Essential Partial Differential Equations: Analytical and Computational Aspects Solutions to odd numbered exercises

David F. Griffiths, John W. Dold and David J. Silvester

Springer International Publishing Switzerland, 2015

Solutions to all exercises are available to approved instructors by contacting the publishers.

Springer Undergraduate Mathematics Series

ISBN: 978-3-319-22568-5, e-ISBN: 978-3-319-22569-2

Essential Partial Differential Equations: Analytical and Computational Aspects Solutions to odd numbered exercises

David F. Griffiths, John W. Dold and David J. Silvester

Exercises

Exercises 1 Introduction

1.1

Function	Comment	Conclusion
$u(x,y) = A(y)$	$u_y = A'(y)$	False
$u(x,y) = A(y)$	$u_{xy} = 0$	True
$u(x,t) = A(x)B(t)$	$u_{xy} = 0$ concerns different independent variables!	False
$u(x,t) = A(x)B(t)$	$uu_{xt} = ABA'B' = u_x u_t$	True
$u(x,t,y) = A(x,y)$	$u_t = \partial_t A(x,y) = 0$	True
$u(x,t) = A(x{+}ct) + B(x{-}ct)$	$u_{tt} + c^2 u_{xx} = 2c^2(A'' + B'')$	False

1.3
These are not the only possible cases; you might find other PDEs:

(a) $u(x,t) = A(x{+}ct) + B(x{-}ct)$: $u_t = cA'(x{+}ct) - cB'(x{-}ct)$, $u_x = A'(x{+}ct) + B'(x{-}ct)$, $u_{tt} = c^2A''(x{+}ct) + c^2B''(x{-}ct)$, $u_{xx} = A''(x{+}ct) + B''(x{-}ct)$ so $u_{tt} - c^2u_{xx} = 0$ (wave equation).

(b) $u(x,t) = A(x) + B(t)$: $u_t = B'(t)$ so $u_{tx} = 0$.

(c) $u(x,t) = A(x)/B(t)$: $\ln u = \ln A(x) - \ln B(x)$ so $(\ln u)_{tx} = 0$ or $uu_{tx} - u_t u_x = 0$.

(d) $u(x,t) = A(xt)$: $u_t = xA'(xt)$, $u_x = tA'(xt)$, so $tu_t - xu_x = 0$.

(e) $u(x,t) = A(x^2t)$: $u_t = x^2A'(x^2t)$, $u_x = 2xtA'(x^2t)$ so $2tu_t - xu_x = 0$.

(f) $u(x,t) = A(x^2/t)$: $u_t = -\frac{x^2}{t^2}A'(x^2/t)$, $u_x = -\frac{2x}{t}A'(x^2/t)$ so $2tu_t + xu_x = 0$.

1.5
Suppose that $u(x,t) = \frac{1}{2}c\,\mathrm{sech}^2(z)$, where $z = \frac{1}{2}\sqrt{c}(x{-}ct{-}x_0)$. Then $z_t = -\frac{1}{2}c^{3/2}$ and $z_x = \frac{1}{2}c^{1/2}$ so, by the chain rule, we find

$$\partial_t u(x,t) = \tfrac{1}{2}c^{5/2}\frac{\sinh z}{\cosh^3 z}, \qquad \partial_x u(x,t) = -\tfrac{1}{2}c^{3/2}\frac{\sinh z}{\cosh^3 z}$$

$$\partial_t u + 6u\partial_x u = \tfrac{1}{2}c^{5/2}(\sinh z)\frac{\cosh^2(z) - 3}{\cosh^5 z} = -\partial_x^3 u(x,t)$$

and so $u_t + 6uu_x + u_{xxx} = 0$.

1.7
Since $u = -2\partial_x\phi = -2\phi_x/\phi$, the partial derivatives are

$$u_t = -2\frac{\phi_{xt}}{\phi} - 2\frac{\phi_x\phi_t}{\phi^2}, \qquad u_x = -2\frac{\phi_{xx}}{\phi} + 2\frac{(\phi_x)^2}{\phi^2},$$

$$u_{xx} = -2\frac{\phi_{xxx}}{\phi} + 2\frac{\phi_{xx}\phi_x}{\phi^2} + 4\frac{\phi_x}{\phi}\left(\frac{\phi_{xx}}{\phi} - \frac{(\phi_x)^2}{\phi^2}\right)$$

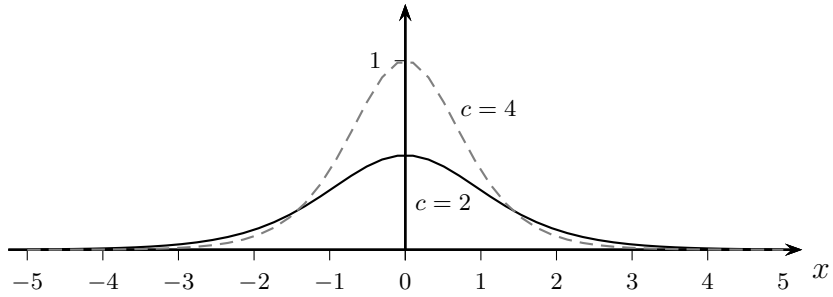


Figure 1: Soliton solutions of the KdV equation with $c = 2$ (solid) and $c = 4$ (dashed) for Exercise 1.5 travel with speed c.

so, with $\phi_t = \phi_{xx}$ and $\phi_{xt} = \phi_{xxx}$,

$$u_t = -2\frac{\phi_{xxx}}{\phi} - 2\frac{\phi_x \phi_{xx}}{\phi^2}$$

and $u_t + uu_x = u_{xx}$. With $\phi = t^{-1/2}\exp(-x^2/(4t))$ we find $\log\phi = -\frac{1}{2}\log(t) - x^2/(4t)$ and so $u = x/t$, which is a special case of equation (9.32).

1.9

The given function $\phi(x,t)$ is a linear combination of solutions of the heat equation and is therefore a viable candidate for use in the Cole-Hopf transformation. We find

$$u(x,t) = 2a\frac{2\mathrm{e}^{-2a(x-2at)} + \mathrm{e}^{-a(x-x_0-at)}}{1 + \mathrm{e}^{-2a(x-2at)} + \mathrm{e}^{-a(x-x_0-at)}}$$

which is shown in Fig. 2 when $a = 1$ and $x_0 = 10$. Two "step"-like solutions coalesce into one.

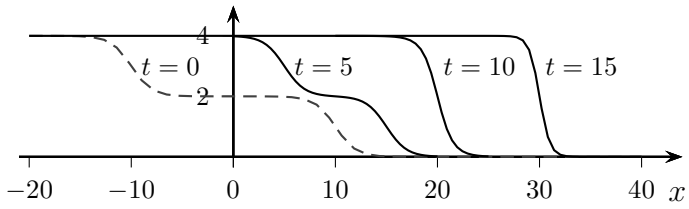


Figure 2: The solution of Burger's equation at times $t = 0, 5, 10, 15$ for Exercise 1.9.

Exercises 2 Boundary and initial data

2.1

(a) $u(x,0) = A(x) + B(x) = f(x)$ so $A(x) + B(x) = f(x)$. There is not enough information to determine both $A(\cdot)$ and $B(\cdot)$.

(b) $u(x,0) = A(x)+B(0) = f(x)$ so $A(x) = f(x)-B(0)$. We would only need to know one value of B, namely $B(0)$ to determine A. However, the initial conditions gives no information about B.

(c) $u(x,0) = A(x)/B(0) = f(x)$ so $A(x) = B(0)f(x)$. We would only need to know one value of B, namely $B(0)$ to determine A, but we have no information about B.

(d) $u(x,0) = A(0) = f(x)$ so $A(0) = f(x)$. This tries to set a constant $A(0)$ to something that is not constant $f(x)$, which is not possible!

(e) (exactly the same)

(f) $u(x,0) = A(\infty) = f(x)$ so $A(\infty) = f(x)$. This tries to set a constant $A(\infty)$ to something that is not constant $f(x)$, which is again not possible!

2.3

From (1.2) with $g(x)$ as given,

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \mathrm{e}^{-(x-s)^2/4t} g(s)\,\mathrm{d}s = \frac{1}{\sqrt{4\pi t}} \int_{0}^{\infty} \mathrm{e}^{-(x-s)^2/4t}\,\mathrm{d}s = \frac{1}{\sqrt{\pi}} \int_{-x/\sqrt{4t}}^{\infty} \mathrm{e}^{-z^2}\,\mathrm{d}z,$$

where we have made the change of variable $s = x + z\sqrt{4t}$ in the integrand. Thus $u(0,t) = 1/2$ from the given result.

$$u(x,0) = \lim_{t\to 0} \frac{1}{\sqrt{\pi}} \int_{-x/\sqrt{4t}}^{\infty} \mathrm{e}^{-z^2}\,\mathrm{d}z = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-z^2}\,\mathrm{d}z = 2\frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \mathrm{e}^{-z^2}\,\mathrm{d}z = 1,$$

since the integrand is an even function of z.

2.5

First note that $\mathcal{L}(au) = -\kappa(au)_{xx} = a(-\kappa u_{xx}) = a\mathcal{L}u$. Similarly, $\mathcal{B}(au) = a\mathcal{B}u$. Then

$$(\mathscr{L}au)(x,t) = \begin{cases} (au)_t(x,t) + (\mathcal{L}au)(x,t) \\ (\mathcal{B}au)(x,t) \\ au(x,0) \end{cases} = \begin{cases} a(u_t(x,t) + \mathcal{L}u(x,t)) & \text{for } (x,t) \in (0,1)\times(0,T) \\ a\mathcal{B}u(x,t) & \text{for } (x,t) \in \{0,1\}\times(0,T) \\ au(x,0) & \text{for } t = 0, x \in [0,1] \end{cases}.$$

and the right hand side is equal to $a\mathscr{L}u(x,t)$.

