\ dx = 2\int_{0}^{L} f(x)\ dx,$$

and if f is an odd function, then

$$\int_{-L}^{L} f(x)\ dx = 0.$$

Using these integral properties, we can simplify the Fourier series of even and odd functions. Assume that f is an even $2L$-periodic function, with Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos\frac{m\pi x}{L} + \sum_{m=1}^{\infty} b_m \cos\frac{m\pi x}{L},$$

where the Fourier coefficients are

$$a_m = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{m\pi x}{L}dx, \quad m = 0, 1, 2, \ldots, \quad b_m = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{m\pi x}{L}dx, \quad m = 1, 2, \ldots$$

Since f is even, so is $f(x)\cos\frac{m\pi x}{L}$ for each $m = 0, 1, 2, \ldots$, where as $f(x)\sin\frac{m\pi x}{L}$ is odd for each $m = 1, 2, \ldots$. Thus, from the lemma above, we have

$$a_m = \frac{2}{L}\int_{0}^{L} f(x)\cos\frac{m\pi x}{L}dx, \quad m = 0, 1, 2, \ldots, \quad b_m = 0, \quad m = 1, 2, \ldots,$$

and the Fourier series simplifies to a **cosine series**

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos\frac{m\pi x}{L}.$$

Notice also that we found a different (and some times a little simpler) formula for the coefficients a_m.

Similarly, if f is and $2L$-periodic, then $f(x)\cos\frac{m\pi x}{L}$ is odd for each $m = 0, 1, 2, \ldots$, and $f(x)\sin\frac{m\pi x}{L}$ is even for each $m = 1, 2, \ldots$, so this time

$$a_m = 0, \quad m = 0, 1, \ldots, \quad b_m = \frac{2}{L}\int_0^L f(x)\sin\frac{m\pi x}{L}dx,$$

and the Fourier series simplifies to a **sine series**

$$f(x) = \sum_{m=1}^{\infty} b_m \sin\frac{m\pi x}{L}.$$

Example 18.1.1. The "sawtooth wave" is given by the function $f(x) = x, -L < x < L, f(-L) = f(L) = 0$, and $f(x + 2L) = f(x)$. This is an odd function, so its Fourier series is a sine series. In other words, $a_n = 0, n = 0, 1, 2, \ldots$, and

$$b_n = \frac{2}{L}\int_0^L x\sin\frac{n\pi x}{L}dx = (-1)^{n+1}\frac{2L}{n\pi},$$

and thus

$$f(x) = \frac{2L}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\frac{n\pi x}{L}.$$

18.2 Extending a function

If we are given a function f defined on the interval $[0, L]$, we frequently want to represent f as a Fourier series. For this, we have to make f a $2L$-periodic functions. This means that we must first of all define f on $(-L, 0)$, and then extend it periodically to all of $\mathbb{R}$. There are of course infinitely many ways of extending f to $(-L, 0)$ without changing its values on $[0, L]$, but some convenient choices are the following.

1. Extending f to an even function: define $f(x) = f(-x)$ for $-L < x < 0$. The resulting Fourier series is a *cosine series.*
2. Extending f to an odd function: define $f(x) = -f(-x)$ for $-L < x < 0$. The resulting Fourier series is a *sine series.*
3. Extending f by zero: define $f(x) = 0$ for $-L < x < 0$. The resulting Fourier series will have both sine and cosine terms.

Example 18.2.1. Consider the function

$$f(x) = \begin{cases} 1 - x, & 0 \le x \le 1 \\ 0, & 1 < x \le 3. \end{cases}$$

defined on $[0,3]$, so that $L=3$. If we extend f to an even function on $[-1,1]$, then f is given by

$$f(x)=\begin{cases} 0, & -3<x<-1 \\ 1+x, & -1\le x<0 \\ 1-x, & 0\le x\le 1 \\ 0, & 1<x\le 3. \end{cases}$$

If we extend f to an odd function, then f is given by

$$f(x)=\begin{cases} 0, & -3<x<-1 \\ -1-x, & -1\le x<0 \\ 1-x, & 0\le x\le 1 \\ 0, & 1<x\le 3. \end{cases}$$

If we extend f by zero, then f is given by

$$f(x)=\begin{cases} 0, & -3<x<0 \\ 1-x, & 0\le x\le 1 \\ 0, & 1<x\le 3. \end{cases}$$

You will compute the Fourier series for the first two extensions in homework.

