Numerical Solutions to PDEs

Lecture Notes #16
— Analysis of Well-Posed and Stable Problems —
A Quick Overview

Peter Blomgren,
⟨blomgren.peter@gmail.com⟩

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

http://terminus.sdsu.edu/

Spring 2018

Outline

Introduction

In the next ≈ 3 lectures we will cover the **high-lights** of chapters 9–11: **"Analysis of Well-Posed and Stable Problems"**, **"Convergence Estimates for Initial Value Problems"**, and **"Well-Posed and Stable Initial-Boundary Value Problems."**

The purpose is to showcase some of the theoretical results and tools which may be useful to a computational scientist, *without* delving into all the finer details of every proof...

We start out with **well-posedness**, a key concept in scientific modeling and the understanding of finite difference schemes used in computations.

Many of the ideas go back to Jacques S. Hadamard (1865–1963), and make plenty use of Fourier (von Neumann) analysis. The culmination of our discussion of well-posedness is the statement of the **Kreiss matrix theorem**.

Well-Posed Initial Value Problems

Some equations, *e.g.*

$$\textbf{Wave Eqn.:} \quad u_{tt} - a^2 u_{xx} = 0, \qquad \textbf{Heat Eqn.:} \quad u_t = b u_{xx},$$

and variants thereof, arise frequently in applied mathematics, but other equations, such as

$$u_{tt} = u_x,$$

do not show up as governing equations of physical systems. It is natural to ask **why?!**

In order to be a useful model of a well-behaved physical process, a PDE must have several properties, one of which is that the solution should depend on initial (boundary) data in a continuous way, so that small errors due to physical experimentation and numerical representation do not overwhelm the solution; here the definition of **"small"** must be reasonable ($\|u_{xxx}\| \leq \epsilon$ is usually not...)

A Note on Higher-Order Time Derivatives

Spatial derivatives up to 4th order are quite common (*e.g.* Beam Equation(s)).

It is quite rare (we have to venture outside of classical mechanics) to see time-derivatives beyond 2nd order; however we can give useful interpretations up to order 4:

- $\vec{x}$ — Position
- $\frac{\partial}{\partial t}\vec{x}$ — Velocity
- $\frac{\partial^2}{\partial t^2}\vec{x}$ — Acceleration
- $\frac{\partial^3}{\partial t^3}\vec{x}$ — Jerk (Jolt)
- $\frac{\partial^4}{\partial t^4}\vec{x}$ — Snap (Jounce)

The Jerk shows up in the description of the Abraham–Lorentz force (electromagnetism), which appears in the context of Wheeler–Feynman absorber theory (an interpretation of electrodynamics derived from the assumption that the solutions of the electromagnetic field equations must be invariant under time-reversal transformation, as are the field equations themselves.)

Wikipedia has some interesting rabbit-holes to explore...

The Continuity Condition

Here we are concerned with **linear problems** (the story for non-linear problems is quite different), the continuity condition is satisfied if the solution to the PDE satisfies

$$\|u(t,\circ)\| \leq C_T \|u(0,\circ)\|, \quad t \leq T,$$

measured in some norm, *i.e.* L^p, $W^{k,p}$, $H^k = W^{k,2}$ ($L^2 = H^0 = W^{0,2}$), where C_T is a constant independent of the solution.

If we have two solutions $v(t,x)$, and $w(t,x)$, then by the linearity

$$\|v(t,\circ) - w(t,\circ)\| \leq C_T \|v(0,\circ) - w(0,\circ)\|,$$

which shows that small changes in initial data results in small (bounded by a multiplicative constant) changes in the solution at time $t \leq T$.

Well-Posedness of the IVP — Definition

Definition (Well-Posedness of the IVP)

The initial value problem for a first-order equation is well-posed if for each positive T there is a constant C_T such that the inequality

$$\|u(t,\circ)\| \leq C_T \|u(0,\circ)\|,$$

holds for all initial data $u(0,x)$.

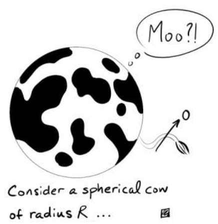
Generally, we use the L^2-norm in the estimate: — This allows us to use Fourier analysis to get sufficient and necessary conditions for the IVP to be well-posed.