2.7

First we confirm that $u = AT^{1/2}(T-t)^{-1/2}\mathrm{e}^{-x^2/4(T-t)}$ is a solution of $u_t = -u_{xx}$:

$$u_t = \frac{1}{2}AT^{1/2}(T-t)^{-3/2}\mathrm{e}^{-x^2/4(T-t)} - \frac{1}{4}AT^{1/2}x^2(T-t)^{-5/2}\mathrm{e}^{-x^2/4(T-t)},$$

$$u_x = -\frac{1}{2}AT^{1/2}x(T-t)^{-3/2}\mathrm{e}^{-x^2/4(T-t)},$$

$$u_{xx} = -\frac{1}{2}AT^{1/2}(T-t)^{-1/2}\mathrm{e}^{-x^2/4(T-t)} + \frac{1}{4}AT^{1/2}x^2(T-t)^{-5/2}\mathrm{e}^{-x^2/4(T-t)} = -u_t$$

so that $u_t = -u_{xx}$.
Note that with this solution $|u(x,0)| \leq A$ and $u(0,t) \to \infty$ as $t \to T$.
Hence, given any $\varepsilon > 0$ and any position (X,T), the solution $u = \varepsilon T^{1/2}(T-t)^{-1/2}\mathrm{e}^{-(x-X)^2/4(T-t)}$, satisfying the initial condition $u(x,0) = \varepsilon \mathrm{e}^{-(x-X)^2/4T}$, for which $|u(x,0)| \leq \varepsilon$, becomes infinite as $(x,t) \to (X,T)$.
Because of this, the negative heat equation is ill-posed for $t > 0$ when subjected to initial conditions at $t = 0$; we can always find solutions that are arbitrarily small initially but that become infinite at any chosen time later on.
Suppose that the negative heat equation is subjected to *final conditions* $u(x,t_f) = g(x)$ at some time $t = t_f$, say. Then it is well posed for times *before* the final time, $t < t_f$.

2.9
The highest derivatives take the form $au_{tt} + bu_{tx} + cu_{xx}$ (or different subscripts for different independent variables):

(a) $u_t + u_{tx} - u_{xx} + u_x^2 = \sin u$: semilinear.

(b) $u_x + u_{xx} + u_y + u_{yy} = \sin(xy)$: linear and inhomogeneous.

(c) $u_x + u_{xx} - u_y - u_{yy} = \cos(xyu)$: semilinear.

(d) $u_{tt} + xu_{xx} + u_t = f(x,t)$: linear and inhomogeneous.

(e) $u_t + uu_{xx} + u^2u_{tt} - u_{tx} = 0$: quasilinear.

Exercises 3 Origins of PDEs

3.1
Since H does not depend on t, we can differentiate $h_t + Hv_x$ with respect to t:

$$h_{tt} + Hv_{xt} = 0 \quad \Rightarrow \quad h_{tt} + H\partial_x v_t = 0.$$

But $v_t = -gh_x$, so $\partial_x v_t = -gh_{xx}$ and $h_{tt} - c^2 h_{xx} = 0$, where $c^2 = gH$.

3.3
This simply requires the vector form of differentiation of a product:

$$\vec{\nabla} \cdot (\vec{v}T) = T\vec{\nabla} \cdot \vec{v} + \vec{v} \cdot \vec{\nabla}T.$$

For example, suppose that $\vec{v} = (u, v)$, then in $\mathbb{R}^2$,

$$\begin{aligned} \vec{\nabla} \cdot (\vec{v}T) &= (uT)_x + (vT)_y = u_x T + uT_x + v_y T + vT_y \\ &= T(u_x + v_y) + (uT_x + vT_y) = T\vec{\nabla} \cdot \vec{v} + \vec{v} \cdot \vec{\nabla}T. \end{aligned}$$

Exercises 4 Classification of PDEs

4.1
If $\varphi(x) = x/(1+|x|)$, then φ is a continuous function for $x \in \mathbb{R}$ and

$$\varphi(x) = \begin{cases} \frac{x}{1+x} & x \geq 0 \\ \frac{x}{1-x} & x < 0 \end{cases} \Rightarrow \varphi'(x) = \begin{cases} \frac{1}{(1+x)^2} & x \geq 0 \\ \frac{1}{(1-x)^2} & x < 0 \end{cases}.$$

Therefore

$$\lim_{x\to 0^-} \varphi'(x) = \lim_{x\to 0} \frac{1}{(1-x)^2} = 1 = \lim_{x\to 0} \frac{1}{(1+x)^2} = \lim_{x\to 0^+} \varphi'(x).$$

That is, $\varphi'(0^-) = \varphi'(0^+)$ showing that φ is continuously differentiable at the origin.

4.3
$u(x,t) = F(x-ct) + G(x+ct)$ and so the initial conditions give

$$\begin{aligned} u(x,0) = g_0(x) = &\quad F(x) + G(x) \\ u_t(x,0) = g_1(x) = &-cF'(x) + cG'(x) \end{aligned}$$

Integrating the second of these over the interval $(x-ct, x+ct)$ leads to

$$-F(x+ct) + F(x-ct) + G(x+ct) - G(x-ct) = \frac{1}{c}\int_{x-ct}^{x+ct} g_1(s)\,\mathrm{d}s$$

while, from the first,

$$\begin{aligned} -F(x+ct) + G(x+ct) &= g_0(x+ct) \\ -F(x-ct) + G(x-ct) &= g_0(x-ct). \end{aligned}$$

Combining these gives d'Alembert's formula (4.20).

4.5
The highest derivatives take the form $au_{tt} + bu_{tx} + cu_{xx}$ (or different subscripts for different independent variables):

(a) $u_t + u_{tx} - u_{xx} + u_x^2 = \sin u$. $b^2 - 4ac = 1 > 0$ so hyperbolic.

(b) $u_x + u_{xx} + u_y + u_{yy} = \sin(xy)$. $b^2 - 4ac = -4 < 0$ so elliptic.

(c) $u_x + u_{xx} - u_y - u_{yy} = \cos(xyu)$. $b^2 - 4ac = 4 > 0$ so hyperbolic.

(d) $u_{tt} + xu_{xx} + u_t = f(x,t)$. $b^2 - 4ac = -4x$ so: elliptic for $x > 0$, hyperbolic for $x < 0$, parabolic for $x = 0$.

(e) $u_t + uu_{xx} + u^2u_{tt} - u_{tx} = 0$. $b^2 - 4ac = 1 - 4u^3$ so elliptic for $u^3 > \frac{1}{4}$, hyperbolic for $u^3 < \frac{1}{4}$, parabolic for $u^3 = \frac{1}{4}$.

4.7
With the change of variables $s = x + y$, $t = x - y$, it follows from the previous answer that $u_{xx} - u_{yy} = x + y$ becomes $4u_{st} = s$. Integrating with respect to t gives $u_s = \frac{1}{4}st + f(s)$ (where f is an arbitrary function) then integrating with respect to s, $u = \frac{1}{8}s^2t + F(s) + G(t)$, where G

is an arbitrary function and $F(s) = \int_s f(s)\,\mathrm{d}s$ is also an arbitrary function. The general solution of $u_{xx} - u_{yy} = x + y$ is, therefore $u(x,y) = \frac{1}{8}(x+y)^2(x-y) + F(x+y) + G(x-y)$.
The boundary conditions give:

$$\begin{aligned} u(x,0) = &\quad \tfrac{1}{8}x^3 + F(x) + G(x) = x \\ u(0,y) = &-\tfrac{1}{8}y^3 + F(y) + G(-y) = -\tfrac{1}{2}y^3 \end{aligned}$$

and, when we replace y by x in the second of these, we find that $F(x) + G(-x) = -\frac{3}{8}x^3$ and so

$$G(x) - G(-x) = x + \tfrac{1}{4}x^3.$$

With the identity[1] $G(x) = \frac{1}{2}(G(x) - G(-x)) + \frac{1}{2}(G(x) + G(-x))$ we have

$$G(x) = \tfrac{1}{2}x + \tfrac{1}{8}x^3 + E(x)$$

where we have used $E(x)$ for the even part of G (which is, as yet, undetermined). Then

$$F(x) = \tfrac{1}{2}x - \tfrac{1}{4}x^3 - E(x)$$

so that the solution is (after some algebra)

$$u(x,y) = x(1-xy) - \tfrac{1}{2}(x+y)y^2 - E(x+y) + E(x-y)$$

and involves an arbitrary even function $E(\cdot)$.

