

Numerical Analysis - Part II

Anders C. Hansen

Lecture 10

*The diffusion equation in two space
dimensions*

The diffusion equation in two space dimensions

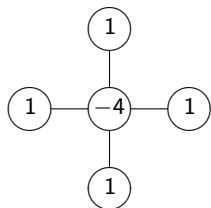
We are solving

$$\frac{\partial u}{\partial t} = \nabla^2 u, \quad 0 \leq x, y \leq 1, \quad t \geq 0, \quad (1)$$

where $u = u(x, y, t)$, together with initial conditions at $t = 0$ and Dirichlet boundary conditions at $\partial\Omega$, where $\Omega = [0, 1]^2 \times [0, \infty)$. It is straightforward to generalize our derivation of numerical algorithms, e.g. by the method of lines.

Recall the five point formula

We have the *five-point method*



$$u_{i,j} = u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j},$$

discretising the two dimensional Laplacian.

The diffusion equation in two space dimensions

Thus, let $u_{\ell,m}(t) \approx u(\ell h, mh, t)$, where $h = \Delta x = \Delta y$, and let $u'_{\ell,m} \approx u_{\ell,m}(nk)$ where $k = \Delta t$. The five-point formula results in

$$u'_{\ell,m} = \frac{1}{h^2} (u_{\ell-1,m} + u_{\ell+1,m} + u_{\ell,m-1} + u_{\ell,m+1} - 4u_{\ell,m}),$$

or in the matrix form

$$\mathbf{u}' = \frac{1}{h^2} \mathbf{A}_* \mathbf{u}, \quad \mathbf{u} = (u_{\ell,m}) \in \mathbb{R}^N, \quad (2)$$

where \mathbf{A}_* is the block TST matrix of the five-point scheme:

$$\mathbf{A}_* = \begin{bmatrix} H & I & & & \\ & I & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ & & & & I & H \end{bmatrix}, \quad H = \begin{bmatrix} -4 & 1 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 1 & -4 \end{bmatrix}.$$

The diffusion equation in two space dimensions

Thus, the Euler method yields

$$u_{\ell,m}^{n+1} = u_{\ell,m}^n + \mu(u_{\ell-1,m}^n + u_{\ell+1,m}^n + u_{\ell,m-1}^n + u_{\ell,m+1}^n - 4u_{\ell,m}^n), \quad (3)$$

or in the matrix form

$$\mathbf{u}^{n+1} = A\mathbf{u}^n, \quad A = I + \mu A_*$$

where, as before, $\mu = \frac{k}{h^2} = \frac{\Delta t}{(\Delta x)^2}$. The local error is

$\eta = \mathcal{O}(k^2 + kh^2) = \mathcal{O}(h^4)$. To analyse stability, we notice that A is symmetric, hence normal, and its eigenvalues are related to those of A_* by the rule

$$\lambda_{k,\ell}(A) = 1 + \mu\lambda_{k,\ell}(A_*) \stackrel{\text{Prop. 1.12}}{=} 1 - 4\mu \left(\sin^2 \frac{\pi kh}{2} + \sin^2 \frac{\pi \ell h}{2} \right).$$

Consequently,

$$\sup_{h>0} \rho(A) = \max\{1, |1 - 8\mu|\}, \quad \text{hence} \quad \mu \leq \frac{1}{4} \Leftrightarrow \text{stability.}$$

Fourier analysis in 2D

Fourier analysis generalizes to two dimensions: of course, we now need to extend the range of (x, y) in (1) from $0 \leq x, y \leq 1$ to $x, y \in \mathbb{R}$. A 2D Fourier transform reads

$$\hat{u}(\theta, \psi) = \sum_{\ell, m \in \mathbb{Z}} u_{\ell, m} e^{-i(\ell\theta + m\psi)}$$

and all our results readily generalize.

Fourier analysis in 2D

In particular, the Fourier transform is an isometry from $\ell_2[\mathbb{Z}^2]$ to $L_2([-\pi, \pi]^2)$, i.e.

$$\left(\sum_{\ell, m \in \mathbb{Z}} |u_{\ell, m}|^2 \right)^{1/2} =: \|\mathbf{u}\| = \|\widehat{u}\|_* := \left(\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |\widehat{u}(\theta, \psi)|^2 d\theta d\psi \right)^{1/2},$$

and the method is stable iff $|H(\theta, \psi)| \leq 1$ for all $\theta, \psi \in [-\pi, \pi]$. The proofs are an easy elaboration on the one-dimensional theory.

