Massachusetts Institute of Technology

Introduction to partial differential equations (PDEs)

A gentle approach

18.303 Linear Partial Differential Equations: Analysis and Numerics

Practical information

- All the important information on class website
  `https://github.com/mitmath/18303/`
- The recorded lectures can be found from the class Canvas website
- We will use the Canvas site when needed (submitting assignments, announcements etc.)
- For the computational parts we will use JULIA (Steve Johnson's tutorial on Friday at 5 pm)
- The class consists of the psets, the midterm, and a final project (I'll decide on the dates soon)

Why are PDEs important?

Law of Motion

Newton's law of motion

$$m \frac{\mathrm{d}^2 \mathsf{x}}{\mathrm{d}t^2} = \mathsf{f}$$

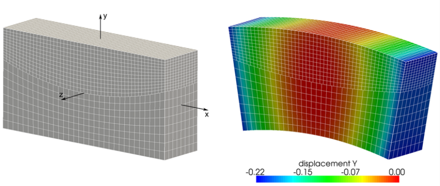
- Describes dynamics at low velocities ($\ll$ speed of light)
- Is used to derive classical dynamics resulting in various differential equations
- Can be extended to relativistic velocities (Einstein)

Hooke's law

Hooke's law

$$\mathbf{f} = -k\mathbf{x}$$

- Elementary constituents of materials are in *equilibrium* when their distance is set
- When this distance changes due to deformations, the system energy increases and stress sets in
- Hooke's law is used to derive a large family of different equations (dynamical or static) to describe elasticity



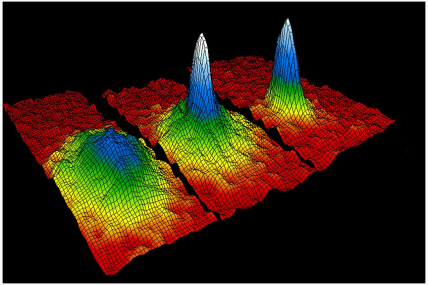
A finite element calculation of a bending beam.

Quantum mechanics

Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi(t, \mathrm{x}) + V(\mathrm{x})\psi$$

- PDEs describe time-evolution and static properties of quantum mechanics
- Different PDEs for relativistic phenomena e.g. spin degrees of freedom
- Also subatomic physics are written in terms of PDEs



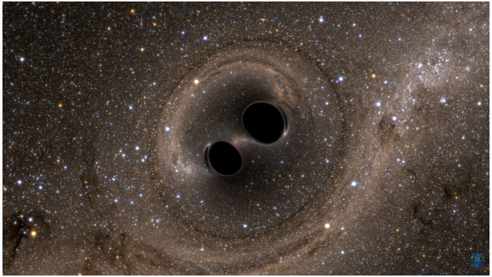
Bose-Einstein condensation of rubidium atoms. Image courtesy of NIST/JILA/CU-Boulder.

General relativity

Einstein's field equation

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

- Describes gravity's relation to mass and spacetime
- Used to derive precise trajectories for planets and galaxies
- Predicts e.g. black holes
- $G_{\mu\nu}$ and $g_{\mu\nu}$ are related to the geometry of the spacetime through differential equations (tensors are not covered in this class)



Simulation of a black hole merger event [SXS lensing].

Finance

Black-Scholes equation for the price of an option

$$\frac{\partial V}{\partial t} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} = rV$$

- Derived from assumed stochastic (random) dynamics of the stock market
- Describes the pricing of an option V as a function of time (t) and the price of the underlying asset S
- Is a sort of an *diffusion equation* (important later on)

Vectors

- During this class we use objects called vectors (abstract vectors are denoted with a bold symbol e.g. $\mathbf{f}$)
- Vectors are elements of a vector space V defined over a field F
- This field could be e.g. the set of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$

- The vectors are defined in a way that adding the vectors creates another vector
- Let $\mathbf{f}, \mathbf{g} \in V$ be vectors and $\alpha, \beta \in F$
- It follows that $\alpha\mathbf{f} + \beta\mathbf{g} \in V$