Example 18.2.2. Let's find the Fourier series of the 4-periodic even extension of the function

$$f(x)=\begin{cases} x, & 0\le x<1, \\ 1, & 1\le x\le 2. \end{cases}$$

The extension is given by

$$f(x)=\begin{cases} 1, & -2\le x<-1, \\ -x, & -1\le x<0, \\ x, & 0\le x<1, \\ 1, & 1\le x<2, \end{cases}$$

and the Fourier coefficients are

$$a_n=\frac{1}{2}\int_0^2 f(x)\cos\frac{n\pi x}{2}dx=\begin{cases} -\frac{2}{n^2\pi^2}, & n \text{ odd} \\ 0, & n=4k, k \text{ integer} \\ -\frac{4}{n^2\pi^2}, & n=4k+2, k \text{ integer}. \end{cases}$$

19 Lecture 19

19.1 Heat equation

We next study the problem of heat conduction in a thin solid rod. In other words, we study how the temperature distribution changes as a function of time. We assume that the points of the rod can be parameterized by a single spatial variable x. The temperature distribution is given by a function $u(x,t)$ depending on the point x and also on time t.

Assume that the rod has length L and is placed on the x-axis so that one end is at $x = 0$ and the other is at $x = L$. Assume that the sides of the rod are perfectly insulated, and that the two ends are kept at zero temperature. Assume also that at time $t = 0$ the temperature distribution is given by some function $f = f(x)$. The function $u(x,t)$ satisfies the conditions of the following **heat conduction problem**:

$$\begin{aligned} u_t &= \alpha^2 u_{xx}, && 0 < x < L, \quad t > 0, \\ u(0,t) = u(L,t) &= 0, && t > 0, \\ u(x,0) &= f(x), && 0 \le x \le L. \end{aligned}$$

The equation $u_t = \alpha^2 u_{xx}$ is called the **heat equation**. It is an example of a **partial differential equation**, or **PDE**, an equation that involves an unknown function depending on more than one variable that expresses a relationship between its various partial derivatives. The positive constant α is a physical parameter. In terms of the time variable t, this problem is an initial value problem, whereas in terms of the spatial variable x, it is a boundary value problem. Alternatively, we can view it as a boundary value problem on the half-infinite strip $\{(x,t) \,:\, 0 \le x \le L, t \ge 0\}$.

19.2 Separation of variables

Our task now is to solve the heat equation. The first observation that we make is that the heat equation is linear and homogeneous, and also the boundary conditions $u(0,t) = u(L,t) = 0$ are homogeneous, so any linear combination of functions satisfying the differential equation and the boundary conditions also satisfies them. In other words, the heat equation together with the homogeneous boundary conditions satisfies the superposition principle. Thus, we could hope to solve the whole problem by finding simple solutions to the heat equation that satisfy the boundary conditions, and superimposing these solutions to find a solution that satisfies the initial condition $u(x,0) = f(x)$.

To find these simple solutions to the problem $u_t = \alpha^2 u_{xx}, u(0,t) = u(L,t) = 0$, we use a technique called **separation of variables**. In this method, we assume that the solution $u(x,t)$ has a simple form: it is the product of two functions, one that depends only on x and one that depends only on t. In other words, we assume

$$u(x,t) = X(x)T(t).$$

Substituting this into the heat equation gives

$$XT' = \alpha^2 X''T,$$

which leads to

$$\frac{1}{\alpha^2}\frac{T'}{T} = \frac{X''}{X}.$$

Since the left hand side only depends on x and the right hand side only depend on t, this equation can only be satisfied if both sides are equal to the same constant. We call this constant the **separation constant** and denote it by $-\lambda$. Thus, we obtain two ordinary differential equations

$$X'' + \lambda X = 0, \qquad T' + \lambda\alpha^2 T = 0.$$

Consider now the two boundary condition:

$$u(0,t) = X(0)T(t) = 0, \quad u(L,t) = X(L)T(t) = 0, \quad t \geq 0.$$

If $X(0) \neq 0$ or $X(L) \neq 0$, then we must have $T(t) = 0$ for $t > 0$, hence $u(x,t) = 0$ for $t > 0$. But we are interested in *nonzero* solutions of the boundary value problem so that we can later superimpose them to satisfy the initial condition. Thus, we must have $X(0) = X(L) = 0$. This means that X has to be an eigenfunction of the boundary value problem

$$X'' + \lambda X = 0, \quad X(0) = X(L) = 0.$$

We already know what the eigenfunctions and corresponding eigenfunctions are:

$$X_n(x) = \sin\frac{n\pi x}{L}, \quad \lambda_n = \frac{n^2\pi^2}{L^2}, \quad n = 1, 2, \ldots.$$