For L^p ($p \neq 2$) norms, there is no relation like **Parseval's relation** for the L^2-norm, which makes the analysis harder; *e.g.* with the L^1 and L^∞-norms it is usually possible to get **sufficient or necessary** conditions, but **not** (**sufficient and necessary**) conditions.

Robustness — Lower Order Terms — 1 of 2

Another important property for a PDE to be relevant model of a physical process is that the **qualitative** (overall / general) behavior of the solution is largely unaffected by the addition of, or changes in, lower order terms.

This **robustness** condition is not always met, but is highly desirable. — Almost all derivations of equations which are meant to model physical processes make certain **assumptions**, *e.g.* **"assume a spherical cow"**, **"assume that the temperature of the body is constant"**, **"we may ignore gravitational forces"**, **"consider a homogeneous body"**, etc. etc. etc.



These assumptions really only work when small deviations in said quantities, *i.e.* the non-sphericalness of a cow, may be ignored without impacting the analysis.

Robustness — Lower Order Terms

Robustness is also important in the view of numerical solutions, since errors introduced by finite differencing, floating point computations, and/or measured (or simulated) initial data may be viewed as perturbations to, or addition of, lower order terms.

For non-robust equations, greater care must be taken when devising numerical schemes.

For now, we restrict our discussion to linear PDEs with **constant coefficients**, with **one** time-derivative, *e.g.*

$$\begin{array}{ll} u_t + a u_x = 0 & u_t - b u_{xx} + a u_x = 0 \\ u_t - c u_{txx} + b u_{xxxx} = 0 & u_t + c u_{tx} + a u_x = 0 \\ u_t = b u_{xx} & \end{array}$$

SAN DIEGO STATE UNIVERSITY

First Order (Time) PDEs

Any linear equation of first order (time) can, with the help of the Fourier transform, be written in the form

$$\widehat{u}_t(t,\omega) = q(\omega)\widehat{u}(t,\omega)$$

which gives the solution to the initial value problem

$$\widehat{u}(t,\omega) = e^{q(\omega)t}\widehat{u}_0(\omega)$$

in the Fourier domain.

With this notation, we can formalize what is required for well-posedness for these problems:

SAN DIEGO STATE UNIVERSITY

Well-posedness for First Order (Time) PDEs

Theorem (Well-Posedness for First Order PDEs)

The necessary and sufficient condition for

$$\widehat{u}_t(t,\omega) = q(\omega)\widehat{u}(t,\omega),$$

to be well-posed, that is, to satisfy the estimate

$$\|u(t,\circ)\| \le C_T \|u(0,\circ)\|,$$

is that there is a constant $\overline{q}$ such that

$$\mathrm{Re}\big(q(\omega)\big) \le \overline{q},$$

for all real values of ω.

If the theorem does not hold, then small errors of high frequency $|\omega|$ can dominate the true solution.

SAN DIEGO STATE UNIVERSITY

Well- and Ill-Posed Equations

The following equations are robust

$$\begin{array}{rlrl} u_t & = a u_x + c u & u_t & = u_{xx} + c u_x \\ u_t - u_{txx} & = a u_x + c u & u_t + u_{tx} & = b u_{xx} + c u_x \end{array}$$

for all values of c. Notice that $\overline{q}$ may depend on c, but not on ω.

The following equation

$$u_t = u_{xxx} + c u_{xx}$$

satisfies the well-posedness condition $\mathrm{Re}\big(q(\omega)\big) \le \overline{q}$ for non-negative values of c, but not if c is negative. Hence this equation is not robust when $c = 0$, since a small perturbation may send it in the wrong direction.

SAN DIEGO STATE UNIVERSITY

Higher Order Equations

The Return of the Symbol

When we have more than one time-derivative in the PDE, the **symbol** $p(s, \omega)$ is a polynomial in s. If the roots of the symbol are $\{q_1(\omega), q_2(\omega), \ldots, q_r(\omega)\}$ then any function of the form

$$e^{q_\nu(\omega)t} e^{i\omega x} \Psi(\omega)$$

is a solution of the PDE.

