Introduction

Opening remarks

All models are wrong, only some models are useful.

Slides of the presentation: These summarise the following paragraphs and the content of the course in a short presentation. todo: add slide link here

Constructing suitable **approximations to model the world** around us sits at the **heart of all sciences**. As the above famous quote implies, modelling a scientific problem inevitable comes with the introduction of errors. Even the most accurate models of physics, e.g. the standard model of particle physics or relativistic quantum mechanics have their limitations. Moreover these models are typically too involved to be applied to all but the simplest systems. In particular for "real" systems of macroscopic size additional approximations. Typically this not only involves the selection of **more approximate**, thus simpler, **physical models**, but also employing **numerical** instead of an analytical techniques for solving the underlying equations.

Which physical effects can be neglected and which numerical techniques are most appropriate is often application-specific and **many alternatives exist**. Understanding the **dominant sources of error** in a simulation is thus a crucial component in understanding whether each respective **approximation choice** has been appropriate and —- most importantly —- which of them **needs to be refined** in order to improve the result.

The goal of this lecture is to discuss mathematical techniques which aid with the estimation of error in scientific simulations. Within the lecture we will distinguish various forms of error and consider (some) of them in separate discussions throughout the semester. For example we will distinguish the **model error**, i.e. the error introduced the selected physical model compared to the ground truth, as well as the **numerical error**, i.e. all additional error from the numerical procedure employed for solving the chosen physical model. Of particular focus in the lecture will be understanding the various aspects of the numerical error, that is in particular the errors due to

the chosen **discretisation basis**, the **algorithms** to solve the numerical problem as well as the chosen **floating-point arithmetic**.

Looking back into the **history of scientific modelling** a number of incidents with **hundreds of millions of dollars of damage** could have been **prevented** if better numerical error control measures would have been implemented. One example is the sinking of the Sleipner A oil rig platform due to an insufficient finite element discretisation. A second example is the self-destruction of an Ariane 5 rocket in its test flight due to an error in the performed floating-point arithmetic. In this lecture we will discuss techniques how such errors could have been prevented. Some guiding questions for our lecture are:

- What can go wrong in a simulation?
- How can we detect this and what can we do about it?
- If we know errors how can this be used to make simulations faster and more reliable?

Eigenvalue problems in scientific modelling

Eigenvalue problems occur naturally in many domains of science. In physics and engineering problems they are frequently related to vibrations, e.g. bridges or buildings can swing under wind load. A famous example is the Tacoma narrows bridge, which collapsed in 1940 just half a year after its opening due to strong winds bringing the bridge into strong vibrations. Similarly the London millennium foot bridge had to be closed due to its strong wobbeling by the passing pedestrians.

In my own field, atomistic modelling, the underlying workhorse theory is the Schrödinger equation. In quantum mechanics the eigenpairs (states) of an operator describing the physical system under study form the key quantity of interest in simulations.

Moreover, as we will discuss in the next examples in more detail, there is a natural connections from solving partial differential equations~(PDE) to eigenvalue problems involving the PDE operator.

Another natural connection arises when analysing the convergence of iterative numerical schemes, such as multidimensional fixed-point problems, iterative schemes like Newton or the conjugate-gradient algorithm. The convergence of CG, for example, is related to the condition number of the underlying linear system, which in turn are related to singular values of the system matrix —- an Quantity, which can be obtained by solving an eigenvalue problem.

The connection between linear and eigenvalue problems : the heat equation

Let $\Omega \subset \mathbb{R}^n$ with a Lipshitz (smooth) boundary. We seek $u : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ such that

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u = 0 \qquad & \text{in } \Omega \times \mathbb{R}^+ && \text{(Heat equation)} \\ u(\cdot, 0) = u_0 \qquad & \text{in } \Omega && \text{(Initial conditions)} \\ u = 0 \qquad & \text{on } \partial\Omega \times \mathbb{R}^+ && \text{(Dirichlet b.c.)} \end{aligned}$$

Where the initial conditions $u_0 : \Omega \to \mathbb{R}$ are of appropriate regularity.

To solve this problem, we perform separation of variables

$$u(x,t) = \sum_{n=1}^{\infty} c_n(t) v_n(x).$$

We will see later that this postulate is justified.