For each n, we get the corresponding equation

$$T' + \frac{n^2\pi^2\alpha^2}{L^2}T = 0,$$

which has solution (up to multiplicative constant)

$$T_n(t) = e^{-\frac{n^2\pi^2\alpha^2 t}{L^2}}.$$

Thus, for each n, the function

$$u_n(x,t) = X_n(x)T_n(t) = e^{-\frac{n^2\pi^2\alpha^2 t}{L^2}} \sin\frac{n\pi x}{L}$$

is a solution to the heat equation $u_t = \alpha^2 u_{xx}$ satisfying the homogeneous boundary conditions $u(0,t) = u(L,t) = 0, t > 0$. These functions are called the **fundamental solutions** of the problem.

Now we use these simple solutions to find a solution that also satisfies the initial condition $u(x,0) = f(x)$. Recall that since the heat equation is linear and homogeneous, and the boundary conditions $u(0,t) = u(L,t) = 0, t > 0$ are homogeneous, any linear combination of solutions is also as solution. We extend this idea to infinite linear combinations: if the coefficients c_n are chosen suitably, the series

$$u(x,t) = \sum_{n=1}^{\infty} c_n u_n(x,t) = \sum_{n=1}^{\infty} c_n e^{-\frac{n^2\pi^2\alpha^2 t}{L^2}} \sin\frac{n\pi x}{L}$$

converges, and is a solution of the heat equation and satisfies the boundary conditions. We now try to choose the coefficients $c_n, n = 1, 2, \ldots$ so that the series satisfies the initial condition $u(x,0) = f(x)$. This means that we must have

$$u(x,0) = \sum_{n=1}^{\infty} c_n \sin\frac{n\pi x}{L} = f(x).$$

This is precisely the sine series associated to the odd $2L$-periodic extension of f, so the coefficients must be given by the Euler-Fourier formula

$$c_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}dx.$$

To recap our findings, the solution to the heat conduction problem

$$\begin{aligned}
u_t &= \alpha^2 u_{xx}, && 0 < x < L, \quad t > 0, \\
u(0,t) = u(L,t) &= 0, && t > 0, \\
u(x,0) &= f(x), && 0 \le x \le L.
\end{aligned}$$

is

$$u(x,t) = \sum_{n=1}^{\infty} c_n e^{-\frac{n^2\pi^2\alpha^2 t}{L^2}} \sin\frac{n\pi x}{L}, \quad \text{where} \quad c_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}dx.$$

Example 19.2.1. Consider a 50 cm long metal rod, insulated on the sides, which initially has uniform temperature 20°C throughout and whose ends are maintained at 0°C for all $t > 0$. The temperature $u(x,t)$ satisfies the above heat conduction problem with $L = 50$ and $f(x) = 20, \quad 0 < x < 50$. Now

$$c_n = \frac{2}{50}\int_0^{50} 20\sin\frac{n\pi x}{50}dx = \frac{40}{n\pi}(1-\cos n\pi) = \begin{cases} \frac{80}{n\pi}, & n \text{ odd}, \\ 0, & n \text{ even}, \end{cases}$$

so the solution of the heat conduction problem is

$$u(x,t) = \frac{80}{\pi}\sum_{n=1,3,5,\ldots} \frac{1}{n} e^{\frac{-n^2\pi^2\alpha^2}{2500}t} \sin\frac{n\pi x}{50}.$$

20 Lecture 20

20.1 Heat equation: recap

Last time we studied the heat conduction problem

$$\begin{aligned}
u_t &= \alpha^2 u_{xx}, && 0 < x < L, \quad t > 0, \\
u(0,t) = u(L,t) &= 0, && t > 0, \\
u(x,0) &= f(x), && 0 \le x \le L.
\end{aligned}$$