A **necessary condition** for well-posedness is that all roots satisfy

$$\mathrm{Re}\big(q_\nu(\omega)\big) \leq \overline{q}$$

for some $\overline{q} \in \mathbb{R}$. For second-order equations this is also **sufficient**.

We restrict our discussion of higher-order equations to some typical cases rather than develop a full theory for well-posedness...

Briefly Returning to the Question of Well-Posedness of $u_{tt} = u_x$

We can now answer why the equation

$$u_{tt} = u_x$$

does not show up as a useful model for any well-behaved physical process.

The corresponding symbol is

$$p(s, \omega) = s^2 - i\omega,$$

which has the roots

$$q_\pm(\omega) = \pm \frac{1+i}{\sqrt{2}} |\omega|^{1/2},$$

for which we cannot bound the real part independent of ω.

A Note on the Square-Root in $\mathbb{C}$

If we think of a complex number, $z \in \mathbb{C}$ in terms of its magnitude $r = |z|$, and angle θ, where $\tan(\theta) = \mathrm{imag}(z)/\mathrm{real}(z)$; we have

$$z = r\, e^{i\theta} = (r \cos(\theta) + i\, r \sin(\theta)),$$

and we can define

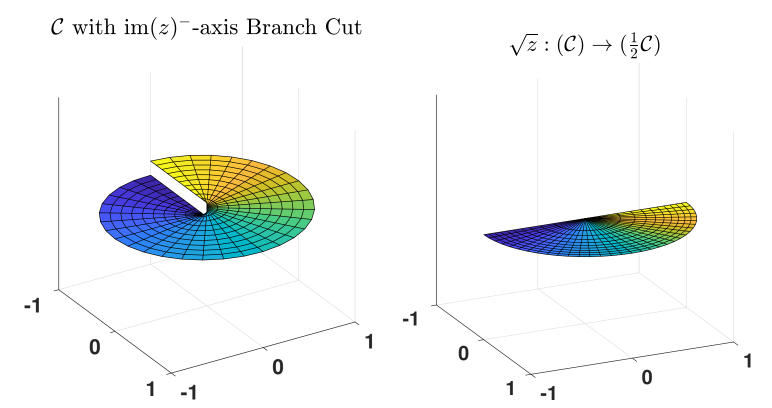
$$\sqrt{z} = \sqrt{r}\, e^{i\theta/2}.$$

This all makes (unique) sense once we restrict the angle to any 2π-interval by introducing a *branch cut*.

One possibility is to cut along the imaginary axis, and let $\theta \in (\pi, pi]$.

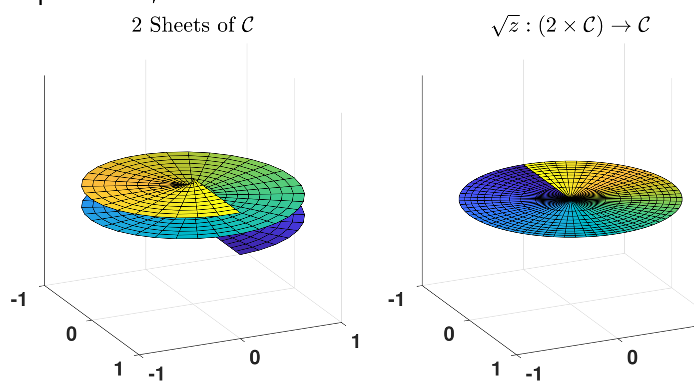
A Note on the Square-Root in $\mathbb{C}$

Figure: $\mathbb{C}$ with a branch-cut along the negative imaginary axis; $\theta \in (-\pi, \pi]$. The square-root associated with this branch ends up in the right half plane, with $\theta_{sqrt} \in (-\pi/2, \pi/2]$.



A Note on the Square-Root in $\mathbb{C}$

Figure: We can take 2 copies (sheets) of $\mathbb{C}$ with a branch-cuts, and glue them together; this way we get a surface with $\theta \in (-2\pi, 2\pi]$. The square-root associated with cork-screw space fills $\mathbb{C}$, with $\theta_{sqrt} \in (-\pi, \pi]$. If we identify the "loose ends" (-2π) and (2π) we see that the square root will map a trip "around the cork-screw space" into a unique trip around the complex plane... Interesting 1-to-1 correspondence, eh?