4.9
Comparing the equation $2u_{xx} + 5u_{xt} + 3u_{tt} = 0$ with the template (4.12) we see that $a = 2$, b= 5/2 and $c = 3$ so $b^2 - ac = \frac{1}{4} > 0$, so the equation is hyperbolic. The factorisation

$$2u_{xx} + 5u_{xt} + 3u_{tt} = (2\partial_x + 3\partial_y)(\partial_x + \partial_y)u$$

suggests the change of variables $y = x - t$, $s = 3x - 2t$ and the chain rule gives

$$\left.\begin{aligned} \partial_x &= (\partial_x y)\partial_y + (\partial_x s)\partial_s = 3\partial_y + \partial_s \\ \partial_t &= (\partial_t y)\partial_y + (\partial_t s)\partial_s = \partial_y - \partial_s \end{aligned}\right\} \Rightarrow \begin{aligned} \partial_x + \partial_t &= 4\partial_y \\ \partial_x - 3\partial_t &= 4\partial_s. \end{aligned}$$

Hence $2u_{xx} + 5u_{xt} + 3u_{tt} = (2\partial_x + 3\partial_t)(\partial_x + \partial_t)u = 16u_{ys}$ which, on integrating twice gives the general solution $u = F(y) + G(s) = F(x-t) + G(3x-2t)$.
The initial conditions give

$$F(x) + G(3x) = 0, \quad -F'(x) - 2G'(3x) = x\mathrm{e}^{-x^2}$$

and integrating the second of these we find $F(x) + \frac{2}{3}G(3x) = \frac{1}{2}\mathrm{e}^{-x^2} + C$, where C is the constant of integration. It follows that

$$F(x) = \tfrac{3}{2}\mathrm{e}^{-x^2} + 3C, \quad G(3x) = -\tfrac{3}{2}\mathrm{e}^{-x^2} - 3C \Rightarrow G(x) = -\tfrac{3}{2}\mathrm{e}^{-x^2/9} - 3C.$$

Hence $u(x,y) = \frac{3}{2}\left(\mathrm{e}^{-(x-t)^2} - \mathrm{e}^{-(3x-2t)^2/9}\right)$ in which there is no arbitrary constant.

4.11
Substituting $u(r,t) = r^m f(t-r)$ into the PDE and collecting terms leads to

$$r^{m-2}f(t-r)m(m+n-2) + r^{m-1}f'(t-r)(2m+n-1) = 0.$$

[1] Every function can be written as the sum of an odd and an even function.

This will hold for all differentiable functions f if, and only if,

$$m(m+n-2)=0 \text{ and } 2m+n-1=0.$$

Thus, either $m=0$ and $n=1$ or $n=3$ and $m=-1$ giving the solutions $u(r,t)=f(t-r)$ when $n=1$ and $u(r,t)=f(t-r)/r$ when $n=3$ (see the previous question).

4.13
The change of independent variables $x=s\cos\alpha-t\sin\alpha$, $y=s\sin\alpha+t\cos\alpha$ gives

$$\left.\begin{aligned}\partial_s &= (\partial_s x)\partial_x+(\partial_s y)\partial_y = \quad \cos\alpha\partial_x+\sin\alpha\partial_y\\ \partial_t &= (\partial_t x)\partial_x+(\partial_t y)\partial_y = -\sin\alpha\partial_x+\cos\alpha\partial_y\end{aligned}\right\}$$

$$(\partial_s^2+\partial_t^2)u=(\cos\alpha\partial_x+\sin\alpha\partial_y)^2u+(-\sin\alpha\partial_x+\cos\alpha\partial_y)^2u=u_{xx}+u_{yy}$$

since α is constant and $\cos^2\alpha+\sin^2\alpha=1$.

4.15

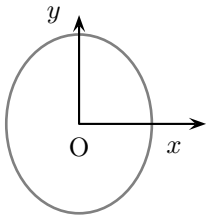


Figure 3: With $a=2$, $b=0$ and $c=3$ then $\widehat{Q}=\frac{1}{6}(3x^2+2y^2)$ in the solution (4.29). The figure shows a typical elliptical curve $\boldsymbol{x}^{\mathsf{T}}\widehat{Q}\boldsymbol{x}=$ constant.

4.17
With $G(x,y,t)=\frac{1}{4\pi}\left(\log(x^2+(y+t)^2)-\log(x^2+(y-t)^2)\right)$ we find

$$\partial_t G(x,y,t)=\frac{1}{4\pi}\left(\frac{2(y+t)}{(x^2+(y+t)^2}-\frac{-2(y-t)}{x^2+(y-t)^2}\right)$$

$$\partial_t G(x,y,t)|_{t=0}=\frac{1}{\pi}\left(\frac{y}{x^2+y^2}\right)=k(x,y)$$

from (4.32).
With the change of variables $s=x+y\tan\theta$ the interval $-\infty<s<\infty$ becomes $-\frac{1}{2}\pi<\theta<\frac{1}{2}\pi$ and $\mathrm{d}s=y\sec^2\theta\,\mathrm{d}\theta$ so that

$$\int_{-\infty}^{\infty}k(x-s,y)\,\mathrm{d}s=\int_{-\infty}^{\infty}\frac{y}{(x-s)^2+y^2}\,\mathrm{d}s=\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta=\pi,$$

where we have used the identity $1+\tan^2\theta=\sec^2\theta$.

4.19
The identity

$$\tan(a-b)=\frac{\tan a-\tan b}{1+\tan a\tan b}$$

with $\tan a=(x+2)/y$ and $\tan b=(x-2)/y$ gives, after a little manipulation,

$$\tan(a-b)=\frac{4y}{x^2+y^2-4}=\tan(\pi u)$$

and, in the limit $u \to 1/2$, we have $x^2 + y^2 \to 4$. In particular, when $y \to 0$ and $x \to \pm 2$ we see that $u(\pm 2, 0) = 1/2$. However, the boundary condition is discontinuous at $(\pm 2, 0)$: $g(2^-, 0) = 0$, $g(2^+, 0) = 1$, $g(-2^-, 1) = 0$, $g(-2^+, 0) = 0$ from which we see that

$$u(2,0) = \tfrac{1}{2}(g(2^-,0) + g(2^+,0)) \text{ and } u(-2,0) = \tfrac{1}{2}(g(-2^-,0) + g(-2^+,0)).$$

4.21
(a) From Exercise (4.22) with $u = F(r)$, we find $\nabla^2 u = \frac{1}{r}(rF'(r))'$ so $-\nabla^2 u = 1$ leads to

$$(rF'(r))' = -r \Rightarrow rF'(r) = -\tfrac{1}{2}r^2 + A \Rightarrow F'(r) = -\tfrac{1}{2}r + \frac{A}{r} \Rightarrow F(r) = -\tfrac{1}{4}r^2 + A\log r + B,$$

which is the general solution with arbitrary constants A and B. The boundary condition $F(1) = 0$ requires $B = \frac{1}{4}$ and, since $\log r \to -\infty$ as $r \to 0$, a bounded solution on any region contain the origin, requires $A = 0$. Thus $u(r, \theta) = \frac{1}{4}(1 - r^2)$.
(b) We look for a solution of the form $u(r, \theta) = F(r)\cos\theta$. Substituting into Laplace's equation gives

$$\nabla^2 u = \left(F''(r) + \frac{1}{r}F'(r) - \frac{1}{r^2}F(r)\right)\cos\theta = 0.$$

We try a solution in the form $F(r) = Ar^n$ and find $\nabla^2 u = A(n^2 - 1)\cos\theta$. This will be a solution if $n = \pm 1$. This gives two linearly independent solutions $Ar\cos\theta$ and $(B/r)\cos\theta$ but the second of these is unbounded when $r \to 0$. Therefore $u(r, \theta) = Ar\cos\theta$ and the boundary condition $u(1, \theta) = \cos\theta$ requires $A = 1$. Hence $u = r\cos\theta$, i.e., $u = x$.

4.23
Under the change of variables $z = x + \mathrm{i}y$, $z^* = x - \mathrm{i}y$, the chain rule gives

$$\partial_x = \partial_z + \partial_{z^*}, \qquad \partial_y = \mathrm{i}\partial_z - \mathrm{i}\partial_{z^*}$$

and so,

$$\begin{aligned} u_{xx} &= \quad (\partial_z + \partial_{z^*})^2 u = \quad u_{zz} + 2u_{zz^*} + u_{z^*z^*} \\ u_{yy} &= -(\partial_z - \partial_{z^*})^2 u = -(u_{zz} - 2u_{zz^*} + u_{z^*z^*}) \end{aligned}$$

giving $u_{xx} + u_{yy} = 4u_{zz^*}$. Integrating the PDE $\mathbf{u}_{zz^*} = 0$ with respect to z^* gives $u_z = f(z)$ (where f is an arbitrary function) then integrating with respect to z, $u = F(z) + G(z^*)$, where G is an arbitrary function and $F(s) = \int_s f(s)\,\mathrm{d}s$ is also an arbitrary function. The general solution of $u_{xx} + u_{yy} = 0$ is, therefore, $u(x, y) = F(x + \mathrm{i}y) + G(x - \mathrm{i}y)$.
If F is a real function then, since $\Re F(z) = \frac{1}{2}(F(z) + (F(z))^*) = \frac{1}{2}(F(z) + F(z^*)$. This will be a solution if we choose $G(s) = F(s)$. (The PDE is homogeneous, so $u = c(F(x + \mathrm{i}y) + G(x - \mathrm{i}y)$ is also a solution for any constant c.)
Since $\Im F(z) = \frac{1}{2}(F(z) - (F(z))^*) = \frac{1}{2}(F(z) - F(z^*)$. This will be a solution if we choose $G(s) = -F(s)$.

4.25
Since $-\nabla^2 u = v$ and $-\nabla^2 v = f$,

$$-\nabla^2\big(-\nabla^2 u\big) = -\nabla^2 v = f$$

and so $\nabla^4 u = f$.
With $u = Cr^2 \log(r)$ and $\nabla^2 u = (1/r)(ru_r)_r$

$$u_r = 2Cr \log r + Cr, \Rightarrow ru_r = 2Cr^2 \log r + Cr^2$$
$$(ru_r)_r = 4Cr \log r + 4Cr \Rightarrow \nabla^2 u = 4C \log r + 4C$$

so $v = -\nabla^2 u = -4C \log r - 4C$ which satisfies $\nabla^2 v = 0$ (see Exercise 4.20). Hence u satisfies the biharmonic equation.

