255 Summary

1 General Framework

- *Domain of dependence*: physical/numerical.
- $u_t = \mathcal{L}u$ with IC and periodic BC on a Hilbert space $\mathcal{H}$. $\mathcal{L}: \mathcal{H} \to \mathcal{H}$.
- Discretized to N-dimensional space $\mathcal{B}_N$, with projection operator $\mathcal{P}_N$.
- *Numerical solution:* $u_N \in \mathcal{B}_N$ solves

$$\frac{\partial u_N}{\partial t} = \mathcal{P}_N \mathcal{L} u_N, \quad u_N(0) = \mathcal{P}_N u_0.$$

- *Convergence:*

$$\lim_{N \to \infty} \|u_N(t) - \mathcal{P}_N u(t)\| = 0 \quad (0 \leqslant t \leqslant T).$$

- *Accuracy:*

$$\lim_{N \to \infty} \|\mathcal{P}_N \mathcal{L}(\mathrm{Id} - \mathcal{P}_N) u(t)\| = 0 \quad (0 \leqslant t \leqslant T).$$

- *Stability:*

$$\|\exp(\mathcal{P}_N \mathcal{L} \mathcal{P}_N t)\| \leqslant K \quad (0 \leqslant t \leqslant T).$$

- *Lax Equivalence Theorem (semidiscrete):* If the above IVP is well-posed and the scheme is stable and accurate, then it converges.

  Look at evolution of the error: $e_N = u_N - \mathcal{P}_N u$

$$\frac{\partial}{\partial t} e_N = \mathcal{P}_N \mathcal{L} \mathcal{P}_N e_N - \mathcal{P}_N \mathcal{L}(\mathrm{Id} - P_N)$$

  Integrate as ODE, estimates using accuracy and stability.

- *Order of convergence*: Equal to order of accuracy.

2 Well-Posedness

- *Solution operator* $S(t, t_0)$:
  - Semigroup property,
  - $S(t_0, t_0) = \mathrm{Id}$,
  - $\|S(t, t_0)\| \leqslant K e^{\alpha(t - t_0)}$.
- *Evolution Equation*:

$$u_t = \mathcal{P}\left(x, t, \frac{\partial}{\partial x}\right) u(x, t), \quad u(x, 0) = u_0(x)$$

  with $\mathcal{P}$ a polyomial of degree r. (use multi-indices in n-d)

- *Autonomous* if $\mathcal{P}$ does not depend on t. Then $S(t, t_0) = S(t - t_0)$.
- *Well-posed:* The above IVP is weakly well-posed (of order p, with $p \leqslant r$) if for every $f \in C_0^r$ and all $T_0 > 0$ there is a unique solution $u(x, t)$ satisfying

$$\|u(t)\| \leqslant C e^{\alpha t} \|f\|_p, \quad (0 \leqslant t \leqslant T_0).$$

  If $p = 0$, then *well-posed*. $\|\cdot\|_p$ is the Sobolev norm

$$\|u\|_p := \sum_{|\alpha| \leqslant p} \|\partial^\alpha u\|_{L^2} \sim \int (1 + |\omega|^p)^2 |\hat{f}(\omega)|^2 \mathrm{d}\omega.$$

- Well-posedness allows *defining* solutions by approximation.
- *Proving well-posedness:*
  - *Constant coefficients*:
    - Diagonalize systems, treat each equation separately if possible.
    - Fourier on Jordan block: try to turn into single derivative (multiply by $e^{\pm i\omega t}$?)
      True Jordan blocks become weakly well-posed.
    - Unbounded eigenvalues of symbol $\Rightarrow$ not well-posed.
    - Small perturbations of a weakly well-posed symbol can make that PDE not well-posed
  - *Several t-derivatives*: make it a system.
  - *Non-constant coefficients*:
    - Get an energy estimate: Multiply the equation by u, consider $\mathrm{d}/\mathrm{d}t\, E(t)$, use equation, integrate by parts to put derivatives *only* on coefficient.

2.1 Lower Order Perturbations

- *Duhamel Principle:* $u_t = \mathcal{P}(x,t,\partial_x)u + F(x,t)$ also has a solution, namely

$$u(x,t) = S(t,0)u_0(x) + \int_0^t S(t,\tau)F(x,\tau)d\tau.$$

  Proof: Differentiate solution.
- *Perturbed problems are well-posed:* $v_t = \mathcal{P}(x,t,\partial_x)v$ strongly well-posed.

$$u_t = \mathcal{P}(x,t,\partial_x)u + \mathcal{B}(x,t)u, \quad u(x,0) = f(x)$$

  has solution for $f \in C^\infty$. $\sup_{0\leqslant\tau\leqslant t} \|\mathcal{B}(x,\tau)u(\tau)\| \leqslant b_0\|u(t)\| \Rightarrow$ strongly well-posedness.
  (Proof: examine $y := e^{-\beta t}u(x,t)$ for $\beta \geqslant 0$, write down evolution, Duhamel that)

3 Convergence, Stability and Accuracy

Assume $\Delta x = h_i(\Delta t)$. $\|\cdot\|_N$ is discrete L^2.