Insofar as the Euler method (3) is concerned,

$$H(\theta, \psi) = 1 + \mu (e^{-i\theta} + e^{i\theta} + e^{-i\psi} + e^{i\psi} - 4) = 1 - 4\mu \left(\sin^2 \frac{\theta}{2} + \sin^2 \frac{\psi}{2} \right),$$

and we again deduce stability if and only if $\mu \leq \frac{1}{4}$.

Parseval's identity

Lemma 1 (Parseval's identity)

For any $\mathbf{v} \in \ell_2[\mathbb{Z}]$, we have $\|\mathbf{v}\| = \|\widehat{\mathbf{v}}\|_*$.

Proof. By definition,

$$\begin{aligned}\|\widehat{\mathbf{v}}\|_*^2 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{m \in \mathbb{Z}} e^{-im\theta} v_m \right|^2 d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{m \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} v_m \bar{v}_k e^{-i(m-k)\theta} d\theta \\ &= \frac{1}{2\pi} \sum_{m \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} v_m \bar{v}_k \int_{-\pi}^{\pi} e^{-i(m-k)\theta} d\theta \stackrel{(*)}{=} \sum_{m \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} v_m \bar{v}_k \delta_{m-k} = \|\mathbf{v}\|^2,\end{aligned}$$

where equality (*) is due to the fact that

$$\int_{-\pi}^{\pi} e^{-i\ell\theta} d\theta = \begin{cases} 2\pi, & \ell = 0, \\ 0, & \ell \in \mathbb{Z} \setminus \{0\}, \end{cases}$$

□

The implication of the lemma is that the Fourier transform is an *isometry* of the Euclidean norm. This is an important reason underlying its many applications in mathematics and beyond.

Crank-Nicolson for 2D

Applying the trapezoidal rule to our semi-discretization (2) we obtain the two-dimensional Crank-Nicolson method:

$$(I - \frac{1}{2}\mu A_*) \mathbf{u}^{n+1} = (I + \frac{1}{2}\mu A_*) \mathbf{u}^n, \quad (4)$$

in which we move from the n -th to the $(n+1)$ -st level by solving the system of linear equations $B\mathbf{u}^{n+1} = C\mathbf{u}^n$, or $\mathbf{u}^{n+1} = B^{-1}C\mathbf{u}^n$. For stability, similarly to the one-dimensional case, the eigenvalue analysis implies that $A = B^{-1}C$ is normal and shares the same eigenvectors with B and C , hence

$$\lambda(A) = \frac{\lambda(C)}{\lambda(B)} = \frac{1 + \frac{1}{2}\mu\lambda(A_*)}{1 - \frac{1}{2}\mu\lambda(A_*)} \Rightarrow |\lambda(A)| < 1 \text{ as } \lambda(A_*) < 0$$

and the method is stable for all μ . The same result can be obtained through the Fourier analysis.

Crank-Nicolson for 2D

We would like to find a fast solver to the system (4). The matrix $B = I - \frac{1}{2}\mu A_*$ has a structure similar to that of A_* , where

$$A_* = \begin{bmatrix} H & I & & & \\ & I & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ & & & & I & H \end{bmatrix}, \quad H = \begin{bmatrix} -4 & 1 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 1 & -4 \end{bmatrix}.$$

so we may apply the Hockney method.

The total computational cost per iteration is $\mathcal{O}(M^2 \log M)$ for a $M \times M$ discretization grid.

Matlab demo: Download the Matlab GUI for *Solving the Wave and Diffusion Equations in 2D* from http://www.damtp.cam.ac.uk/user/hf323/M21-II-NA/demos/pdes_2d/pdes_2d.html and solve the diffusion equation (1) for different initial conditions. For the numerical solution of the equation you can choose from the Euler method and the Crank-Nicolson scheme. The GUI allows you to solve the wave equation as well. Compare the behaviour of solutions!

In all the examples of semi-discretization we have seen so far, we always reach a linear system of ODE of the form:

$$\mathbf{u}' = A\mathbf{u}, \quad \mathbf{u}(0) = \mathbf{u}_0. \quad (5)$$

The solution of this linear system of ODE is given by

$$\mathbf{u}(t) = e^{tA}\mathbf{u}_0 \quad (6)$$

where the *matrix exponential* function is defined by $e^B := \sum_{k=0}^{\infty} \frac{1}{k!} B^k$. It is easily verified that $de^{tA}/dt = Ae^{tA}$, therefore (6) is indeed a solution of (5).