Exercises 5 Boundary value problems in $\mathbb{R}^1$

5.1
First we look for a particular solution $u(x) = C$, where C is constant. Substituting into the ODE gives $-4C = 1$ and so $C = -1/4$.
Next we look for a solution of the homogeneous equation $u'' + 3u' - 4u = 0$ in the form $u = A\mathrm{e}^{\lambda x}$, where A and λ are constants. Substituting into the ODE gives

$$A\mathrm{e}^{\lambda x}(\lambda^2 + 3\lambda - 4) = 0.$$

Thus, either $A = 0$ (leading to the trivial solution $u(x) = 0$) or $\lambda^2 + 3\lambda - 4 = (\lambda + 4)(\lambda - 1) = 0$, that is, $\lambda = -4$ or $\lambda = 1$. The ODE is linear and homogeneous so the principle of superposition applies and the general solution of homogeneous ODE (often referred to as the complementary function) is $u = A\mathrm{e}^{-4x} + B\mathrm{e}^{x}$. The general solution of the given ODE is therefore

$$u(x) = A\mathrm{e}^{-4x} + B\mathrm{e}^{x} - \tfrac{1}{4}.$$

Applying the boundary conditions $u(0) = 1$ and $u(1) = 3$ leads to

$$\left.\begin{array}{l} A + B - \frac{1}{4} = 1 \\ A\mathrm{e}^{-4} + B\mathrm{e} - \frac{1}{4} = 3 \end{array}\right\} \Rightarrow 4A = \frac{5\mathrm{e} - 13}{\mathrm{e} - \mathrm{e}^{-4}}, \qquad 4B = \frac{13 - 5\mathrm{e}^{-4}}{\mathrm{e} - \mathrm{e}^{-4}}$$

which uniquely determine $u(x)$.

5.3
The key identities are

$$\cosh A - \cosh B = 2\sinh\tfrac{1}{2}(A+B)\sinh\tfrac{1}{2}(A-B), \quad \cos A - \cos B = -2\sin\tfrac{1}{2}(A+B)\sin\tfrac{1}{2}(A-B).$$

The first of these with $A = \frac{1}{2}\sqrt{b}$ and $B = \sqrt{b}(x - \frac{1}{2})$ allows us to rewrite (5.6) as

$$u(x) = -\frac{2\varepsilon}{b}\frac{\sinh(\frac{1}{2}\sqrt{b}x)\sinh\frac{1}{2}\sqrt{b}(1-x)}{\cosh\frac{1}{2}\sqrt{b}} = -2\frac{\varepsilon}{\cosh\frac{1}{2}\sqrt{b}}\left(\frac{\sinh\frac{1}{2}\sqrt{b}x}{\sqrt{b}}\right)\left(\frac{\sinh\frac{1}{2}\sqrt{b}(1-x)}{\sqrt{b}}\right)$$

Now as $z \to 0$, $\cosh z \to 1$, $\sinh(az)/z \to a$ (for any constant a) and consequently $u(x) \to -\frac{1}{2}\varepsilon x(1-x)$.
Similarly, (5.8) may be rewritten as

$$u(x) = -\frac{2\varepsilon}{b}\frac{\sin\frac{1}{2}\sqrt{|b|}x\sin\frac{1}{2}\sqrt{|b|}(1-x)}{\cos\frac{1}{2}\sqrt{|b|}} = -2\frac{\varepsilon}{\cos\frac{1}{2}\sqrt{|b|}}\left(\frac{\sin(\frac{1}{2}\sqrt{|b|}x)}{\sqrt{|b|}}\right)\left(\frac{\sin\frac{1}{2}\sqrt{|b|}(1-x)}{\sqrt{|b|}}\right)$$

and, since $\cos z \to 1$, $\sin(az)/z \to a$ as $z \to 0$ we again obtain $u(x) \to -\frac{1}{2}\varepsilon x(1-x)$.

5.5
It follows from the inverse monotonicity of $\mathscr{L}$ that there is a comparison function $\varphi(x)$ such that $\mathscr{L}\varphi \geq 1$. Then

$$\mathscr{L}u = \mathscr{F} \geq -\|\mathscr{F}\| = -\|\mathscr{F}\| \times 1 \geq -\|\mathscr{F}\| \times (\mathscr{L}\varphi)$$

and so, using the linearity of $\mathscr{L}$, $\mathscr{L}(u + \|\mathscr{F}\|\varphi) \geq 0$ which, by inverse monotonicity, implies that $u + \|\mathscr{F}\|\varphi) \geq 0$, as required.

5.7
The given equations imply $\dfrac{a_0}{b_0} = \dfrac{u'(0)}{u(0)} = \dfrac{v'(0)}{v(0)}$ from which $u'(0)v(0) - u(0)v'(0)) = 0$ follows.

5.9
Multiplying both sides of $-2xu'' - u' = 2f(x)$ by $\frac{1}{2}x^{-1/2}$ we obtain

$$-x^{1/2}u'' - \tfrac{1}{2}x^{-1/2}u' = x^{-1/2}f(x) \Rightarrow -(2x^{1/2}u')' = x^{-1/2}f(x)$$

which has the form (5.28) with $p(x) = x^{1/2}$, $q(x) = 0$ and $g(x) = x^{-1/2}f(x)$. The change of independent variable $\xi = \xi(x)$ given by (5.30), that is,

$$\xi = \int_0^x \frac{1}{p(s)}\,\mathrm{d}s = 2x^{1/2}$$

then leads to (5.31), i.e.,

$$-\frac{\mathrm{d}^2u}{\mathrm{d}\xi^2} = \tilde{g}(\xi) = p(x)g(x) = 2f(x) = 2f(\tfrac{1}{4}\xi^2).$$

in which $\tilde{q}(\xi) = 0$. This boundary value problem with Dirichlet boundary conditions may be written in the form $\mathscr{L}u = \mathscr{F}$, where $\mathscr{L}$ is identical to the operator defined in Example 5.2 except that the independent variable is ξ rather than x. The arguments used there establish that this equation has a unique solution.

5.11
With $\langle u, v\rangle = \int_0^L u(x)v^*(x)\,\mathrm{d}x$:

(a) $\langle v, u\rangle^* = \Big(\displaystyle\int_0^L v(x)u^*(x)\,\mathrm{d}x\Big)^* = \int_0^L v^*(x)u(x)\,\mathrm{d}x = \langle u, v\rangle$ since $(u^*)^* = u$.

(b) $\langle u, u\rangle = \displaystyle\int_0^L |u(x)|^2\,\mathrm{d}x \geq 0$ since the integrand is non-negative.

(c) Suppose that $\langle u, u\rangle = 0$ but that $u(x)$ is not identically zero. There must therefore be point $a \in (0, L)$ where $u(a) \neq 0$. By continuity there must be an $\varepsilon > 0$ such that u is non-zero in the interval $(a-\varepsilon, a+\varepsilon)$ so $\langle u, u\rangle = \int_{a-\varepsilon}^{a+\varepsilon} |u(x)|^2\,\mathrm{d}x > 0$ giving a contradiction.

(d) $\langle c_1u_1 + c_2u_2, v\rangle = \displaystyle\int_0^L (c_1u_1 + c_2u_2)v^*\,\mathrm{d}x = c_1\int_0^L u_1v^*\,\mathrm{d}x + c_2\int_0^L u_2v^*\,\mathrm{d}x$ by the property of integrals and so $\langle c_1u_1 + c_2u_2, v\rangle = c_1\langle u_1, v\rangle + c_2\langle u_2, v\rangle$.

(e) Suppose that u and v are orthogonal with respect to the inner product $\langle u, v\rangle$ and are linearly dependent. Thus $\langle u, v\rangle = 0$ and there are non-zero constants a, b such that $au(x) + bv(x) = 0$. However, since $v(x) = -(a/b)u(x)$,

$$\langle u, v\rangle = \langle u, -(a/b)u(x)\rangle = -(a/b)u(x)\langle u, u\rangle \neq 0$$

(having used properties (d) with $c_1 = -b/a$, $c_2 = 0$ and (c) 5) giving a contradiction.

5.13
$\phi(x) = \sin 2\pi x$ satisfies the differential equation and boundary conditions.
Multiplying $-u'' - 4\pi^2 u = f(x)$ by ϕ and integrating over $(0,1)$, then integrating by parts twice, gives

$$\begin{aligned}\int_0^1 \phi(x) f(x)\,\mathrm{d}x &= \int_0^1 \phi(x)(-u'' - 4\pi^2 u)\,\mathrm{d}x \\ &= -\phi(x)u'(x)\Big|_0^1 + \phi'(x)u(x)\Big|_0^1 + \int_0^1 u(x)(-\phi'' - 4\pi^2\phi)\,\mathrm{d}x = 0.\end{aligned}$$

The ODE cannot be satisfied with the given BCs unless $\int_0^1 \phi(x) f(x)\,\mathrm{d}x = 0$.
This condition is clearly violated when $f(x) = \phi(x) = \sin 2\pi x$. The ODE $-u'' - 4\pi^2 u = \sin(2\pi x)$ has the general solution

$$u(x) = x\frac{\cos 2\pi x}{4\pi} + A\sin 2\pi x + B\cos 2\pi x$$

and applying the BCs we find $u(0) = B = 0$ and $u(1) = \frac{1}{4\pi} + B = 0$ which are contradictory. There cannot therefore be a solution.
When $f(x) = 1$ we find $\int_0^1 \phi(x) f(x)\,\mathrm{d}x = 0$. In this case the ODE $-u'' - 4\pi^2 u = 1$ has the general solution

$$u(x) = -\frac{1}{4\pi^2} + A\sin 2\pi x + B\cos 2\pi x$$

and, applying the BCs we find $u(0) = -\frac{1}{4\pi^2} + B$ and $u(1) = -\frac{1}{4\pi^2} + B$
Thus,

$$u(x) = -\tfrac{1}{4\pi^2} + A\sin 2\pi x - \tfrac{1}{4\pi^2}\cos 2\pi x$$

satisfies the differential equation and boundary conditions for any choice of constant A so we have a non-unique solution: the solution is unique up to an arbitrary multiple of $\phi(x)$.