- *Abstract FD scheme*: V^n a vector of point evaluations, E_k is the shift operator in the kth dimension.

$$B_0(E_1, \ldots, E_s)V_\alpha^{n+1} = B_1(E_1, \ldots, E_s)V_\alpha^n.$$

  Explicit iff $B_0 = \mathrm{Id}$.
- $V^{n+1} = C(\Delta t, \Delta x, \bar{x}, t)V^n$.
- $Q_{\Delta x}$ projection onto the point evaluation space.
- If, like in Leapfrog, we have dependency on two previous time steps: Interpret V as a vector of $(V^n, V^{n-1})^T$.
- *Accuracy:* Schme $C(\Delta t)$ is accurate of degree q_1 in space and q_2 in time: $\Leftrightarrow$

$$\underbrace{\frac{1}{\Delta t}\|[C(\Delta t)Q_{\Delta x} - Q_{\Delta x}S(\Delta t)]u(x,t)\|_N}_{\text{truncation error}} \leqslant K(t)(|\Delta x|^{q_1} + \Delta t^{q_2}).$$

- *Convergence:* For arbitrary t and $n\Delta t = t$,

$$\lim_{\Delta t\downarrow 0, \Delta x\downarrow 0} \|\{C^n(\Delta t)Q_{\Delta x} - Q_{\Delta x}S^n(\Delta t)\}f(x)\|_N = 0.$$

- *Stability:* For all n, Δt,
$$|C(\Delta t)^n| \leqslant K\, e^{\alpha n \Delta t}.$$
- The difference between accuracy and convergence (which is stability) is a promise about what happens if I shrink the timestep a lot.
- *Proving accuracy*: Plug true solution into the above.
- *Lax Equivalence Theorem:* $\exists$classical solution, scheme stable $\Rightarrow$ order of convergence $=$ order of accuracy, in both space and time.

  Proof: Write error evolution $\varepsilon^{n+1} = C(\Delta t)\varepsilon^n + \delta_n$, write $\varepsilon^n = \sum_k C(\Delta t)^{n-k-1}\delta_k$, estimate that using stability and accuracy.

  Can be generalized even if the IC is only in L^2 by approximation.
- *Kreiss Perturbation Theorem:* $V^{n+1} = C(\Delta t)V^n$ stable, $|D(\Delta t)|$ bounded

  $\Rightarrow$ perturbed scheme $V^{n+1} = \{C(\Delta t) + \Delta t\, D(\Delta t)\}V^n$ stable.

  Proof: $W^n = e^{-n\Delta t\beta}V^n$, write down evolution for it, Duhamel that.

4 Constant Coefficient Problems

- Depend neither on x nor t.
$$\begin{aligned} u_t &= \mathcal{P}(\partial_x)u \quad u(x,0) = f(x), \\ \hat{u}_t &= \mathbb{P}(i\omega)\hat{u} \quad \hat{u}(\omega,0) = \hat{f}(\omega). \end{aligned}$$
  $\mathbb{P}(i\omega)$ is called the *symbol* of the PDE.
- *Well-posedness:* Weakly (strongly for $p=0$) w-p $\Leftrightarrow \exists K, \alpha, p$ independent of ω:
$$|e^{\mathbb{P}(i\omega)t}| \leqslant K(1 + \|\omega\|^p)e^{\alpha t}.$$
  Proof: Use Fourier description of Sobolev norm: $\|(\|\omega\|^p + 1)^2|\hat{f}(\omega)|^2\|$.
- $A \leqslant B$ for two matrices A, B: $\Leftrightarrow A - B$ negative definite.
- *Sufficient condition for well-posedness:*
$$\exists\alpha: \mathbb{P}(i\omega) + \mathbb{P}(i\omega)^* \leqslant \alpha I.$$
  Proof: $\partial/\partial t(\hat{u}, \hat{u}) < \alpha(\hat{u}, \hat{u})$. (Adjoint-stuff)
- *Sharp criterion for well-posedness:* $\exists H(\omega)$ hermitian with $|H(\omega)|, |H^{-1}(\omega)| \leqslant K$
$$H(\omega)\mathbb{P}(i\omega) + \mathbb{P}(i\omega)^* H(\omega) \leqslant H(\omega).$$
  Proof: $\partial/\partial t(\hat{u}, H(\omega)\hat{u}) < \alpha(\hat{u}, \hat{u})$. (Adjoint-stuff)