(See how this works for your usual vectors, say, in $\mathbb{R}^3$)

Vector spaces

Norm

- We define an operation $V \to \mathbb{R}_+$ called the norm denoted by $\|\cdot\|$ ($\mathbb{R}_+$ is the set of nonnegative real numbers)
- The norm has the following properties for $\mathbf{f} \in V$ and $\alpha \in F$:
  1. It is nonnegative i.e. $\|\mathbf{f}\| \geq 0$
  2. For nonzero vectors it is positive: $\|\mathbf{f}\| = 0 \Leftrightarrow \mathbf{f} = 0$
  3. $\|\alpha\mathbf{f}\| = |\alpha| \, \|\mathbf{f}\|$
  4. The triangle inequality holds (**important**): $\|\mathbf{f} + \mathbf{g}\| \leq \|\mathbf{f}\| + \|\mathbf{g}\|$

Assume $\|\mathbf{f}_k\|$ converges like a normal sequence of numbers (Cauchy). If there is $\mathbf{f} \in V$ s.t. $\lim_{k\to\infty} \|\mathbf{f}_k - \mathbf{f}\| = 0$ for all such sequences, V is called a *Banach space* and V is said to be *complete* (not important for this class but our vectors will be in a Banach space). [I'll use this color to denote optional material]

The norm defines a *metric* i.e. a notion of distance for vectors (if the norm of $\mathbf{f} - \mathbf{g}$ is zero, they're the same vector). We will introduce a notion of angle between vectors by defining an inner product:

Inner product

- Inner product is a bilinear operation $\langle \cdot \rangle : V \times V \to F$ with the following properties
  1. Linearity: $\langle \mathbf{h}, \alpha \mathbf{f} + \beta \mathbf{g} \rangle = \alpha \langle \mathbf{h}, \mathbf{f} \rangle + \beta \langle \mathbf{h}, \mathbf{g} \rangle$
  2. Conjugate symmetry: $\langle \mathbf{f}, \mathbf{g} \rangle = \overline{\langle \mathbf{g}, \mathbf{f} \rangle}$ ($\overline{\alpha}$ is the complex conjugate of α)
  3. Positive-definiteness: $\langle \mathbf{f}, \mathbf{f} \rangle > 0$ if $\mathbf{f} \neq 0$

For inner products $\langle \mathbf{f}, \mathbf{f} \rangle = \|\mathbf{f}\|^2$ i.e. it induces a norm on the vector space.

Banach spaces with an inner product are called *Hilbert spaces* (very important mathematical structure in quantum mechanics).

Exercise 1

Inner product induces a norm

Show that $\sqrt{\langle \mathbf{f}, \mathbf{f} \rangle} = \|\mathbf{f}\|$ is actually a norm

Solution

The first three properties of the norm follow pretty easily. Let's show the triangle inequality i.e. $\|\mathbf{f}+\mathbf{g}\| \leq \|\mathbf{f}\| + \|\mathbf{g}\|$. We assume $\mathbf{f}, \mathbf{g} \neq 0$ (these cases work trivially).

Since both sides are positive, this is equivalent to $\|\mathbf{f}+\mathbf{g}\|^2 \leq \|\mathbf{f}\|^2 + \|\mathbf{g}\|^2 + 2\|\mathbf{f}\|\|\mathbf{g}\|$.

We have $\|\mathbf{f}+\mathbf{g}\|^2 = \langle \mathbf{f}+\mathbf{g}, \mathbf{f}+\mathbf{g}\rangle = \langle \mathbf{f}, \mathbf{f}\rangle + \langle \mathbf{g}, \mathbf{g}\rangle + \langle \mathbf{f}, \mathbf{g}\rangle + \langle \mathbf{g}, \mathbf{f}\rangle = \|\mathbf{f}\|^2 + \|\mathbf{g}\|^2 + 2\operatorname{Re}(\langle \mathbf{f}, \mathbf{g}\rangle)$.