5.15
(a) The standard argument shows that ϕ cannot satisfy the boundary conditions unless $\lambda > 0$. So, with $\lambda = \mu^2$, the general solution of the ODE is $\phi(x) = A\sin\mu x + B\cos\mu x$. The boundary condition $\phi(0) = 0$ requires $B = 0$ and $\phi'(1) = 0$ requires $A\cos\mu = 0$. Since $A \neq 0$ (since it would lead to a trivial solution), it follows that $\lambda = \lambda_n := (n - \frac{1}{2})^2\pi^2$, $n = 1, 2, \ldots$, with corresponding eigenfunctions $\phi_n(x) = \sin(n - \frac{1}{2})\pi x$.
(b) Suppose that $\omega^2 = \lambda_n$ for some value of n. If $\phi(x)$ is the corresponding eigenfunction, then

$$\int_0^1 \phi(x)(-u''(x) + \omega^2 u)\,\mathrm{d}x = \int_0^1 \phi(x) f(x)\,\mathrm{d}x.$$

Integrating the left hand side by parts twice and using the boundary conditions on both u and ϕ, we find

$$\begin{aligned}-\phi(x)u'(x)\Big|_0^1 + \int_0^1 (\phi'(x)u'(x) + \omega^2\phi u)\,\mathrm{d}x &= \int_0^1 \phi(x) f(x)\,\mathrm{d}x, \\ (-\phi(x)u'(x) + \phi'(x)u(x))\Big|_0^1 + \int_0^1 (-\phi''(x) + \omega^2\phi)u(x)\,\mathrm{d}x &= \int_0^1 \phi(x) f(x)\,\mathrm{d}x, \\ 2\phi(1) - \phi'(0) &= \int_0^1 \phi(x) f(x)\,\mathrm{d}x,\end{aligned}$$

since $-\phi''(x) + \omega^2\phi = 0$. The data for the problem are inconsistent unless this condition is satisfied.
When $\omega^2 = \lambda_1 = \frac{1}{2}\pi$, then $\phi(x) = \sin\frac{1}{2}\pi x$ and, with $f(x) = c$, we find that $c = \frac{1}{2}\pi(2 - \frac{1}{2}\pi)$. The general solution of $-u''(x) = \omega^2 u(x) + c$ is $u(x) = A\sin\frac{1}{2}\pi x + B\cos\frac{1}{2}\pi x - c/\omega^2$. Applying the BCs:

$$u(0) = 1 = B - c/\omega^2, \qquad u'(1) = -2 = -\tfrac{1}{2}\pi B$$

both of which give the same value $B = 4/\pi$ because the value of c was carefully chosen. The solution is, therefore, $u(x) = 1 + A\sin\frac{1}{2}\pi x + (4/\pi)\big(\cos(\frac{1}{2}\pi x) - 1\big)$ which is unique up to an arbitrary multiple of ϕ.

5.17
When $u(x) = M(x)w(x)$ we find

$$u'(x) = M'(x)w(x) + M(x)w'(x), \quad u''(x) = M''(x)w(x) + 2M'(x)w'(x) + M(x)w''(x)$$

so that

$$u'' - au' - bu = Mw'' + (2M' - aM)w' + (M'' - aM' - bM)w.$$

The coefficient of w' can be made to vanish by choosing M such that $2M' = aM$ in which case $2M'' = a'M + aM' = (a' + \frac{1}{2}a^2)M$ and

$$M'' - aM' - bM = -\tfrac{1}{2}(2b + \tfrac{1}{2}a^2 - a')M.$$

Thus, $u'' - au' - bu = f$ becomes

$$-w'' + Q(x)w(x) = G(x),$$

where $Q(x) = -(M'' - aM' - bM)/M = \frac{1}{2}(2b + \frac{1}{2}a^2 - a')$, $G(x) = f(x)/M(x)$ and $M(x) = A\exp(\int^x a(s)\,\mathrm{d}s)$.

5.19
With $\phi_n(x) = \mathrm{e}^{2\pi\mathrm{i}nx/L}$ then, for $m \neq n$,

$$\begin{aligned}\big\langle \phi_n, \phi_m \big\rangle &= \int_0^L \mathrm{e}^{2\pi\mathrm{i}nx/L}\mathrm{e}^{-2\pi\mathrm{i}mx/L}\,\mathrm{d}x = \int_0^L \mathrm{e}^{2\pi\mathrm{i}(n-m)x/L}\,\mathrm{d}x \\ &= \frac{L}{2\pi\mathrm{i}(n-m)}\mathrm{e}^{2\pi\mathrm{i}(n-m)x/L}\Big|_0^L = \frac{L}{2\pi\mathrm{i}(n-m)}\Big(\mathrm{e}^{2\pi\mathrm{i}(n-m)} - 1\Big) = 0\end{aligned}$$

since $n - m$ is an integer and so $\mathrm{e}^{2\pi\mathrm{i}(n-m)} = 1$. Also, when $m = n$,

$$\big\langle \phi_n, \phi_n \big\rangle = \int_0^L \mathrm{e}^{2\pi\mathrm{i}nx/L}\mathrm{e}^{-2\pi\mathrm{i}nx/L}\,\mathrm{d}x = \int_0^L \mathrm{d}x = L.$$

5.21
Suppose that the eigenvalue problem is defined by $\mathcal{L}u = \lambda u$ with boundary conditions $\mathcal{B}u = 0$. Let $v = cu$, where c is a constant. By the linearity of $\mathcal{L}$ and $\mathcal{B}$, $\mathcal{L}v = \mathcal{L}(cu) = c\mathcal{L}u = c\lambda u = \lambda v$ and $\mathcal{B}(v) = \mathcal{B}(cu) = c\mathcal{B}u = 0$. Thus, v satisfies the same equations as u: $\mathcal{L}v = \lambda v$ and $\mathcal{B}v = 0$.

5.23
With $\phi_n = \cos(n-\frac{1}{2})x$ for $0 < x < \pi$ and $m \neq n$,

$$\begin{aligned}\langle \phi_n, \phi_m \rangle &= \int_0^\pi \cos(n-\tfrac{1}{2})x \cos(m-\tfrac{1}{2})x \, \mathrm{d}x \\ &= \tfrac{1}{2}\int_0^\pi (\cos(n+m-1)x + \cos(n-m)x) \, \mathrm{d}x \\ &= \Big(\frac{\sin(n+m-1)x}{2(n+m-1)} + \frac{\sin 2(n-m)x}{(n-m)}\Big)\Big|_0^\pi = 0.\end{aligned}$$

5.25
When $\lambda \leq 4$ the argument from the previous exercise is easily adapted to show that only trivial solutions are possible for the ODE $-u'' + 4u = \lambda u$ with BCs $u(0) = u(\pi) = 0$. When $\lambda > 4$ we have the general solution

$$u(x) = A \sin\sqrt{\lambda - 4}x + B\cos\sqrt{\lambda-4}x$$

and the BCs give

$$u(0) = 0 = B \text{ and } u(\pi) = 0 = A\sin\sqrt{\lambda-4}\pi = 0.$$

Since A cannot be zero (this would imply that $u(x) = 0$, the trivial solution) so λ must satisfy $\sin\sqrt{\lambda-4}\pi = 0$. Hence $\sqrt{\lambda-4}\pi = n\pi$, $n = \pm 1, \pm 2, ...$ giving $\lambda_n = 4 + n^2$ with corresponding eigenfunction $u_n(x) = \sin n\pi x$, $n = 1, 2, 3, ..$ (we do not include negative values of n, since $\sin(-n\pi x) = -\sin(n\pi x)$ and we would have linearly dependent eigenfunctions).