To start, let's just consider one term of the sum, i.e. $u(x,t) = c(t)v(x)$.

- From the heat equation

$$c'(t)v(x) - c(t)\Delta v(x) = 0$$

- Assuming that $v(x)$ and $c(t)$ are non-zero, we can divide

$$\frac{c'(t)}{c(t)} = \frac{\Delta v(x)}{v(x)} \qquad \forall\, t > 0, x \in \Omega.$$

- The left only depends on t, and the right only on x. As a result, we need

$$\frac{c'(t)}{c(t)} = \frac{\Delta v(x)}{v(x)} = \mu \in \mathbb{R}$$

  for some μ.

- Consider the time-dependant part (c)

$$\begin{aligned} c'(t) &= \mu c(t) \\ \Rightarrow c(t) &= \alpha \exp(\mu t). \end{aligned}$$

- Separately, consider the space-dependant part (v)

$$\begin{aligned} \Delta v &= \mu v \qquad && \text{in } \Omega \\ v &= 0 \qquad && \text{on } \partial\Omega \end{aligned}$$

which is the Dirchlet-Laplace eigenvalue problem !

- One can show that the operator $-\Delta$ with homogeneous Dirichlet boundary conditions has a sequence of eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \ldots \to \infty$ and orthonormal eigenfunctions $\{\phi_n\}_{n=1}^{\infty}$, i.e.

$$-\Delta \phi_n = \lambda_n \phi_n$$

with

$$\langle \phi_n, \phi_m \rangle = \int_\Omega \phi_n(x)\phi_m(x)dx = \delta_{mn}.$$

- These form a complete basis of $H_0^1(\Omega)$, the relevant Hilbert space for this problem.

This justifies the ansatz

$$u(x,t) = \sum_{n=1}^{\infty} c_n(t)\phi_n(x)$$

we made before, where ϕ_n are now eigenfunctions of the Laplace operator.

To fully translate the problem to the eigenfunction setting, we need to consider the initial condition u_0.

- On the one hand, in the eigenbasis,

$$u_0 = \sum_{i=1}^{\infty} \underbrace{\langle \phi_i, u_0 \rangle}_{u_{0,i}} \phi_i.$$

- On the other hand,

$$c_i(t) = \alpha_i \exp(-\lambda_i t).$$

which yields

$$u(x,t) = \sum_{i=1}^{\infty} \alpha_i \exp(-\lambda_i t)\phi_i(x).$$

- Taking $t \to 0$, we obtain

$$\sum_{i=1}^{\infty} u_{0,i}\phi_i(x) = u_0(x) = u(0,x) = \sum_{i=1}^{\infty} \alpha_i \phi_i(x).$$

- Since ϕ_i is a basis, we deduce $\alpha_i = u_{0,i}$ and finally obtain

$$u(x,t) = \sum_{i=1}^{\infty} u_{0,i} \exp(-\lambda_i t)\phi_i(x)$$

We observe the following :

- The solution decays to $u_\infty = 0$ as $t \to \infty$.
- The decay rate depends on λ_i.

More importantly for this course, this shows that eigenvalue problems, which we will focus on, are key in understanding linear problems, even though they might seem unrelated at a first glance. One of these linear problems of particular relevance is solving Schrödinger equation, which we treat in the next section.

The Schrödinger equation and quantum mechanics

At the microscopic level, the physics and chemistry of materials is governed by the interaction of electrons and nuclei. At this scale, the regime of quantum mechanics applies.

In quantum mechanics the state of a system is described the complex-valued square-integrable wave function $\Psi : (x,t) \mapsto \Psi(x,t) \in \mathbb{C}$, where x corresponds to the degrees of freedom (e.g. position or spin) of the particles (e.g. electrons) in the system.

For simplicity, take $x \in \mathbb{R}^3$, i.e. one particle in 3D. Part of the meaning of Ψ appears through its modulus squared. Indeed, for all t, $|\Psi(x,t)|^2$ corresponds to the probability distribution of finding the particle at position x at time t.