We used separation of variables to find fundamental solutions

$$u_n(x,t) = e^{-\frac{n^2\pi^2\alpha^2}{L^2}} \sin\frac{n\pi x}{L}$$

that satisfy the heat equation with boundary conditions

$$u_t = \alpha^2 u_{xx}, \qquad u(0,t) = u(L,t) = 0,$$

and took an "infinite linear combination" of the fundamental solutions to obtain a solution that satisfies the initial condition $u(x,0) = f(x)$:

$$u(x,t) = \sum_{n=1}^{\infty} c_n e^{-\frac{n^2\pi^2\alpha^2 t}{L^2}} \sin\frac{n\pi x}{L}, \quad \text{where} \quad c_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}dx.$$

The coefficients c_n turned out to be the Fourier coefficients of the odd $2L$-periodic extension of the function $f(x)$.

Example 20.1.1. Consider a 50 cm long metal rod, insulated on the sides, which initially has uniform temperature 20°C throughout and whose ends are maintained at 0°C for all $t > 0$. The temperature $u(x,t)$ satisfies the above heat conduction problem with $L = 50$ and $f(x) = 20, \quad 0 < x < 50$. Now

$$c_n = \frac{2}{50}\int_0^{50} 20 \sin\frac{n\pi x}{50}dx = \frac{40}{n\pi}(1-\cos n\pi) = \begin{cases} \frac{80}{n\pi}, & n \text{ odd}, \\ 0, & n \text{ even}, \end{cases}$$

so the solution of the heat conduction problem is

$$u(x,t) = \frac{80}{\pi}\sum_{n=1,3,5,\ldots} \frac{1}{n} e^{\frac{-n^2\pi^2\alpha^2}{2500}t} \sin\frac{n\pi x}{50}.$$

20.2 Heat equation with inhomogeneous boundary conditions

We next consider the heat conduction problem in a thin, insulated bar of length L, whose one end is kept at temperature $u(0,t) = T_1$ and the other end at $u(L,t) = T_2$ for all $t > 0$, and that has initial temperature distribution given by $u(x,0) = f(x), \; 0 \le x \le L$. The idea is to reduce this problem to the one with homogeneous boundary conditions, whose solution we found last time.

We try to find the temperature distribution in the form $u(x,t) = v(x) + w(x,t)$: here v is a **steady state solution**, a solution of the heat conduction problem that is obtained as $t \to \infty$ and that does not depend on t, as the temperature distribution has had enough time to settle; and w is a "transient solution" that is needed for u to satisfy the initial condition and that approaches zero as $t \to \infty$.

Let's first find the steady state solution v. We require that v satisfies the heat equation and the boundary conditions, so since v_t equals zero, v is the solution of the boundary value problem

$$v'' = 0, \quad v(0) = T_1, \; v(L) = T_2.$$

Thus, v must be the linear function

$$v(x) = \frac{T_2 - T_1}{L}x + T_1.$$

Now we find the boundary value problem that w has to satisfy. Since the heat equation is linear, and both u and v satisfy it, also the linear combination $w = u - v$ must satisfy it. Moreover,

$$w(0,t) = u(0,t) - v(0) = T_1 - T_1 = 0, \qquad w(L,t) = u(L,t) - v(L) = T_2 - T_2 = 0,$$

and

$$w(x,0) = u(x,0) - v(x) = f(x) - \frac{T_2 - T_1}{L}x - T_1$$

Thus, w is the solution of the heat conduction problem

$$\begin{aligned} w_t &= \alpha^2 w_{xx}, && 0 < x < L, \quad t > 0, \\ w(0,t) = w(L,t) &= 0, && t > 0, \\ w(x,0) &= f(x) - \tfrac{T_2 - T_1}{L}x - T_1, && 0 \le x \le L, \end{aligned}$$

which has homogeneous boundary conditions, so we know how to solve it from last time. We now put the pieces together to obtain the solution u:

$$u(x,t) = \frac{T_2 - T_1}{L}x + T_1 + \sum_{n=1}^{\infty} c_n e^{-\frac{n^2\pi^2\alpha^2 t}{L^2}} \sin\frac{n\pi x}{L},$$

where

$$c_n = \frac{2}{L}\int_0^L \left(f(x) - \frac{T_2 - T_1}{L}x - T_1\right)\sin\frac{n\pi x}{L}dx.$$