Now, it suffices to show that $\operatorname{Re}(\langle \mathbf{f}, \mathbf{g}\rangle) \leq \|\mathbf{f}\|\|\mathbf{g}\|$ i.e. $\operatorname{Re}(\langle \mathbf{f}/\|\mathbf{f}\|, \mathbf{g}/\|\mathbf{g}\|\rangle) =: \operatorname{Re}\left(\langle \hat{\mathbf{f}}, \hat{\mathbf{g}}\rangle\right) \leq 1$.

Let $\hat{\mathbf{g}} = \hat{\mathbf{f}} + \mathbf{h}$. Now we have $\operatorname{Re}\left(\langle \hat{\mathbf{f}}, \hat{\mathbf{g}}\rangle\right) = \operatorname{Re}\left(\underbrace{\langle \hat{\mathbf{f}}, \hat{\mathbf{f}}\rangle}_{=1}\right) + \operatorname{Re}\left(\langle \hat{\mathbf{f}}, \mathbf{h}\rangle\right)$. We see that it's enough to show that $\operatorname{Re}\left(\langle \hat{\mathbf{f}}, \mathbf{h}\rangle\right) \leq 0$.

We know that $1 = \langle \hat{\mathbf{g}}, \hat{\mathbf{g}}\rangle = \langle \hat{\mathbf{f}}+\mathbf{h}, \hat{\mathbf{f}}+\mathbf{h}\rangle = \underbrace{\left\|\hat{\mathbf{f}}\right\|^2}_{=1} + \|\mathbf{h}\|^2 + 2\operatorname{Re}\left(\langle \hat{\mathbf{f}}, \mathbf{h}\rangle\right)$. It follows that $\operatorname{Re}\left(\langle \hat{\mathbf{f}}, \mathbf{h}\rangle\right) = -\|\mathbf{h}\|^2/2 \leq 0$, which completes the proof.

Linear operators

- Linear operator on a vector space $\mathcal{L}(\cdot) : V \to V$ maps vectors to vectors
- We write $\mathcal{L}(\mathbf{f}) = \mathcal{L}\mathbf{f}$
- They are linear i.e. $\mathcal{L}(\alpha\mathbf{f} + \beta\mathbf{g}) = \alpha\mathcal{L}\mathbf{f} + \beta\mathcal{L}\mathbf{g}$
- Linearity of operators: for two operators $\mathcal{L}$ and $\mathcal{G}$ we have $(\alpha\mathcal{L} + \beta\mathcal{G})\,\mathbf{f} = \alpha\mathcal{L}\mathbf{f} + \beta\mathcal{G}\mathbf{g}$

Null space and range

- Null space (kernel) N of an operator $\mathcal{L}$ is the set $\{\mathbf{f} \in V : \mathcal{L}\mathbf{f} = 0\}$
- Range R of an operator $\mathcal{L}$ is the set $\{\mathbf{f} \in V : \mathcal{L}\mathbf{g} = \mathbf{f} \quad \text{for some } \mathbf{g} \in V\}$

Exercise 2

Let $\mathbf{f}, \mathbf{g} \in V$. What can you say about them if for a given operator $\mathcal{L}$ we have $\mathcal{L}\mathbf{f} = \mathcal{L}\mathbf{g}$? What if the null space of $\mathcal{L}$ is $\{0\}$?

Eigenvalues and eigenvectors

- If $\mathcal{L}\mathbf{f} = \lambda\mathbf{f}$, we say that $\mathbf{f}$ is an eigenvector of the operator $\mathcal{L}$ with an eigenvalue λ

Adjoints

- The adjoint of an operator $\mathcal{L}$, $\mathcal{L}^*$ is defined through the property $\langle \mathbf{f}, \mathcal{L}\mathbf{g} \rangle = \langle \mathcal{L}^*\mathbf{f}, \mathbf{g} \rangle$ for all $\mathbf{f}, \mathbf{g} \in V$
- If $\mathcal{L} = \mathcal{L}^*$, the operator is called self-adjoint (think of symmetric (or Hermitian) matrices).