Further information appears through its Fourier transform, as $|\hat{\Psi}(p,t)|^2$ corresponds to the probability distribution of finding the particle at momentum p at time t, where the Fourier transform is given by

$$\hat{\Psi}(p,t) = \int_{\mathbb{R}^3} \Psi(x,t) e^{-2\pi i\, p\cdot x} dx.$$

For those already familiar with quantum mechanics, note that we set $\hbar = 1$.

In analogy to classical mechanics, one may obtain the dynamics of the particle by investigating its total energy (considering the mass of the particle to be 1).

- The total energy is given by

$$\begin{aligned} E(t) &= \textbf{kinetic energy} + \textbf{potential energy} \\ &= \int_{\mathbb{R}^3} \frac{|p|^2}{2} |\hat{\Psi}(p,t)|^2 dp + \int_{\mathbb{R}^3} V(x) |\Psi(x,t)|^2 dx \end{aligned}$$

  where $V : \mathbb{R}^2 \to \mathbb{R}$ is the potential inducing the particle dynamics.
- From the Fourier identity $p\hat{\Psi}(p) = -i\widehat{(\nabla \Psi)}(p)$, we can develop the first term into

$$\begin{aligned} \int_{\mathbb{R}^3} |p|^2 |\hat{\Psi}(p,t)|^2 dp &= \int_{\mathbb{R}^3} |p\hat{\Psi}(p,t)|^2 dp \\ &= \int_{\mathbb{R}^3} |-i\widehat{(\nabla \Psi)}(p,t)|^2 dp \\ &= \int_{\mathbb{R}^3} |\widehat{(\nabla \Psi)}(p,t)|^2 dp \\ &\overset{\text{Parceval}}{=} \int_{\mathbb{R}^3} |\nabla \Psi(x,t)|^2 dx \end{aligned}$$

  where we used Parceval's theorem to go from the third to the fourth line.

- With this we obtain,

$$\begin{aligned} E(t) &= \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \Psi(x,t)|^2 dx + \int_{\mathbb{R}^3} V(x) |\Psi(x,t)|^2 dx \\ &\overset{IBP}{=} -\frac{1}{2} \int_{\mathbb{R}^3} \overline{\Psi(x,t)} \Delta \Psi(x,t) dx + \int_{\mathbb{R}^3} \overline{\Psi(x,t)} V(x) \Psi(x,t) dx \\ &= \int_{\mathbb{R}^3} \overline{\Psi(x,t)}\, \mathcal{H}\, \Psi(x,t) dx \\ &= \langle \Psi, \mathcal{H}\Psi \rangle_{L^2(\mathbb{R}^3)} \end{aligned}$$

  Where we used integration by parts to go from the first to the second line.
- In this, the *Hamiltonian*

$$\mathcal{H} := -\frac{1}{2}\Delta + V(x)$$

gives the total energy of the system when applied to the wavefunction Ψ, where
- $-\Delta/2$ is the kinetic energy operator
- $V(x)$ is the potential energy, applied point-wise in real space.

The dynamics of the system are now given by the time-dependant Schrödinger equation (TDSE)

$$i\frac{\partial \Psi}{\partial t} = \mathcal{H}\Psi \qquad \textbf{(TDSE)}$$

In general, solving this equation is quite complicated. However, since it is linear, we know that superpositions of multiple solutions are also a solution. This motivates a separation of variables similar to the heat equation.

- Inserting $\Psi(x,t) = c(t)\varphi(x)$ into the TDSE yields

$$i\frac{c'(t)}{c(t)} = \frac{\mathcal{H}\varphi(x)}{\varphi(x)} = E = \text{cst}$$

- For the time-dependence we obtain the ODE

$$\begin{aligned} c'(t) &= -iEc(t) \\ \Rightarrow c(t) &= \exp(-iEt) \end{aligned}$$

- The space dependence yields the time-independent Schrödinger equation (TISE)

$$\mathcal{H}\varphi(x) = E\varphi(x) \qquad \textbf{(TISE)}$$

- As before, if we are able to find (orthonormal) eigenpairs (E_n, ϕ_n), then the TDSE can be solved as

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n e^{-iE_n t}\phi_n(x)$$

  with $c_n = \langle \phi_n, \Psi \rangle$.
- Of key importance in quantum mechanics in particular is the computation of the eigenpair corresponding to the smallest eigenvalue of $\mathcal{H}$, called the *ground state*.

Table of Contents