5.27
With $u = Mw$ the ODE $-x^2u'' + 2xu' - 2u = \lambda x^2 u$ becomes (see Exercise 5.17)

$$\begin{aligned}-x^2\big(M'' - aM' - bM\big) + 2x\big(M'w + Mw'\big) - 2Mw &= \lambda x^2 Mw \\ -x^2Mw'' + (2xM - 2x^2M')w' + (-x^2M'' + 2xM' - 2M)w &= \lambda x^2 Mw\end{aligned}$$

and the coefficient of w' vanishes when $M - xM' = 0$. Therefore we may choose $M(x) = x$ and the eigenvalue problem becomes $-w'' = \lambda w$. Since $u' = w + xw'$, the BC $u'(0) = 0$ becomes $w(0) = 0$ and $u(1) = u'(1)$ becomes $w'(1) = 0$.
The standard argument can be applied to show that $\lambda = 0$ and $\lambda < 0$ both lead to trivial solutions. For the case $\lambda > 0$, let $\lambda = \mu^2$ then $-w'' = \mu^2 w$ has general solution $w = A\sin\mu x + B\cos\mu x$. The BC $w(0) = 0$ implies $B = 0$ and $w'(1) = 0$ then leads to $A\mu\cos\mu = 0$. Choosing $A = 0$ gives immediately the trivial solution, choosing $\mu = 0$ leads to the earlier case $\lambda = 0$ so we are left with $\cos\mu = 0$. Therefore $\mu = (n - \frac{1}{2})\pi$, $n = 1, 2, \ldots$ and the eigenvalues are $\lambda_n = (n-\frac{1}{2})^2\pi^2$ with corresponding eigenfunction $w_n = \sin(n-\frac{1}{2})\pi x$, i.e., $u_n(x) = x\sin(n-\frac{1}{2})\pi x$, $n = 1, 2, \ldots$.

5.29
We shall work from first principles. Multiplying the differential equation $-u''(x) = \lambda w(x)u(x)$ by $u^*(x)$ and integrating by parts gives, on applying the boundary conditions,

$$\begin{aligned}\int_0^1 -u''(x)u^*(x)\,\mathrm{d}x &= \lambda\int_0^1 w(x)u(x)u^*(x)\,\mathrm{d}x \\ -u'(x)u^*(x)\Big|_0^1 + \int_0^1 u'(x)(u'(x))^*\,\mathrm{d}x &= \lambda\int_0^1 w(x)|u(x)|^2\,\mathrm{d}x \\ |u(1)|^2 + \int_0^1 |u'(x)|^2\,\mathrm{d}x &= \lambda\int_0^1 w(x)|u(x)|^2\,\mathrm{d}x\end{aligned}$$

and so

$$\lambda = \frac{|u(1)|^2 + \int_0^1 |u'(x)|^2}{\int_0^1 w(x)|u(x)|^2 \,\mathrm{d}x}$$

in which both numerator and denominator are real and positive.

5.31
This requires on the insertion of a factor $w(x)$ in each of the integrands in the solution of Exercise 5.11.

5.33
Choosing $u = \phi_m(x)$ and $v = \phi_n(x)$ so that

$$\mathcal{L}u = \lambda_m w\phi_m \text{ and } \mathcal{L}v = \lambda_n w\phi_n$$

then by Lagrange's identity (5.25) $\int_0^1 \left(v^*\mathcal{L}u - u\mathcal{L}v^*\right) \mathrm{d}x = 0$. However, the eigenfunctions are real by Exercise 5.30, so

$$\begin{aligned} 0 = \int_0^1 \left(v^*\mathcal{L}u - u\mathcal{L}v^*\right) \mathrm{d}x &= \int_0^1 \left(\phi_n\lambda_m w\phi_m - \phi_m\lambda_n w\phi_n\right) \mathrm{d}x \\ &= (\lambda_m - \lambda_n)\left\langle \phi_n, \phi_m \right\rangle_w \end{aligned}$$

and therefore $\left\langle \phi_n, \phi_m \right\rangle_w = 0$ provided $\lambda_m \neq \lambda_n$.

Exercises 6 Finite difference methods in $\mathbb{R}^1$

6.1
Using Taylor series expansions with remainder terms, (6.13) with $x = x_m$ becomes

$$v(x_m \pm h) = v(x_m) \pm hv'(x_m) + \tfrac{1}{2}h^2v''(x_m) \pm \tfrac{1}{6}h^3v'''(x_m) + \tfrac{1}{24}h^4v''''(\xi_m^{\pm}),$$

where $x_m - h < \xi_m^- < x_m < \xi_m^+ < x_m + h$. Adding these series together we obtain

$$v(x_{m+1}) + v(x_{m-1}) = 2v(x_m) + h^2v''(x_m) + \tfrac{1}{24}h^4\big(v''''(\xi_m^-) + v''''(\xi_m^+)\big).$$

but, by the Intermediate Value Theorem, there must be a point $\xi_m \in (\xi_m^-, \xi_m^+)$ such that $\frac{1}{2}(v''''(\xi_m^-) + v''''(\xi_m^+)\big) = v''''(\xi_m)$. Consequently, on rearranging,

$$v''(x_m) = h^{-2}\big(v(x_{m+1}) - 2v(x_m) + v(x_{m-1})\big) - \tfrac{1}{12}h^2v''''(\xi_m).$$

6.3

$$\begin{aligned}\triangle^+\triangle^+v_m &= \triangle^+(\triangle^+v_m) = \triangle^+(v_{m+1} - v_m) = \triangle^+v_{m+1} - \triangle^+v_m \\ &= (v_{m+2} - v_{m+1}) - (v_{m+1} - v_m) = v_{m+2} - 2v_{m+1} + v_m = \delta^2v_{m+1}.\end{aligned}$$

Then solution to Exercise 6.1 gives

$$h^{-2}\triangle^+\triangle^+v_m = h^{-2}\delta^2v_{m+1} = v''_{m+1} + \mathcal{O}(h^2)$$

but $v''_{m+1} = v''(x_m + h) = v''_m + hv'''_m + \mathcal{O}(h^3) = v''_m + \mathcal{O}(h)$, Hence, $h^{-2}\triangle^+\triangle^+v_m = v''_m + \mathcal{O}(h)$. The corresponding results for $\triangle^-\triangle^-v_m$ are:

$$h^{-2}\triangle^-\triangle^-v_m = h^{-2}\delta^2v_{m-1} = v''_{m-1} + \mathcal{O}(h^2) = v''_m + \mathcal{O}(h).$$

6.5
At $m = 0$, the equation $\mathscr{L}_hU_m = \mathscr{F}_{h,m}$ gives $U_0 = \alpha$.
For $0 < m < M$,

$$-a_mU_{m-1} + b_mU_m - c_mU_{m+1} = f_m$$

and at $m = M$, $U_M = \beta$. With $\boldsymbol{u} = [U_1, U_2, \ldots, U_{M-1}]^{\mathsf{T}}$, these equations may be combined to give $A\boldsymbol{u} = \boldsymbol{f}$, where

$$A = \frac{1}{h^2}\begin{bmatrix} b_1 & -c_1 & & & \\ -a_2 & b_2 & -c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -a_{M-2} & b_{M-2} & -c_{M-2} \\ & & & -a_{M-1} & b_{M-1} \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} f_1 + \alpha a_1 \\ f_2 \\ \vdots \\ f_{M-2} \\ f_{M-1} + \beta c_{M-1} \end{bmatrix}.$$

6.7
$a_m = h^{-2}$, $b_m = 2h^{-2} + x_m^2$, $c_m = h^{-2}$, $d_m = x_m$ for $m = 1, 2, \ldots, M-1$.

Referring to Definition 6.9, these coefficients are all positive and $b_m = a_m + c_m + x_m^2 \geq 0$ and the corresponding operator $\mathcal{L}_h$ is of positive type.

6.9
Since $\Phi(x)$ is a quadratic function of x, $h^{-2}\delta^2\Phi_m = \Phi''(x_m) = 2c$. Also,

$$\Phi_M - \Phi_{M-1} = \Phi(1) - \Phi(1-h) = h\Phi'(1) - \tfrac{1}{2}h^2\Phi''(1) = -ch(1-a) + ch^2$$

Consequently,

$$-h^{-2}\delta^2\Phi_m = -2c \text{ and } h^{-1}(\Phi(1) - \Phi(1-h)) = -c(1-a) + ch.$$

Since $h \leq \frac{1}{2}$ we choose $a = \frac{1}{4}$ so that $1 - a - h = \frac{3}{4} - h \geq \frac{1}{4}$ and therefore $c(1-a) - ch \geq 1$ with $c = -4$. Then $-h^{-2}\delta^2\Phi_m = -2c = 8 \geq 1$. Thus $\Phi(x) = x(4-x)$ is a possible comparison function for Exercise 6.8.
When $U_m = \frac{1}{2}x_m(3 - x_{m-1})$ we find $U_0 = 0$, $-h^{-2}\delta^2 U_m = 1$ and $2h^{-1}(U_M - U_{M-1}) = 1$ and so U satisfies exactly the finite difference equations from Exercise 6.8. This is the only solution since $\mathscr{L}_h$ is inverse monotone.
The general solution of the differential equation $-u''(x) = 1$ is $u = A + Bx - \frac{1}{2}x^2$. The BC $u(0) = 0$ implies $A = 0$ while $u'(1) = \frac{1}{2}$ requires $B = 3/2$. Thus $u(x) = \frac{1}{2}x(3-x)$ and the global error is

$$E_m = u(x_m) - U_m = \tfrac{1}{2}x_m(x_m - x_{m-1}) = \tfrac{1}{2}hx_m.$$

and so $\|E\|_{h,\infty} = \mathcal{O}(h)$—convergence is at a first order rate.

6.11
The ODE is approximated by $\mathcal{L}_h U_m = f_m$ for $m = 1, 2, \ldots, M-1$, where

$$\begin{aligned}\mathcal{L}_h U_m &:= -h^{-2}\delta^2 U_m + 20h^{-1}\triangle U_m \\ &= -h^{-2}(1+10h)U_{m-1} + 2h^{-2}U_m - h^{-2}(1-10h)U_{m+1}\end{aligned}$$

and $f_m = (mh)^2$. Comparing the coefficients of $\mathcal{L}_h$ with those of Definition 6.9, $a_m = h^{-2}(1 + 10h) \geq 0$ for all $h > 0$, $c_m = h^{-2}(1-10h) \geq 0$ for $h \leq 1/10$ and $b_m = 2h^{-2} = a_m + c_m$. Hence $\mathcal{L}_h$ is a positive type operator for $h \leq 1/10$, i.e., $M \geq 10$.