Second Order Equations

For second order equations of the form $u_{tt} = R(\partial_x)u$, with symbols $p(s,\omega) = s^2 - r(\omega)$, we get the roots

$$q_{\pm} = \pm\sqrt{r(\omega)},$$

and we must require that $r(\omega)$ must be close to (or on) the negative real axis — otherwise the square-root may end up "too deep" into the right half-plane.

The Wave- and Euler-Bernoulli equations

$$u_{tt} - a^2 u_{xx} = 0, \qquad u_{tt} = -b^2 u_{xxxx},$$

provide examples of this type.

The Euler-Bernoulli Equation: Lower Order Terms

Lower order terms can severely impact the well-posedness of the IVP for the Euler-Bernoulli equation, *consider*

$$u_{tt} = -b^2 u_{xxxx} + \mathbf{c u_{xxx}}.$$

The corresponding symbol is

$$p(s,\omega) = s^2 - r(\omega) = s^2 + b^2\omega^4 + \mathbf{i c}\omega^3$$

so that, with a little help from Taylor

$$q_{\pm}(\omega) = \pm\left[ib\omega^2 - \frac{\mathbf{c}\omega}{\mathbf{2b}} + \mathcal{O}(1).\right]$$

When $c \neq 0$, each root violates $\mathrm{Re}(q_{\pm}(\omega)) \leq \bar{q}$ for either positive or negative values of ω.

Well-Posedness of the Second-Order IVP

Theorem (Well-Posedness of the Second-Order IVP)

The initial value problem for the second-order equation

$$u_{tt} = R(\partial_x)u,$$

(where $r(\omega)$, the symbol of $R(\partial_x)$, is a polynomial of degree 2ρ) is well-posed if for each positive $t > 0$ there is a constant C_t such that for all solutions u

$$\|u(t,\circ)\|_{H^\rho} + \|u_t(t,\circ)\|_{H^0} \leq C_t\left(\|u(0,\circ)\|_{H^\rho} + \|u_t(0,\circ)\|_{H^0}\right).$$

Recall

$$\|u\|_{H^r}^2 := \int_{-\infty}^{\infty} \left(1 + |\omega|^2\right)^r |\hat{u}(\omega)|^2 \, d\omega.$$

Well-Posed Systems of Equations

With the help of the Fourier transform, any $d \times d$-system of first order can be put in the form

$$\widehat{u}_t = Q(\omega)\widehat{u},$$

where $\widehat{u} \in \mathbb{C}^d$, and $Q(\omega) \in \mathbb{C}^{d\times d}$. We can also let $\omega \in \mathbb{R}^n$, $n > 1$ if we are considering multiple space dimensions.

The solution of the IVP is given by

$$\widehat{u}(t,\omega) = e^{Q(\omega)t}\widehat{u}_0(\omega).$$

We formalize the well-posedness requirements in a theorem: ...

Well-Posedness for First Order Systems

Theorem (Well-Posedness for First Order Systems)

The necessary and sufficient condition for

$$\widehat{u}_t = Q(\omega)\widehat{u}$$

to be well-posed is that for each $t \geq 0$, there is a constant C_t such that

$$\|e^{Q(\omega)t}\| \leq C_t$$

for all $\omega \in \mathbb{R}^n$. A necessary condition for this to be true is that $\mathrm{Re}(q_\nu(\omega)) \leq \overline{q}$ holds for all eigenvalues of $Q(\omega)$.

The theorem is hard to use for general systems, since finding the eigenvalues may require a lot of work.

Special Case: $Q(\omega) = U(\omega)$, Upper Triangular

Lemma

Let U be an upper triangular matrix $\in \mathbb{C}^{d\times d}$ and let

$$\overline{u} = \max_{1\leq i\leq d} \mathrm{Re}(u_{ii}), \quad u^* = \max_{j>i} |u_{ij}|.$$

Then there is a constant C_d, such that

$$\|e^{Ut}\| \leq C_d e^{\overline{u}t}\left(1 + (tu^*)^{d-1}\right).$$

This lemma is used in conjunction with Schur's lemma (Math 543, or Strikwerda appendix A), which states that for any matrix $Q(\omega)$ we can find a unitary matrix $O(\omega)$, ($O(\omega)^H O(\omega) = I$, and $\|O(\omega)\|_2 = 1$), such that

$$\tilde{Q}(\omega) = O(\omega)Q(\omega)O(\omega)^{-1}$$

is an upper triangular matrix.