If A can be diagonalized $A = VDV^{-1}$, then $e^{tA} = Ve^{tD}V^{-1}$ where e^{tD} is the diagonal matrix consisting $\text{diag}(e^{tD_{ii}})$. As such one can compute the solution of (5) exactly. However computing an eigenvalue decomposition can be costly, and so one would like to consider more efficient methods, based on the solution of sparse linear systems instead.

Observe that one-step methods for solving (5) are approximating a matrix exponential. Indeed, with $k = \Delta t$, we have:

$$\text{Euler:} \quad \mathbf{u}^{n+1} = (I + kA)\mathbf{u}^n, \quad e^z = 1 + z + \mathcal{O}(z^2);$$

$$\text{Implicit Euler:} \quad \mathbf{u}^{n+1} = (I - kA)^{-1}\mathbf{u}^n, \quad e^z = (1 - z)^{-1} + \mathcal{O}(z^2);$$

$$\text{Trapezoidal:} \quad \mathbf{u}^{n+1} = \left(I - \frac{1}{2}kA\right)^{-1} \left(I + \frac{1}{2}kA\right) \mathbf{u}^n, \quad e^z = \frac{1 + \frac{1}{2}z}{1 - \frac{1}{2}z} + \mathcal{O}(z^3).$$

In practice the matrix A is very sparse, and this can be exploited when solving linear systems e.g., for the implicit Euler or Trapezoidal Rule.

Splitting

In many cases, the matrix A is naturally expressed as a *sum of two matrices*, $A = B + C$. For example, when discretizing the diffusion equation in 2D with zero boundary conditions, we have $A = \frac{1}{h^2}(A_x + A_y)$ where $\frac{1}{h^2}A_x \in \mathbb{R}^{M^2 \times M^2}$ corresponds to the 3-point discretization of $\frac{\partial^2}{\partial x^2}$, and $\frac{1}{h^2}A_y \in \mathbb{R}^{M^2 \times M^2}$ corresponds to the 3-point discretization of $\frac{\partial^2}{\partial y^2}$. In matrix notations, if the grid points are ordered by columns, then we have:

$$A_x = \begin{bmatrix} -2I & I & & & \\ & I & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ & & & & I & -2I \end{bmatrix}, A_y = \begin{bmatrix} G & & & \\ & G & & \\ & & \ddots & \\ & & & G \end{bmatrix}, G = \begin{bmatrix} -2 & 1 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 1 & -2 \end{bmatrix} \in \mathbb{R}^{M \times M}. \quad (7)$$

Kronecker product

Remark: It is convenient to note that $A_x = G \otimes I$ and $A_y = I \otimes G$, where \otimes is the Kronecker product of matrices (`kron` in Matlab) defined by

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \dots & A_{1m_A}B \\ A_{21}B & A_{22}B & \dots & A_{2m_A}B \\ \vdots & & & \\ A_{n_A1}B & \dots & \dots & A_{n_Am_A}B \end{bmatrix} \in \mathbb{R}^{n_A n_B \times m_A m_B}$$

where $A \in \mathbb{R}^{n_A \times m_A}$ and $B \in \mathbb{R}^{n_B \times m_B}$.

In general, $\exp(t(B + C)) \neq \exp(tB) \exp(tC)$. Equality holds however when B and C commute.

Splitting the exponential

Proposition 2

For any matrices B, C ,

$$e^{t(B+C)} = e^{tB}e^{tC} + \frac{1}{2}t^2(CB - BC) + \mathcal{O}(t^3). \quad (8)$$

If B and C commute, then $e^{B+C} = e^B e^C$.

Proof. We Taylor-expand both expressions $e^{tB}e^{tC}$ and $e^{t(B+C)}$:

$$\begin{aligned} e^{tB}e^{tC} &= (I + tB + t^2B^2/2 + \mathcal{O}(t^3))(I + tC + t^2C^2/2 + \mathcal{O}(t^3)) \\ &= I + t(B + C) + \frac{t^2}{2}(B^2 + C^2 + 2BC) + \mathcal{O}(t^3) \end{aligned}$$

and

$$\begin{aligned} e^{t(B+C)} &= I + t(B + C) + \frac{t^2}{2}(B + C)^2 + \mathcal{O}(t^3) \\ &= I + t(B + C) + \frac{t^2}{2}(B^2 + C^2 + BC + CB) + \mathcal{O}(t^3). \end{aligned}$$

Equation (8) follows.