Example 20.2.1. Consider the heat conduction problem

$$\begin{aligned}
u_t &= u_{xx}, && 0 < x < 30, \quad t > 0,\\
u(0,t) &= 20, && t > 0,\\
u(30,t) &= 50, && t > 0,\\
u(x,0) &= 60 - 2x && 0 \le x \le L,
\end{aligned}$$

Here we have $\alpha = 1, L = 30, T_1 = 20, T_2 = 50, f(x) = 60 - 2x$. Thus, the steady state solution is

$$v(x) = \frac{T_2 - T_1}{L}x + T_1 = x + 20.$$

The function w now satisfies the initial condition

$$w(x,0) = u(x,0) - v(x) = 60 - 2x - (x + 20) = 40 - 3x.$$

The Fourier coefficients for this function are

$$c_n = \frac{2}{30}\int_0^{30}(40 - 3x)\sin\frac{n\pi x}{30}dx = \frac{80 + 100(-1)^n}{n\pi}.$$

Thus, the solution of the heat conduction problem is

$$u(x,t) = x + 20 + \sum_{n=1}^{\infty}\frac{80 + 100(-1)^n}{n\pi}e^{\frac{-n^2\pi^2}{900}t}\sin\frac{n\pi x}{30}.$$

21 Lecture 21

21.1 Heat equation with insulated endpoints

Another variation of the heat conduction problem is when the endpoints of the rod are insulated so that no heat flows through them. This changes the boundary conditions to $u_x(0,t) = u_x(L,t) = 0$, so we obtain the heat conduction problem

$$\begin{aligned}
u_t &= \alpha^2 u_{xx}, && 0 < x < L, \quad t > 0,\\
u_x(0,t) = u_x(L,t) &= 0, && t > 0,\\
u(x,0) &= f(x), && 0 \le x \le L.
\end{aligned}$$

Separation of variables by $u(x,t) = X(x)T(t)$ leads again to the two ordinary differential equations

$$X'' + \lambda X = 0, \qquad T' + \lambda\alpha^2 T = 0.$$

The boundary conditions are now

$$u_x(0,t) = X'(0)T(t) = 0, \quad u_x(L,t) = X'(L)T(t) = 0, \quad t > 0,$$

so to find nonzero solutions, we must have $X'(0) = X'(L) = 0$. Thus, X must be an eigenfunction of the eigenvalue problem

$$X'' + \lambda X = 0, \quad X'(0) = X'(L) = 0.$$

We found them in homework, and the eigenfunctions and corresponding eigenvalues are

$$X_n(x) = \cos\frac{n\pi x}{L}, \quad \lambda_n = \frac{n^2\pi^2}{L^2}, \quad n = 0, 1, 2, \ldots.$$

For each n, we find $T_n(t)$ from the equation

$$T' + \frac{n^2\pi^2\alpha^2}{L^2}T = 0,$$

so

$$T_n(t) = e^{-\frac{n^2\pi^2\alpha^2}{L^2}t},$$

and we get the fundamental solutions

$$u_n(x,t) = X_n(x)T_n(t) = e^{-\frac{n^2\pi^2\alpha^2}{L^2}t}\cos\frac{n\pi x}{L}, \quad n = 0, 1, 2, \ldots$$

Note that also $u_0(x,t) \equiv 1$ is a fundamental solution corresponding to the eigenvalue $\lambda_0 = 0$. We look for a solution of the full heat conduction problem in the form

$$u(x,t) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n e^{-\frac{n^2\pi^2\alpha^2}{L^2}t}\cos\frac{n\pi x}{L}.$$

To satisfy the initial condition $u(x,0) = f(x)$, we must have

$$f(x) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos\frac{n\pi x}{L},$$

which is just the cosine series of the even extension of f. Thus, the coefficients c_n are determined by

$$c_n = \frac{2}{L}\int_0^L f(x)\cos\frac{n\pi x}{L}\,dx.$$

Note that the solution can be viewed as the sum of a steady state solution $u_0(x,t) = \frac{c_0}{2}$, which is just a constant, and a transient solution that tends to zero as $t \to \infty$. The steady state solution

$$\frac{c_0}{2} = \frac{1}{L}\int_0^L f(x)dx$$

is in fact the average of the initial temperature distribution. This agrees with our physical intuition: since no heat passes through the sides or ends of the rod, the total heat should remain constant, but as time progresses, heat conduction averages out the differences in temperature in different parts of the rod.