  Remark: $H^{1/2}$ is a change of variables recovering the sufficient condition.
- *Well-posedness for normal matrices:* If $\mathbb{P}(i\omega)$ normal, then the IVP is well-posed iff
$$\operatorname{Re}\lambda_j(\omega) \leqslant \alpha.$$
  Proof: Norm coincides with the spectral radius.
- *Last criterion without normality:* You only get equivalence to weak well-posedness.

4.1 Hyperbolic Equations

- General form:
$$u_t = \sum_{j=1}^{s} A_j \partial_{x_j} u, \quad u(x,0) = u_0(x).$$
$$\mathbb{P}(i\omega) = \sum_{j=1}^{s} i\, A_j \omega_j.$$

- *Weakly hyperbolic*: purely imaginary eigenvalues.
- *Strongly hyperbolic*:
  - $\exists T(\omega)\colon |T(\omega)|, |T^{-1}(\omega)| \leqslant K$, T diagonalizes $\mathbb{P}(i\omega)$
  - purely imaginary eigenvalues.
- *Strictly hyperbolic*: weakly hyperbolic with pairwise distinct eigenvalues.
- *Symmetric hyperbolic*: $\exists S\colon S^{-1}A_jS$ *symmetric* (!)
- strictly $\Rightarrow$ strongly.
- symmetric $\Rightarrow$ strongly.
- weakly/strongly hyperbolic $\Rightarrow$ weakly/strongly well-posed
  Proof: non-normal criterion for weakly, otherwise $H = T^{-H}T^{-1}$.
- *Time reversal*: You may invert the sign on the A_j without affecting strong/weak hyperbolicity.
- *Calculating a symmetrizer:* Grab a diagonalizer for A_1, multiply by a well-chosen diagonal matrix.

5 Stability of Constant Coefficient Schemes

- *Obtaining a stability estimate*: Use Fourier ansatz
  $$V_j^n = \sum_{k=-\infty}^{\infty} \hat{V}^n(k) e^{ik\cdot(j\Delta x)}$$
  in the scheme.
- *Parseval's identity, discrete*:
  $$\frac{1}{N}\sum_{j=0}^{N=1} |V_j^n|^2 = \sum_{k=-\infty}^{\infty} |\hat{V}(k)|^2.$$
- *Amplification matrix:* $\mathcal{G}(\Delta t, k)$ in
  $$\hat{V}^{n+1} = \mathcal{G}(\Delta t, k)\hat{V}^n(k).$$
- *Stability condition*:
  $$|\{\mathcal{G}(\Delta t, k)\}^n| \leqslant K\, e^{\alpha n \Delta t}.$$
- *Von-Neumann condition*: Scheme stable $\Rightarrow$
  $$\rho[\mathcal{G}(\Delta t, k)] \leqslant e^{\gamma\Delta t} = 1 + O(\Delta t)$$
  VNC is sufficient if
  - $\mathcal{G}$ is normal ($\rho(\,\cdot\,) = \|\cdot\|$)
  - or diagonalizable by a bounded and inverse-bounded diagonalizer.

5.1 Kreiss Matrix Theorem

- *Stable family of matrices*: $\exists K\, \forall G \in \mathcal{F}\, \forall n \geqslant 0\colon |G^n| \leqslant k$.
- *Kreiss Matrix Theorem*: Equivalent:
  - $\mathcal{F}$ stable family
  - *Resolvent condition*: $\exists C$ $\forall$complex $|z| > 1$
    $$|(A - z\,\mathrm{Id})^{-1}| \leqslant \frac{C}{|z| - 1}.$$

- ∘ $\forall A \in \mathcal{F}\, \exists S \in \mathbb{R}^{p \times p}$ bounded, inverse-bounded s.t. $B = S\,A\,S^{-1}$ upper triangular
$$|b_{i,j}| \leqslant K_S \min\{1 - |b_{i,i}|, 1 - |b_{j,j}|\}$$
- ∘ *Energy Condition*: $\forall A \in \mathcal{F}\, \exists H \geqslant 0$ hermitian, bounded, inverse-bounded,
$$A^* H\, A \leqslant H.$$

Proof: Neumannsche Reihe, $H^{1/2}$ is a change of variables for energy condition.