Splitting the exponential

Proof.

When B and C commute, we can write:

$$\exp(B+C) = \sum_{n=0}^{\infty} \frac{1}{n!} (B+C)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^n \binom{n}{k} B^{n-k} C^k \right) = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{1}{n!} \binom{n}{k} B^{n-k} C^k$$

Recall that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, so

$$\exp(B+C) = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{1}{k!(n-k)!} B^{n-k} C^k = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{k!m!} B^m C^k = e^B e^C.$$

□

Splitting for the 2D diffusion equation

Splitting the exponential

It is straightforward to verify that A_x and A_y commute; namely $A_x A_y = A_y A_x = G \otimes G$ (check out the basic rules of multiplication with the kronecker product

https://en.wikipedia.org/wiki/Kronecker_product). This should not come as a surprise since the operators $\partial^2/\partial x^2$ and $\partial^2/\partial y^2$, which A_x/h^2 and A_y/h^2 approximate, are known to commute. So we can write

$$e^{k(A_x+A_y)/h^2} = e^{kA_x/h^2} e^{kA_y/h^2}.$$

This means that the solution of the semi-discretized diffusion equation in 2D, with zero boundary conditions, satisfies

$$\mathbf{u}^{n+1} = e^{kA_x/h^2} e^{kA_y/h^2} \mathbf{u}^n. \tag{9}$$

Split Crank-Nicolson

In the split Crank-Nicolson scheme, we approximate each exponential map in (9) by the rational function

$$r(z) = (1 + z/2)(1 - z/2)^{-1},$$

which leads to

$$\mathbf{u}^{n+1} = (I + \frac{\mu}{2}A_x)(I - \frac{\mu}{2}A_x)^{-1}(I + \frac{\mu}{2}A_y)(I - \frac{\mu}{2}A_y)^{-1}\mathbf{u}^n. \quad (10)$$

Note that computing $\mathbf{u}^{n+1/2} = (I + \frac{\mu}{2}A_y)(I - \frac{\mu}{2}A_y)^{-1}\mathbf{u}^n$ can be done efficiently in $\mathcal{O}(M^2)$ time as A_y is block-diagonal, and the matrices G are tridiagonal (each tridiagonal solve requires $\mathcal{O}(M)$ time, and we have M of these). Computing $\mathbf{u}^{n+1} = (I + \frac{\mu}{2}A_x)(I - \frac{\mu}{2}A_x)^{-1}\mathbf{u}^{n+1/2}$ can also be done in $\mathcal{O}(M^2)$ time, since A_x is also block-diagonal provided we appropriately permute the rows and columns so that the grid ordering is by rows instead of columns. This means that the update step (10) of Split-Crank-Nicolson can be performed in time $\mathcal{O}(M^2)$ and only requires tridiagonal matrix solves (no FFT needed).

Split Crank-Nicolson

One can easily verify stability of the split Crank-Nicolson scheme. Indeed, we can write

$$\|r(\mu A_x)r(\mu A_y)\|_2 \leq \|r(\mu A_x)\|_2 \|r(\mu A_y)\|_2 \leq 1$$

since, as seen in previous lectures,

$\|r(\mu A_x)\|_2 = \|(I + \frac{\mu}{2}A_x)(I - \frac{\mu}{2}A_x)^{-1}\|_2 \leq 1$ since A_x is symmetric and its eigenvalues are ≤ 0 . (Same for $\|r(\mu A_y)\|_2$.)

Exercise: Check the consistency of the scheme

$$\mathbf{u}^{n+1} = r(\mu A_x)r(\mu A_y)\mathbf{u}^n.$$

In particular, show that split Crank-Nicolson has the 'same' local error as the classical Crank-Nicolson scheme. That is the local error is $\mathcal{O}(k^3 + kh^2)$.