Well-Posedness for Systems

Now, we can write down a general inequality for all matrices $Q(\omega)$

$$\|e^{Q(\omega)t}\| = \|e^{\tilde{Q}(\omega)t}\| \leq C_d e^{\overline{q}(\omega)t}\left(1 + |tq^*(\omega)|^{d-1}\right),$$

where

$$\overline{q}(\omega) = \max_{1\leq\nu\leq d} \mathrm{Re}(q_\nu(\omega)), \quad q^*(\omega) = \max_{j>i} |\tilde{Q}_{ij}(\omega)|.$$

We see that the eigenvalues $(q_\nu(\omega))$ enter the inequality in a very predicable way, but that the well-posedness result also depends on the off-diagonal elements of $\tilde{Q}(\omega)$, which (physically) say something about how the quantities on $\widehat{u}$ interact (are "mixed") over time.

Parabolic Systems: General Definition

Definition (Parabolic System of PDEs)

The system

$$u_t = \sum_{j_1,j_2=1}^{n} B_{j_1,j_2} \frac{\partial^2 u}{\partial x_{j_1} \partial x_{j_2}} + \sum_{j=1}^{n} C_j \frac{\partial u}{\partial x_j} + Du$$

for which

$$Q(\omega) = -\sum_{j_1,j_2=1}^{n} B_{j_1,j_2} \omega_{j_1} \omega_{j_2} + i \sum_{j=1}^{n} C_j \omega_j + D$$

is parabolic if the eigenvalues, q_ν, of $Q(\omega)$ satisfy

$$\mathrm{Re}(q_\nu) \leq a - b|\omega|^2$$

for some constant a, and some positive constant b.

Hyperbolic Systems: General Definition

Definition (Hyperbolic System of PDEs)

The system

$$u_t = \sum_{j=1}^{n} A_j \frac{\partial u}{\partial x_j} + Bu, \quad \text{for which} \quad Q(\omega) = i \sum_{j=1}^{n} A_j \omega_j + B$$

is hyperbolic if the eigenvalues, q_ν, of $Q(\omega)$ satisfy

$$\mathrm{Re}(q_\nu) \leq c$$

for some constant c, and if $Q(\omega)$ is uniformly diagonalizable for large ω, *i.e.* for $|\omega| > K$, $\exists M(\omega)$ such that $M(\omega)Q(\omega)M^{-1}(\omega)$ is diagonal and $\|M(\omega)\| \leq M_b$, $\|M^{-1}(\omega)\| \leq M_b$, independently of ω.

Lower Order Terms: General Results 1 of 2

Theorem (Well-Posedness and Lower Order Terms)

If the system

$$\widehat{u}_t = Q(\omega)\widehat{u}$$

is well-posed and the matrix $Q_0(\omega)$ is bounded independently of ω, then the system

$$\widehat{u}_t = \left(Q(\omega) + Q_0(\omega)\right)\widehat{u}$$

is also well-posed.

The theorem directly tells us that the matrix B in hyperbolic systems, and the matrix D in parabolic systems do not affect the well-posedness of the corresponding systems.

The next theorem takes care of the C_j (first-derivative) terms for parabolic systems:

Lower Order Terms: General Results 2 of 2

Theorem (Well-Posedness and Lower Order Terms)

If the system

$$\widehat{u}_t = Q(\omega)\widehat{u}$$

satisfies

$$\|e^{Q(\omega)t}\| \leq K_t e^{-b|\omega|^\rho t}$$

for some positive constants b and ρ, with K_t independent of ω, and if $Q_0(\omega)$ satisfies

$$\|Q_0(\omega)\| \leq c_0|\omega|^\sigma$$

with $\sigma < \rho$, then the system

$$\widehat{u}_t = \left(Q(\omega) + Q_0(\omega)\right)\widehat{u}$$

is also well-posed.