With $\varphi(x) = Ax + B$ and $\mathcal{L}u(x) = -u''(x) + 20u'(x)$ we find $\mathcal{L}\varphi(x) = 20A \geq 1$ if $A \geq 1/20$. Also $\varphi(0) \geq 1$ and $\varphi(1) \geq 1$ if $B \geq 1$ and $A + B \geq 1$, respectively. All these conditions can be met by choosing $A = 1/20$ and $B = 1$ so $\varphi(x) = 1 + x/20$ is a comparision function for $\mathcal{L}$ with Dirichlet BCs. Since φ is a linear function, then $\mathcal{L}_h\varphi(x_m) \geq 1$ so it also acts as a comparison function for $\mathcal{L}_h$. $C = \max_{0\leq x\leq 1}\varphi(x) = 21/20$ and therefore $\mathcal{L}_h$ with Dirichlet BCs is stable by Lemma 6.8 for $h \leq 1/10$.

6.13
The local truncation error is $\mathcal{R}_h = \mathcal{L}_h U - \mathcal{F}_h$, where $\mathcal{L}_h U_m := -\varepsilon h^{-2}\delta^2 U_m + 2h^{-1}\triangle^- U_m$ and $\mathcal{F}_{h,m} = f_m$. Using the results in Table 6.1,

$$\begin{aligned}\mathcal{R}_{h,m} &= -\varepsilon h^{-2}\delta^2 m_m + 2h^{-1}\triangle^- u_m - f_m \\ &= -\varepsilon\big(u_m'' + \tfrac{1}{12}h^2 u_m'''' + \mathcal{O}(h^4)\big) + 2\big(u_m' + \tfrac{1}{2}hu_m'' + \mathcal{O}(h^2)\big) - f_m \\ &= \big(-\varepsilon u_m'' + 2u_m' - f_m\big) + hu_m'' + \mathcal{O}(h^2) = \mathcal{O}(h)\end{aligned}$$

since $-\varepsilon u'' + 2u' = f$. The order of consistency is therefore first order. Also

$$\mathcal{L}_h U_m = -(\varepsilon h^{-2} + 2h)U_{m-1} + (2\varepsilon h^{-2} + h)U_m - \varepsilon h^{-2} U_{m+1}$$

so, according to Definition 6.9, is of positive type for all $\varepsilon > 0$ and $h > 0$. The comparison function $\varphi(x) = \frac{1}{2}x + 1$, being linear in x, also satisfies $\mathcal{L}_h\varphi(x_m) \geq 1$. Therefore $\mathcal{L}_h$ is stable by Lemma 6.8 (with stability constant $C = 3/2$). Finally, convergence at a first order rate is a consequence of Theorem 6.7.

6.15
From (6.51) the local truncation error $\mathscr{R}_h = \widehat{\mathscr{L}_h}u - \widehat{\mathscr{F}_h}$ and, since $\triangle^- u_m = hu'_m - \frac{1}{2}h^2 u''_m + \mathcal{O}(h^3)$ (see Table 6.1),

$$\begin{aligned}
\mathscr{R}_{h,M} &= (a + \tfrac{1}{2}bhr_M)u_M + bh^{-1}\triangle^- u_M - (\beta + \tfrac{1}{2}bhf_M) \\
&= (a + \tfrac{1}{2}bhr(1))u(1) + b(u'(1) - \tfrac{1}{2}hu''(1) + \mathcal{O}(h^2)) - (\beta + \tfrac{1}{2}bhf(1)) \\
&= au(1) + bu'(1) - \beta + \tfrac{1}{2}hb\big[-u''(1) + r(1)u(1) - f(1) + \mathcal{O}(h^2)\big]
\end{aligned}$$

and so $\mathscr{R}_{h,M} = \mathcal{O}(h^2)$ since $au(1) + bu'(1) = \beta$ and $-u''(1) + r(1)u(1) = f(1)$. Had we used $\triangle^- u_m = hu'_m - \frac{1}{2}h^2 u''_m + \frac{1}{6}h^3 u'''(\xi_m)$, where $x_m - h < \xi_m < x_m$, we would have found that $\mathscr{R}_{h,M} = bh^2 u'''(\xi_m)$. In both cases the local truncation error is of second order.

6.17
The leading term in the local truncation error from the previous question is $-\frac{1}{2}bhu''(0)$ and, using the ODE $-u'' + ru = f$ at $x = 0$, this becomes $\frac{1}{2}bh(f(0) - r(0)u(0))$. Thus, subtracting this term from the left of the BC gives the modified condition

$$aU_0 - bh^{-1}\triangle^+ U_0 - \tfrac{1}{2}bh(f_0 - r_0 U_0) = \alpha.$$

The corresponding local truncation error is

$$\begin{aligned}
\mathscr{R}_{h,M} &= au_0 - bh^{-1}\triangle^+ u_0 - \tfrac{1}{2}bh(f_0 - r_0 u_0) - \alpha \\
&= au_0 - b\big(u'(0) + \tfrac{1}{2}hu''(0) + \mathcal{O}(h^2)\big) - \tfrac{1}{2}bh(f_0 - r_0 u_0) - \alpha \\
&= \big(au(0) - bu'(0) - \alpha\big) - \tfrac{1}{2}h\big(-u''(0) + r_0 u_0 - f_0\big) + \mathcal{O}(h^2) = \mathcal{O}(h^2)
\end{aligned}$$

since $au(0) - bu'(0) = \alpha$ and $-u''(0) + r_0 u_0 = f_0$.

6.19
The proof of Theorem 6.10 can be used to prove that U_m cannot have a negative minimum for $1 \leq m \leq M - 1$. It remains to prove that U_m cannot have a negative minimum for either $m = 0$ or $m = M$.
We begin with the left end-point. Suppose that, contrary to the statement of the theorem, U_m has a negative minimum at $m = 0$ so that $U_1 \geq U_0$ (which implies $-c_0 U_1 \leq -c_0 U_0$) and

$$\mathscr{L}_h U_0 = b_0 U_0 - c_0 U_1 \leq (b_0 - c_0)U_0 \leq 0.$$

If this inequality were strict (because $b_0 > c_0$) it would contradict the assumption $\mathscr{L}_h U \geq 0$ and prove that a negative minimum at $m = 0$ could not occur. Suppose therefore that equality holds which means that $b_0 = c_0$ and $U_1 = U_0 < 0$.
The argument at the right end-point is essentially the same—either a negative minimum cannot occur at $m = M$ or $b_M = c_M$ and $U_{M-1} = U_M < 0$.

We now turn to the interior grid points. Suppose that, contrary to the statement of the theorem, U_m has a negative minimum at m, where $0 < m < M$. The argument in Theorem 6.10 proves that either there is a contradiction of $\mathscr{L}_h U_m \geq 0$ or both $U_{m-1} = U_m = U_{m+1} < 0$ and $b_m = a_m + c_m$. Combining all three cases we see that either there is a contradiction of $\mathscr{L}_h U_m \geq 0$ or the same negative minimum is attained at *all* grid points *and* $b_m = a_m + c_m$ for all m (recall $a_0 = c_M = 0$). However, this last possibility is ruled out by Definition 6.17 which requires that $b_m > a_m + c_m$ for at least one value of m.

6.21
We find that

$$\begin{aligned} a_0 = 0, b_0 = c_0 = 1 \qquad & b_0 = a_0 + c_0 \\ a_m = 1, b_m = 2, c_m = 1 \qquad & b_m = a_m + c_m, \quad 0 < m < M \\ a_M = b_M = 1, c_M = 0 \qquad & b_M = a_M + c_M \end{aligned}$$

so that $b_m = a_m + c_m$ for *all* m and we do not have strict inequality for any m. Thus $\mathscr{L}_h$ is not of positive type.
If $V_m = C$ for $m = 0, 1, \ldots, M$ then $\triangle^+ V_0 = 0$, $\delta^2 V_m = 0$ for $m = 1, 2, \ldots, M-1$ and $\triangle^- V_M = 0$. Consequently $\mathscr{L}_h V_m = 0$ for $m = 0, 1, \ldots, M$. By linearity of $\mathscr{L}_h$,

$$\mathscr{L}_h(U + V) = \mathscr{L}_h U + \mathscr{L}_h V = \mathscr{F}_h$$

so, if a solution exists then there are infinitely many solutions.
With $\mathscr{F}_h$ given by (6.22), the equation $\mathscr{L}_h U = \mathscr{F}_h$ corresponds to

$$\mathscr{L}_h U_m = \begin{cases} -h^{-1}(U_1 - U_0), \\ -h^{-2}\delta^2 U_m, \\ \quad h^{-1}(U_M - U_{M-1}) \end{cases}, \quad \mathscr{F}_h = \begin{cases} \alpha, & m = 0, \\ f_m, & m = 1, 2, \ldots, M-1, \\ \beta, & m = M, \end{cases}$$

which is consistent with the BVP $-u''(x) = f(x)$ $(0 < x < 1)$ with BCs $-u'(0) = \alpha$, $u'(1) = \beta$. If we integrate this ODE over the interval $(0, 1)$ and apply the BCs, we find

$$-u'(1) + u'(0) = \int_0^1 f(x)\,\mathrm{d}x \Rightarrow -\beta - \alpha = \int_0^1 f(x)\,\mathrm{d}x.$$