2D diffusion with variable diffusion coefficient

2D diffusion with variable diffusion coefficient

In general, however, the matrices B and C in $A = B + C$ do not have to commute, as in the following example: The general diffusion equation with a diffusion coefficient $a(x, y) > 0$ is given by:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(a(x, y) \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(a(x, y) \frac{\partial u}{\partial y} \right), \quad (11)$$

together with initial conditions on $[0, 1]^2$ and Dirichlet boundary conditions along $\partial[0, 1]^2 \times [0, \infty)$. We replace each space derivative by *central differences* at midpoints,

$$\frac{dg(\xi)}{d\xi} \approx \frac{g(\xi + \frac{1}{2}h) - g(\xi - \frac{1}{2}h)}{h},$$

resulting in the ODE system

$$\begin{aligned} u'_{\ell, m} = & \frac{1}{h^2} \left[a_{\ell-\frac{1}{2}, m} u_{\ell-1, m} + a_{\ell+\frac{1}{2}, m} u_{\ell+1, m} + a_{\ell, m-\frac{1}{2}} u_{\ell, m-1} + a_{\ell, m+\frac{1}{2}} u_{\ell, m+1} \right. \\ & \left. - \left(a_{\ell-\frac{1}{2}, m} + a_{\ell+\frac{1}{2}, m} + a_{\ell, m-\frac{1}{2}} + a_{\ell, m+\frac{1}{2}} \right) u_{\ell, m} \right]. \end{aligned} \quad (12)$$

2D diffusion with variable diffusion coefficient

Assuming zero boundary conditions, we have a system $\mathbf{u}' = A\mathbf{u}$, and the matrix A can be split as $A = \frac{1}{h^2}(A_x + A_y)$. Here, A_x and A_y are again constructed from the contribution of discretizations in the x - and y -directions respectively, namely A_x includes all the $a_{\ell \pm \frac{1}{2}, m}$ terms, and A_y consists of the remaining $a_{\ell, m \pm \frac{1}{2}}$ components. The resulting operators A_x and A_y do not necessarily commute, and so the splitting scheme

$$\mathbf{u}^{n+1} = e^{kA_x/h^2} e^{kA_y/h^2} \mathbf{u}^n$$

will carry an error of $\mathcal{O}(k^2)$, following Proposition 2.

One can obtain better splitting approximations of $e^{t(B+C)}$. For example it is not hard to prove that $e^{\frac{1}{2}tB}e^{tC}e^{\frac{1}{2}tB}$ gives a $\mathcal{O}(t^3)$ approximation of $e^{t(B+C)}$, i.e.,

$$e^{t(B+C)} = e^{\frac{1}{2}tB}e^{tC}e^{\frac{1}{2}tB} + \mathcal{O}(t^3). \quad (13)$$

Splitting of inhomogeneous systems

Our exposition so far has been limited to the case of zero boundary conditions. In general, the linear ODE system is of the form

$$\mathbf{u}' = A\mathbf{u} + \mathbf{b}, \quad \mathbf{u}(0) = \mathbf{u}^0, \quad (14)$$

where \mathbf{b} originates in boundary conditions (and, possibly, in a forcing term $f(x, y)$ in the original PDE (11)). Note that our analysis should accommodate $\mathbf{b} = \mathbf{b}(t)$, since boundary conditions might vary in time! The exact solution of (14) is provided by the *variation of constants* formula

$$\mathbf{u}(t) = e^{tA}\mathbf{u}(0) + \int_0^t e^{(t-s)A}\mathbf{b}(s) ds, \quad t \geq 0,$$

therefore

$$\mathbf{u}(t_{n+1}) = e^{kA}\mathbf{u}(t_n) + \int_{t_n}^{t_{n+1}} e^{(t_{n+1}-s)A}\mathbf{b}(s) ds.$$

The integral on the right-hand side can be evaluated using quadrature.

Splitting of inhomogeneous systems

For example, the trapezoidal rule

$$\int_0^k g(\tau) d\tau = \frac{1}{2}k[g(0) + g(k)] + \mathcal{O}(k^3) \text{ gives}$$

$$\mathbf{u}(t_{n+1}) \approx e^{kA}\mathbf{u}(t_n) + \frac{1}{2}k[e^{kA}\mathbf{b}(t_n) + \mathbf{b}(t_{n+1})],$$

with a local error of $\mathcal{O}(k^3)$. We can now replace exponentials with their splittings. For example, Strang's splitting (13), together with the rational function approximation $r(z) = (1 + z/2)/(1 - z/2)$ of the exponential map, results in

$$\mathbf{u}^{n+1} = r\left(\frac{1}{2}kB\right) r(kC) r\left(\frac{1}{2}kB\right) [\mathbf{u}^n + \frac{1}{2}k\mathbf{b}^n] + \frac{1}{2}k\mathbf{b}^{n+1}.$$

As before, everything reduces to (inexpensive) solution of tridiagonal systems.