5.2 Lax-Wendroff Condition

- *Numerical range* of a matrix G:
$$\tau(G) = \max_{V \in \mathbb{R}^{n \times n} \setminus \{0\}} \frac{\|V^H G\, V\|}{\|V^2\|}.$$
- G normal $\Rightarrow \tau(G) = \rho(G)$.
- *Lax-Wendroff-Theorem*: $\tau(G) \leqslant 1 \Rightarrow \exists K\colon \|G^n\| \leqslant K$.
  Proof: $\|G^n\| \leqslant \|G^n + (G^H)^n\| + \|G^n - (G^H)^n\|$.

5.3 Dissipative Schemes

- *Scheme dissipative* of order $2r$: $\Leftrightarrow$
$$\rho[\mathcal{G}(\Delta t, k)] \leqslant 1 - \delta |k \Delta x|^{2r}.$$

6 Examples

6.1 Transport

- $u_t = a\, u_x$ $(a > 0)$ Analytic solution: $u(x, t) = f(x + a\, t)$.
  (Left shift $\rightarrow$ Wind from right)
  - ∘ preserves energy $\int u^2$
  - ∘ preserves "mass" $\int |u|$ (chop up integral at sign changes)
- *CFL number*: (Courant, Friedrichs, Lewy)
$$\lambda = a \frac{\Delta t}{\Delta x}.$$
- *Scheme 1:*
$$V_j^{n+1} = V_j^n + \frac{\lambda}{2}(V_{j+1}^n - V_{j-1}^n)$$
  - ∘ (2,1)-accurate (Taylor)
  - ∘ unstable (Fourier; lin. combination of upwind and downwind scheme)
- *Lax-Friedrichs:*
$$V_j^{n+1} = \frac{1}{2}(V_{j+1}^n + V_{j-1}^n) + \frac{\lambda}{2}(V_{j+1}^n - V_{j-1}^n)$$
  - ∘ (1,1)-accurate ($g_k - e^{iakt} = O(\Delta t) + O(\Delta x)$)
  - ∘ stable if $|\lambda| \leqslant 1$
  - ∘ L^2 error at a given point $\rightarrow 0$ as Δt, $\Delta x \rightarrow 0$. (Fourier, Parseval, split tail off Fourier series)
  - ∘ Dissipates energy: $E(n+1) \leqslant E(n)$ (rewrite as $(1 + \lambda) V_{j+1} + (1 - \lambda) V_{j-1}$).
  - ∘ Dissipates mass (again, rewrite as $(1 + \lambda) V_{j+1} + (1 - \lambda) V_{j-1}$)

  - Dissipative of order 2.

- *Upwind Scheme*:

$$V_j^{n+1} = V_j^n + \lambda(V_{j+1}^n - V_j^n)$$

  - $(1,1)$-accurate
  - stable for $0 \leqslant \lambda \leqslant 1$
    (Fourier, $\sin(\xi) = \eta\sqrt{1-\eta^2}$, $\cos(\xi) = 1 - 2\eta^2$, where $\eta = \sin(\xi/2)$)

- *Leap frog scheme*:

$$\frac{V_j^{n+1} - V_j^{n-1}}{2\Delta t} = \frac{V_{j+1}^n - V_{j-1}^n}{2\Delta x}.$$

  - (2,2)-accurate.
  - Stable for $\lambda^2 < 1$.
  - Not dissipative. (conserves energy)

- *Lax-Wendroff*: Plug PDE into Taylor expansion of $u(t + \Delta t)$ until all time derivatives are gone. Use centered differences for spatial part.

$$V_j^{n+1} = V_j^n + \frac{\Delta t}{2\Delta x}(V_{j+1}^n - V_{j-1}^n) + \frac{(\Delta t)^2}{2(\Delta x)^2}(V_{j+1}^n - 2V_j^n + V_{j-1}^n)$$

  - (2,2)-accurate.
  - Dissipative of order 4.

- *Crank-Nicholson*:

$$V_j^{n+1} = V_j^n + \frac{\Delta t}{2\Delta x}(V_{j+1}^{n+1} - V_{j-1}^{n+1} + V_{j+1}^n - V_{j-1}^n)$$

  - (2,2)-accurate.

6.2 Heat

- $u_t = u_{xx}$.
- $\lambda = \Delta t / \Delta x^2 \leqslant 1/2$ for standard centered difference stuff.

6.3 Schrödinger

- $u_t = i\, u_{xx}$.
- $\mathbb{P}(i\omega) + \mathbb{P}(i\omega)^* = 0 \Rightarrow$ Energy conservation.
- centered differences are unstable.