Example 21.1.1. Consider the temperature $u(x,t)$ in a metal rod of length 25 cm that is insulated at the ends as well as on the sides and whose initial temperature distribution is $u(x,0) = x$ for $0 < x < 25$. The function $u(x,t)$ satisfies the heat conduction problem

$$\begin{aligned} u_t &= \alpha^2 u_{xx}, & 0 < x < 25, \quad t > 0, \\ u_x(0,t) = u_x(25,t) &= 0, & t > 0, \\ u(x,0) &= x, & 0 \le x \le 25. \end{aligned}$$

We find the Fourier coefficients

$$c_0 = \frac{2}{25}\int_0^{25} x\,dx = 25, \quad c_n = \frac{2}{25}\int_0^{25} x\cos\frac{n\pi x}{25}dx = \begin{cases} -\frac{100}{n^2\pi^2}, & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases}$$

Thus, the temperature distribution is given by

$$u(x,t) = \frac{25}{2} - \frac{100}{\pi^2}\sum_{n=1,3,5,\ldots} \frac{1}{n^2} e^{-\frac{n^2\pi^2\alpha^2}{625}t}\cos\frac{n\pi x}{25}.$$

22 Lecture 22

22.1 Wave equation

The next partial differential equation we will discuss is called the wave equation, and it describes the propagation of waves in a continuous medium. Consider an elastic string (such as a guitar string) of length L that is tightly stretched between two points. Assume that when the string is at rest, it lies along the x-axis, with endpoints at $x = 0$ and $x = L$. We let $u(x,t)$ denote the vertical displacement of the string at point x and time t. If we ignore damping effects, and the amplitude of the displacement is small enough, the function u satisfies the **wave equation**

$$u_{tt} = a^2 u_{xx}$$

for $0 < x < L$ and $t > 0$. The constant a is a physical constant depending on the material, and turns out to be the velocity of the propagation of waves in the medium.

Since the wave equation is a second order differential equation in the time variable, we need to specify both the initial position and the initial velocity of the string. Assume that these are given by some functions f and g respectively, so that

$$u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad 0 \le x \le L.$$

Since the endpoints are fixed, we assume

$$u(0,t) = 0, \quad u(L,t) = 0, \quad t > 0,$$

so we must also require that

$$f(0) = f(L) = 0, \quad g(0) = g(L) = 0.$$

We will break the solution of the wave propagation problem into two parts, one where $g = 0$, and one where $f = 0$, and then put these solutions together to obtain a solution for arbitrary initial data f, g.

22.2 Zero initial velocity

In this case we assume that the string is disturbed from its equilibrium and then released at $t = 0$ with zero velocity. Then the vertical displacement $u(x,t)$ must satisfy the conditions

$$\begin{aligned} u_{tt} &= a^2 u_{xx}, && 0 < x < L, \quad t > 0, \\ u(0,t) = u(L,t) &= 0, && t \geq 0, \\ u(x,0) &= f(x), && 0 \leq x \leq L, \\ u_t(x,0) &= 0, && 0 \leq x \leq L. \end{aligned}$$

We solve this problem using separation of variables: assume that $u(x,t) = X(x)T(t)$ for some functions X and T. Substituting this into the wave equation gives

$$XT'' = a^2 X'' T,$$

which leads to

$$\frac{X''}{X} = \frac{1}{a^2}\frac{T''}{T} = -\lambda,$$

where λ is the separation constant. This leads to the two ordinary differential equations

$$X'' + \lambda X = 0, \quad T'' + a^2 \lambda T = 0.$$

As with the heat equation, to find nonzero solutions that satisfy the boundary conditions, we must require $X(0) = X(L) = 0$, and this leads to the eigenfunctions and eigenvalues

$$X_n(x) = \sin\frac{n\pi x}{L}, \quad \lambda_n = \frac{n^2\pi^2}{L^2}, \quad n = 1, 2, \ldots.$$