Example 1

Smooth functions

- Let us define the vector space as a space of functions $f : [0, 1] \to \mathbb{R}$ such that arbitrarily high degrees of derivatives are continuous (we write $f \in C^\infty$)
- We can define an inner product of f and g by $\langle f, g \rangle = \int_0^1 f(x)g(x)\mathrm{d}x$
- If in addition we require that $f(0) = f(1) = 0$, we write $f \in C_0^\infty$
- Functions in C_0^∞ form an important category of functions called the *test functions*
- Examples of linear operations for these vectors:
    - The differentiation operator $\frac{\mathrm{d}^n}{\mathrm{d}x^n}$ for any integer n
    - The integral $(\mathcal{I}f)(x) := \int_0^x f(x')\mathrm{d}x'$

In addition to the number of times functions can be differentiated, many times we also need to care if the integral $\int_0^1 |f|^p \mathrm{d}x$ is finite. If this is the case we write $f \in L^p([0, 1])$.

First differential equation

Poisson's equation with Dirichlet boundaries

$$\frac{\partial^2 u(x)}{\partial x^2} = f(x),$$
$$u(0) = u(1) = 0.$$

First differential equation

Poisson's equation with Dirichlet boundaries

$$
\begin{aligned}
\frac{\partial^2 u(x)}{\partial x^2} &= f(x), \\
u(0) = u(1) &= 0.
\end{aligned}
$$

First, we solve the eigenvalue problem

$$
\begin{aligned}
\frac{\partial^2 \phi_n(x)}{\partial x^2} &= -\lambda_n^2 \phi(x), \\
u(0) = u(1) &= 0.
\end{aligned}
$$

First differential equation

Poisson's equation with Dirichlet boundaries

$$\frac{\partial^2 u(x)}{\partial x^2} = f(x),$$
$$u(0) = u(1) = 0.$$

First, we solve the eigenvalue problem

$$\frac{\partial^2 \phi_n(x)}{\partial x^2} = -\lambda_n^2 \phi(x),$$
$$u(0) = u(1) = 0.$$

The eigenfunctions (eigenvectors) of the differential operator ∂_x^2 are exponential functions $e^{\pm i\lambda_n x}$.

First differential equation

The eigenfunctions have to respect the boundary conditions. We have $\phi_n(0) = 0$. How is this possible with the exponential functions?

First differential equation

The eigenfunctions have to respect the boundary conditions. We have $\phi_n(0) = 0$. How is this possible with the exponential functions?

Since we have two eigenfunctions ($e^{\pm i\lambda_n x}$), the linear combination is also an eigenfunction. Now $\phi_n = \alpha e^{i\lambda_n x} + \beta e^{-i\lambda_n x}$ for some complex α and β.

First differential equation

The eigenfunctions have to respect the boundary conditions. We have $\phi_n(0) = 0$. How is this possible with the exponential functions?

Since we have two eigenfunctions ($e^{\pm i\lambda_n x}$), the linear combination is also an eigenfunction. Now $\phi_n = \alpha e^{i\lambda_n x} + \beta e^{-i\lambda_n x}$ for some complex α and β.

Solving for $\phi_n(0) = 0$ gives $\beta = -\alpha$ and requiring that ϕ_n is real gives $\alpha = A/(2i)$ with some real A. For now we can set A to 1 – we just have to keep in mind that multiplying the eigenvector by a constant is also a solution. Now, $\phi_n(x) = \sin(\lambda_n x)$.

First differential equation

The eigenfunctions have to respect the boundary conditions. We have $\phi_n(0) = 0$. How is this possible with the exponential functions?

Since we have two eigenfunctions ($e^{\pm i\lambda_n x}$), the linear combination is also an eigenfunction. Now $\phi_n = \alpha e^{i\lambda_n x} + \beta e^{-i\lambda_n x}$ for some complex α and β.