Inhomogeneous Problems

For inhomogeneous problems, $Pu = f$, all the estimates and bounds will contain the energy added by the forcing function $f(t,x)$, *e.g.* for a first-order problem we can, as usual with the Fourier transform, write

$$\widehat{u}_t(t,\omega) = q(\omega)\widehat{u}(t,w) + r(\omega)\widehat{f}(t,\omega)$$

For well-posedness we need

$$\mathrm{Re}(q(\omega)) \le \overline{q}, \quad |r(\omega)| \le C_1.$$

The solution is given by

$$\widehat{u}(t,\omega) = e^{q(\omega)t}\widehat{u}_0(\omega) + r(\omega)\int_0^t e^{q(\omega)(t-s)}\widehat{f}(s,\omega)\,ds.$$

Inhomogeneous Problems

We quickly get the following bound

$$|\widehat{u}(t,\omega)|^2 \le Ce^{2\overline{q}t}\left[|\widehat{u}_0(\omega)|^2 + \int_0^t |\widehat{f}(s,\omega)|^2\,ds\right],$$

and by Parseval's relation

$$\|u(t,\circ)\|^2 \le Ce^{2\overline{q}t}\left[\|u_0\|^2 + \int_0^t \|f(s,\circ)\|^2\,ds\right].$$

Analogously, for a corresponding finite difference scheme we get

$$\|v^n\|^2 \le C_T\left[\|v^0\|^2 + k\sum_{\ell=0}^{n}\|f^\ell\|^2\right].$$

Duhamel's principle states that the solution to an inhomogeneous problem can be written as a super-position of solutions to homogeneous IVPs... One homogeneous IVP per time-level.

The Kreiss Matrix Theorem

Theorem (Kreiss Matrix Theorem — pt.1)

For a family $\mathcal{F}$ of $M \times M$ matrices, the following statements are equivalent:

A: *There exists a positive constant C_a such that for all $A \in \mathcal{F}$ and each non-negative integer n,*

$$\|A^n\| \le C_a.$$

R: *There exists a positive constant C_r such that for all $A \in \mathcal{F}$ and all complex numbers z with $|z| > 1$,*

$$\|(zI - A)^{-1}\| \le C_r(|z|-1)^{-1}.$$

...

The Kreiss Matrix Theorem

Theorem (Kreiss Matrix Theorem — pt.2)

S: *There exists positive constants C_s and C_b such that for each $A \in \mathcal{F}$ there is a non-singular Hermitian matrix S such that $B = SAS^{-1}$ is upper triangular and*

$$\|S\|,\ \|S^{-1}\| \le C_s$$
$$|B_{ii}| \le 1$$
$$|B_{ij}| \le C_b \min\{1 - |B_{ii}|,\ 1 - |B_{jj}|\}$$

for $i < j$.

H: *There exists a positive constant C_h such that for each $A \in \mathcal{F}$ there is a Hermitian matrix H such that*

$$C_h^{-1} I \le H \le C_h I$$
$$A^* H A \le H.$$

The Kreiss Matrix Theorem

Theorem (Kreiss Matrix Theorem — pt.3)

N: *There exists constants C_n and c_n such that for each $A \in \mathcal{F}$ there is a Hermitian matrix N such that*

$$C_n^{-1} I \le N \le C_n I$$
$$\mathrm{Re}(N(I - zA)) \ge c_n(1 - |z|)I$$

for all complex numbers z with $|z| \le 1$.

Ω: *There exists a positive constant C_ω such that for each $A \in \mathcal{F}$ there is a Hermitian matrix Ω such that*

$$C_\omega^{-1} I \le \Omega \le C_\omega I$$
$$\sup_{x \neq 0} \frac{|(\Omega A^n x, x)|}{(\Omega x, x)} \le 1.$$

The Kreiss Matrix Theorem

The theorem is of theoretical importance since it relates stability (condition A), with equivalent conditions that may be useful in different contexts.

It is of limited **practical** use in determining stability, since verifying any of the conditions is usually as difficult as verifying condition A itself.

In some applications it is important to know when the matrices H, N, and Ω can be constructed by (locally) continuous functions of the members of $\mathcal{F}$.