For each λ_n, we solve T from the equation

$$T'' + \frac{n^2\pi^2 a^2}{L^2} T = 0,$$

so

$$T(t) = k_1 \cos\frac{n\pi a t}{L} + k_2 \sin\frac{n\pi a t}{L}$$

The initial condition $u_t(x,0) = 0$ now translates to $X(x)T'(0) = 0$ for $0 \leq x \leq L$, and since we seek nonzero solutions, we must require $T'(0) = 0$, so we must have $k_2 = 0$, so that

$$T_n(t) = k_1 \cos\frac{n\pi a t}{L},$$

and we get the set of fundamental solutions

$$u_n(x,t) = X_n(x)T_n(t) = \sin\frac{n\pi x}{L}\cos\frac{n\pi a t}{L}.$$

Just like with the heat equation, since the wave equation is linear and the boundary conditions $u(0,t) = u(L,t) = 0$ are homogeneous, the problem satisfies the superposition principle, so linear combinations of solutions are also solutions. Thus, we attempt to find the solution satisfying the initial condition $u(x,0) = f(x)$ as an "infinite linear combination"

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{L} \cos \frac{n\pi a t}{L}.$$

To satisfy the initial condition, we must have

$$u(x,0) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{L} = f(x).$$

This is just the Fourier series of the odd $2L$-periodic extension of f, so the coefficients c_n are determined by

$$c_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} dx.$$

Example 22.2.1. Consider a vibrating string of length $L = 30$ that satisfies the wave equation

$$u_{tt} = 4u_{xx}, \quad 0 < x < 30, \quad t > 0.$$

Assume that the ends of the string are fixed and that the string is set in motion with no initial velocity from the initial position

$$u(x,0) = f(x) = \begin{cases} x/10, & 0 \leq x \leq 10, \\ (30-x)/20, & 10 < x \leq 30. \end{cases}$$

Here we have $a = 2, L = 30$, so the solution is given by

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{30} \cos \frac{2n\pi t}{30},$$

where

$$c_n = \frac{2}{30} \int_0^{30} f(x) \sin \frac{n\pi x}{30} dx = \frac{9}{n^2\pi^2} \sin \frac{n\pi}{3}.$$

22.3 Wave equation: zero initial displacement

Next we modify the problem as follows: we assume that the string starts to vibrate from its equilibrium position with initial velocity given by some function g. Thus, we study the problem

$$\begin{aligned} u_{tt} &= a^2 u_{xx}, && 0 < x < L, \quad t > 0, \\ u(0,t) = u(L,t) &= 0, && t \geq 0, \\ u(x,0) &= 0, && 0 \leq x \leq L, \\ u_t(x,0) &= g(x), && 0 \leq x \leq L. \end{aligned}$$

Separation of variables leads to the same eigenfunctions

$$X_n(x) = \sin \frac{n\pi x}{L}, \quad n = 1, 2, \ldots$$

and equations

$$T'' + \frac{n^2\pi^2 a^2}{L^2} T = 0,$$

whose solution is

$$T(t) = k_1 \cos \frac{n\pi a t}{L} + k_2 \sin \frac{n\pi a t}{L}.$$

This time however, the initial condition $u(x, 0) = X(x)T(0) = 0$ dictates that $T(0) = 0$ in order for us to find nonzero solutions. Thus, we must have $k_1 = 0$, so

$$T_n(t) = \sin \frac{n\pi a t}{L}, \quad n = 1, 2, \ldots,$$

and we get the fundamental solutions

$$u_n(x, t) = \sin \frac{n\pi x}{L} \sin \frac{n\pi a t}{L}, \quad n = 1, 2, \ldots$$

Using the superposition principle, we attempt to find a series solution

$$u(x, t) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{L} \sin \frac{n\pi a t}{L}.$$

For this to satisfy the initial condition $u_t(x, 0) = g(x)$, we must have

$$u_t(x, 0) = \sum_{n=1}^{\infty} \frac{n\pi a}{L} c_n \sin \frac{n\pi x}{L} = g(x).$$

Thus, the numbers $\frac{n\pi a}{L} c_n$ are the Fourier coefficients of the odd $2L$-periodic extension of g, and are determined by

$$\frac{n\pi a}{L} c_n = \frac{2}{L} \int_0^L g(x) \sin \frac{n\pi x}{L},$$

that is,

$$c_n = \frac{2}{n\pi a} \int_0^L g(x) \sin \frac{n\pi x}{L}.$$

23 Lecture 23

23.1 General case: nonzero initial displacement and velocity

Now that we know how to solve the wave propagation problem of a string with nonzero initial position and nonzero initial velocity separately, we can put these together to find the solution of the general wave propagation problem

$$\begin{aligned} u_{tt} &= a^2 u_{xx}, && 0 < x < L, \quad t > 0, \\ u(0,t) = u(L,t) &= 0, && t \geq 0, \\ u(x,0) &= f(x), && 0 \leq x \leq L, \\ u_t(x,0) &= g(x), && 0 \leq x \leq L. \end{aligned}$$