Solving for $\phi_n(0) = 0$ gives $\beta = -\alpha$ and requiring that ϕ_n is real gives $\alpha = A/(2i)$ with some real A. For now we can set A to 1 – we just have to keep in mind that multiplying the eigenvector by a constant is also a solution. Now, $\phi_n(x) = \sin(\lambda_n x)$.

The other boundary condition ($\phi_n(1) = 0$) gives $\sin(\lambda_n) = 0$. This is solved by $\lambda_n = \pi n$ for any integer n. Since $\sin(-x) = -\sin(x)$ (a constant times $\sin(x)$) and $\sin(0) = 0$, it suffices to have $n = 1, 2, 3, \ldots$

First differential equation

The eigenfunctions have to respect the boundary conditions. We have $\phi_n(0) = 0$. How is this possible with the exponential functions?

Since we have two eigenfunctions ($e^{\pm i\lambda_n x}$), the linear combination is also an eigenfunction. Now $\phi_n = \alpha e^{i\lambda_n x} + \beta e^{-i\lambda_n x}$ for some complex α and β.

Solving for $\phi_n(0) = 0$ gives $\beta = -\alpha$ and requiring that ϕ_n is real gives $\alpha = A/(2i)$ with some real A. For now we can set A to 1 – we just have to keep in mind that multiplying the eigenvector by a constant is also a solution. Now, $\phi_n(x) = \sin(\lambda_n x)$.

The other boundary condition ($\phi_n(1) = 0$) gives $\sin(\lambda_n) = 0$. This is solved by $\lambda_n = \pi n$ for any integer n. Since $\sin(-x) = -\sin(x)$ (a constant times $\sin(x)$) and $\sin(0) = 0$, it suffices to have $n = 1, 2, 3, \ldots$

It turns out that ϕ_n is a basis for functions $f : [0, 1] \to \mathbb{R}$. It was finally proven in 60's (Carleson (1966) & Hunt (1968)) that any functions f (even ones that are not continuous) can be expressed in this basis *iff* $\int_0^1 |f(x)|^p \mathrm{d}x < \infty$ for some $p > 1$. To be precise, the sine series converges almost everywhere to f with this condition (not at isolated points).

Basis

- A set of basis vectors $\{\phi_n\}_{n=1}^{\infty}$ is a basis for the vector space V if the following properties hold:
  1. **Linear independence**: $\sum_{n=1}^{\infty} \alpha_n \phi_n = 0 \Leftrightarrow \alpha_n = 0$ for all n
  2. **Spanning property**: any vector $\mathbf{f} \in V$ can be written as a linear combination of the basis vectors i.e. $\mathbf{f} = \sum_{n=1}^{\infty} \alpha_n \phi_n$ for some $\{\alpha_n\}$
- The basis is said to be orthogonal *iff* $\langle \phi_i, \phi_j \rangle = \beta_i \delta_{ij}$ (δ_{ij} is called Kronecker delta; it's 1 if $i = j$ and 0 otherwise)
- If constants $\beta_i = 1$ for all i, the basis is called orthonormal

Exercise 3

Show that the basis $\phi_n(x) = \sin(\pi n x)$ is orthogonal on the interval $x \in [0, 1]$.

Is it orthonormal?

If not, how could you make it orthonormal?

First differential equation

Poisson's equation with Dirichlet boundaries

$$\frac{\partial^2 u(x)}{\partial x^2} = f(x),$$

$$u(0) = u(1) = 0.$$

First differential equation

We write both u and f in the eigenbasis ϕ_n. We have $u(x) = \sum_{n=1}^{\infty} \hat{u}_n \phi_n$ and $f(x) = \sum_{n=1}^{\infty} \hat{f}_n \phi_n$.