Assume that $v(x,t)$ satisfies the wave equation, the boundary conditions $v(0,t) = v(L,t) = 0, t > 0$, and the initial conditions $v(x,0) = f(x), v_t(x,0) = 0, 0 \leq x \leq L$. Assume also that $w(x,t)$ satisfies the wave equation, the same boundary conditions, and the initial conditions $w(x,0) = 0, w_t(x,0) = g(x), 0 \leq x \leq L$. Now consider the function $u(x,t) = v(x,t) + w(x,t)$. We have

$$u_{tt} = v_{tt} + w_{tt} = a^2 v_{xx} + a^2 w_{xx} = a^2 u_{xx},$$

so u satisfies the wave equation. Moreover,

$$u(0,t) = v(0,t) + w(0,t) = 0, \quad t > 0$$

and similarly $u(L,t) = 0, t > 0$, so u satisfies the boundary conditions. Finally,

$$u(x,0) = v(x,0) + w(x,0) = f(x) + 0 = f(x), \quad 0 \leq x \leq L,$$

and

$$u_t(x,0) = v_t(x,0) + w_t(x,0) = 0 + g(x) = g(x), \quad 0 \leq x \leq L,$$

so u satisfies the required initial conditions as well. Thus, u is the solution to the general wave propagation problem. It was easy to find: once we knew how to solve the two slightly simpler problems, with either zero initial displacement or zero initial velocity, we simply added the two solutions. This is a great example of the power of the principle of superposition.

23.2 Zero initial velocity: structure of the solution

Let's look back at the solution to the wave propagation problem with initial velocity:

$$\begin{aligned} u_{tt} &= a^2 u_{xx}, && 0 < x < L, \quad t > 0, \\ u(0,t) = u(L,t) &= 0, && t \geq 0, \\ u(x,0) &= f(x), && 0 \leq x \leq L, \\ u_t(x,0) &= 0, && 0 \leq x \leq L. \end{aligned}$$

We found that the solution is given by

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{L} \cos \frac{n\pi a t}{L},$$

where the coefficients

$$c_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} dx$$

are the Fourier coefficients of the odd $2L$-periodic extension of $f(x)$. Let's denote this function by $h(x)$, so that

$$h(x) = \begin{cases} f(x), & 0 \leq x \leq L \\ -f(-x), & -L < x < 0. \end{cases}$$

The Fourier series of $h(x)$ is given by

$$h(x) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{L}.$$

We now claim that the above solution $u(x,t)$ satisfies

$$u(x,t) = \frac{1}{2}\left(h(x-at) + h(x+at)\right).$$

To establish this, we use the trigonometric formulas

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta, \quad \sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta$$

to obtain

$$\begin{aligned} h(x-at) &= \sum_{n=1}^{\infty} c_n \sin \frac{n\pi(x-at)}{L} \\ &= \sum_{n=1}^{\infty} c_n \left(\sin\frac{n\pi x}{L}\cos\frac{n\pi at}{L} - \cos\frac{n\pi x}{L}\sin\frac{n\pi at}{L} \right) \end{aligned}$$

and

$$\begin{aligned} h(x+at) &= \sum_{n=1}^{\infty} c_n \sin \frac{n\pi(x+at)}{L} \\ &= \sum_{n=1}^{\infty} c_n \left(\sin\frac{n\pi x}{L}\cos\frac{n\pi at}{L} + \cos\frac{n\pi x}{L}\sin\frac{n\pi at}{L} \right). \end{aligned}$$

Adding these series term by term gives precisely $2u(x,t)$, which yields the result.

We can make two important observations from this result. Firstly, from a physical point of view, we can interpret the solution $u(x,t)$ as the superposition of two waves with shape given by the initial function $f(x)$ moving in opposite directions with velocity a. Secondly, any discontinuities in the initial data $f(x)$ or its derivatives will be preserved in the solution $u(x,t)$ for all time. This is in contrast to the solution of the heat equation, whose derivatives of all orders exist when $t > 0$.