First differential equation

We write both u and f in the eigenbasis ϕ_n. We have $u(x) = \sum_{n=1}^{\infty} \hat{u}_n \phi_n$ and $f(x) = \sum_{n=1}^{\infty} \hat{f}_n \phi_n$.

Inserting this in the differential equation gives

$$\frac{\partial^2}{\partial x^2} \sum_{n=1}^{\infty} \hat{u}_n \phi_n = \sum_{n=1}^{\infty} \hat{u}_n \frac{\partial^2}{\partial x^2} \phi_n = -\sum_{n=1}^{\infty} \hat{u}_n \lambda_n^2 \phi_n = \sum_{n=1}^{\infty} \hat{f}_n \phi_n.$$

First differential equation

We write both u and f in the eigenbasis ϕ_n. We have $u(x) = \sum_{n=1}^{\infty} \hat{u}_n \phi_n$ and $f(x) = \sum_{n=1}^{\infty} \hat{f}_n \phi_n$.

Inserting this in the differential equation gives

$$\frac{\partial^2}{\partial x^2} \sum_{n=1}^{\infty} \hat{u}_n \phi_n = \sum_{n=1}^{\infty} \hat{u}_n \frac{\partial^2}{\partial x^2} \phi_n = -\sum_{n=1}^{\infty} \hat{u}_n \lambda_n^2 \phi_n = \sum_{n=1}^{\infty} \hat{f}_n \phi_n.$$

We can take the inner product of both sides of the equation with $\langle \phi_k, \cdot \rangle$. Since ϕ_n are orthogonal we get $-\hat{u}_k \lambda_k^2 = \hat{f}_k$ giving

$$\hat{u}_k = -\frac{\hat{f}_k}{\lambda_k^2} = -\frac{\hat{f}_k}{\pi^2 k^2},$$

where $k = 1, 2, 3, \ldots$

First differential equation

How do we calculate the coefficients $\hat{f}_n$?

First differential equation

How do we calculate the coefficients $\hat{f}_n$?

We found out earlier that $\langle \phi_i, \phi_j \rangle = \frac{1}{2}\delta_{ij}$ and we have

$$\sum_{n=1}^{\infty} \hat{f}_n \phi_n = f(x).$$

Taking the product $\langle \phi_k, \cdot \rangle$ gives

$$\frac{1}{2}\hat{f}_k = \int_0^1 \phi_k(x) f(x) \mathrm{d}x = \int_0^1 \sin(k\pi x) f(x) \mathrm{d}x$$

First differential equation

How do we calculate the coefficients $\hat{f}_n$?

We found out earlier that $\langle \phi_i, \phi_j \rangle = \frac{1}{2}\delta_{ij}$ and we have

$$\sum_{n=1}^{\infty} \hat{f}_n \phi_n = f(x).$$

Taking the product $\langle \phi_k, \cdot \rangle$ gives

$$\frac{1}{2}\hat{f}_k = \int_0^1 \phi_k(x) f(x) \mathrm{d}x = \int_0^1 \sin(k\pi x) f(x) \mathrm{d}x$$

Solving this gives a general formula for the sine series coefficients

$$\hat{f}_k = 2\int_0^1 \sin(k\pi x) f(x) \mathrm{d}x. \tag{1}$$

The operation for calculating the sine series coefficients can be seen as a Fourier transform. We write $\hat{f} = \mathcal{F}(f)$, where

$$(\mathcal{F}f)_k := 2\int_0^1 \sin(k\pi x) f(x)\mathrm{d}x. \tag{2}$$

Now $\mathcal{F} : V \to W$, where W is the vector space of coefficients $\hat{f}_n$ is a linear map (check for yourself). It also has an inverse $\mathcal{F}^{-1}$ defined through

$$\mathcal{F}^{-1}(\hat{f})(x) = \sum_{n=1}^{\infty} \hat{f}_n \sin(n\pi x)$$

i.e. $\mathcal{F}$ is a bijection between spaces V and W (for mathematical nitpicking we require that V are the vectors for which Fourier transform exists).