Thus, no solution is possible unless the data satisfy this compatibility condition. When they do, $u(x) + c$ is a solution for any constant c whenever $u(x)$ is a solution.
Using the identity $\delta^2 = \triangle^+ \triangle^-$ (see Exercise 6.2) we see that

$$\begin{aligned} \sum_{m=1}^{M-1} \delta^2 U_m &= \sum_{m=1}^{M-1} \triangle^+(\triangle^- U_m) \\ &= (\triangle^- U_M - \triangle^- U_{M-1}) + (\triangle^- U_{M-1} - \triangle^- U_{M-2}) + \cdots + (\triangle^- U_2 - \triangle^- U_1) \\ &= \triangle^- U_M - \triangle^+ U_0 \end{aligned}$$

by virtue of a telescoping series and $\triangle^- U_1 = U_1 - U_0 = \triangle^+ U_0$.
Summing the equations $\mathscr{L}_h U_m = \mathscr{F}_{h,m}$ over $m = 1, 2, \ldots, M-1$ gives, using the BCs $-h^{-1}\triangle^+ U_0 = \alpha$ and $h^{-1}\triangle^- U_M = \beta$,

$$\begin{aligned} -h^{-2}\left[\triangle^- U_M - \triangle^+ U_0\right] &= \sum_{m=1}^{M-1} f_m \\ -h^{-1}(\beta + \alpha) = \sum_{m=1}^{M-1} f_m \quad &\Rightarrow \quad -\beta - \alpha = h \sum_{m=1}^{M-1} f_m \end{aligned}$$

which is a discrete analogue of the compatibility condition found earlier for the ODE. Since

$$h\sum_{m=1}^{M-1} f_m \to \int_0^1 f(x)\,\mathrm{d}x$$

for any continuous function f as $h \to 0$, the continuous and discrete satisfy the same compatibility condition in the limit, but not necessarily for finite values of h.

6.23
Comparing the coefficients of $\mathcal{L}_h$ with those in Definition 6.9, $a_m = c_m = 8 > 0$ and $b_m = 65 \geq a_m + c_m$ and so $\mathcal{L}_h$ is of positive type.

$$\mathcal{L}_h A8^m = A\mathcal{L}_h 8^m = A8^{m-1}(8\times 8^2 - 65\times 8 + 8) = 0$$
$$\mathcal{L}_h B8^{-m} = B\mathcal{L}_h 8^{-m} = B8^{-m+1}(8 - 65\times 8 + 8\times 8^2) = 0$$

so both sequences satisfy $\mathcal{L}_h U = 0$. Since $\mathcal{L}_h$ is a linear operator,

$$\mathcal{L}_h(A8^m + B8^{-m}) = A\mathcal{L}_h 8^m + B\mathcal{L}_h 8^{-m} = 0$$

and $U_m = A8^m + B8^{-m}$ is a solution for any A, B. The BCs lead to the equations

$$A + B = \alpha, \qquad A8^M + B8^{-M} = \beta$$

which are readily solved to give

$$A = 8^{-M}\frac{\beta - \alpha 8^{-M}}{1 - 8^{-2M}}, \qquad B = \frac{\alpha - \beta 8^{-M}}{1 - 8^{-2M}}$$

leading to

$$U_m = \frac{1}{1 - 8^{-2M}}\big(8^{m-M}(\beta - \alpha 8^{-M}) + 8^m(\alpha - \beta 8^{-M})\big).$$

When $M = 10$ we find $8^{-10} \approx 9\times 10^{-10}$ and so, when $\alpha = \pm 1$ and $\beta = \pm 2$ we have $U_m \approx \alpha 8^{-m} + \beta 8^{M-m}$. The solutions with $\alpha = \pm 1$, $\beta = 2$ are shown in Fig. 4. On the left $\alpha = 1, \beta = 2$ so that $\min(0,\alpha,\beta) = 0 \leq U_m \leq \max(0,\alpha,\beta) = 2$. On the right $\alpha = -1, \beta = 2$ so that $\min(0,\alpha,\beta) = -1 \leq U_m \leq \max(0,\alpha,\beta) = 2$.

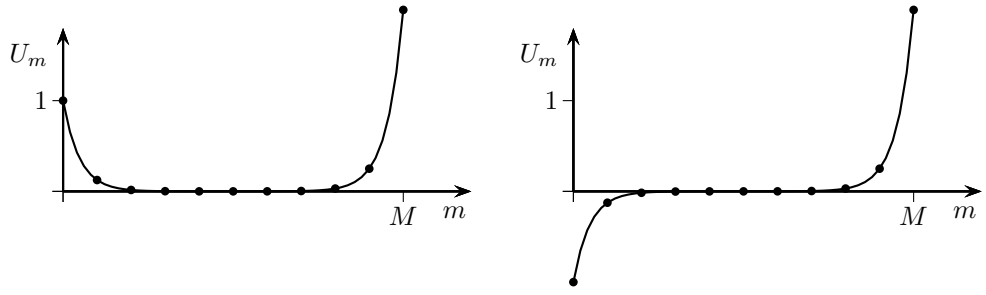


Figure 4: The points $U_m = \alpha 8^{-m} + \beta 8^{m-M}$ for $m = 0, 1, \ldots, M$ and $M = 10$ with $\alpha = 1, \beta = 2$ (left) and $\alpha = -1, \beta = 2$ (right).

6.25
The operator $\mathcal{L}_h$ is of positive type by Example 6.12 (with $r(x) = \sigma^2 > 0$)

$$\mathscr{L}_h U_m = \begin{cases} -h^{-1}\triangle^+ U_0 + \frac{1}{2}h\sigma^2 U_0 \\ \quad \mathcal{L}_h U_m = -h^{-2}\delta^2 U_m + \sigma^2 U_m \\ \quad (\sigma + \frac{1}{2}\sigma^2 h)U_M + h^{-1}\triangle^- U_M \end{cases}$$

$$= \begin{cases} (h^{-1} + \frac{1}{2}h\sigma^2)U_0 - h^{-1}U_1, & m = 0, \\ h^{-2}(-U_{m-1} + (2 + \sigma^2 h^2)U_m - U_{m+1}), & m = 1, 2, \ldots, M-1, \\ (\sigma + \frac{1}{2}\sigma^2 h + h^{-1})U_M - h^{-1}U_{M-1}, & m = M, \end{cases}$$

which satisfy the criteria of Definition 6.17 for $\sigma > 0$.

6.27
The argument used in (6.29) and (6.30) established a second order local truncation error for $0 < m < M$. The local truncation error is identically zero at $m = M$. It remains to examine the situation for $m = 0$. We use the expansion $h^{-1}\triangle^+ u_0 = (u(h) - u(0))/h = u'(0) + \frac{1}{2}hu''(0) + \mathcal{O}(h^2)$ (see Table 6.1)

$$\mathscr{R}_{h,0} = -h^{-1}\triangle^+ u_0 - 1 = (-u'(0) - 1) - \tfrac{1}{2}hu''(0) + \mathcal{O}(h^2) = -\tfrac{1}{2}hu''(0) + \mathcal{O}(h^2)$$

since $-u'(0) = 1$. The order of consistency is therefore first order.
For the second case,

$$\mathscr{R}_{h,0} = -h^{-1}\triangle^+ u_0 - 1 + \tfrac{1}{2}h = (-u'(0) - 1) - \tfrac{1}{2}hu''(0) + \tfrac{1}{2}h + \mathcal{O}(h^2) = \tfrac{1}{2}h(-u''(0) - 1) + \mathcal{O}(h^2).$$

However, the ODE $-u''(x) + xu(x) = 1$ evaluated at $x = 0$ gives $-u''(0) = 1$ so that $\mathscr{R}_{h,0} = \mathcal{O}(h^2)$: the order of consistency is second order.

6.29
Since $h^{-2}\delta^2 u_m = u''_m + \mathcal{O}(h^2)$ and $h^{-1}\triangle u_m = u'_m + \mathcal{O}(h^2)$ (see Table 6.1), the given scheme is consistent of second order.

At $x = 1$, we use a backward difference:

$$h^{-1}\triangle^- u_M = h^{-1}(u(1) - u(1-h)) = u'(1) - \tfrac{1}{2}hu''(1) + \mathcal{O}(h^2)$$

and, since the ODE evaluated at $x = 1$ gives $-u''(1) + 4u'(1) = 0$, the BC $u'(1) = 2$ gives $u''(1) = 8$. Consequently,

$$u'(1) = h^{-1}\triangle^- u_M + 4h + \mathcal{O}(h^2)$$

and the finite difference replacement $h^{-1}\triangle^- U_M = 2 - 4h$ of the BC $u'(1) = 2$ is consistent of second order.

6.31

$$LL^{\mathsf{T}} = h^{-2}\begin{bmatrix} -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & \rho \end{bmatrix}\begin{bmatrix} -1 & & & & \\ 1 & \ddots & & & \\ & \ddots & -1 & & \\ & & 1 & -1 & \\ & & & & \rho \end{bmatrix} = \begin{bmatrix} 2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & \ddots & 2 & -1 & \\ & & -1 & 1+\rho^2 & \end{bmatrix}.$$

When $r(x) \equiv 0$ and $a_{M,M} = 1 + ah/b$, the matrix A in (6.47) becomes

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & \ddots & 2 & -1 & \\ & & -1 & 1+ah/b \end{bmatrix}.$$

Thus $LL^{\mathsf{T}} = A$ by choosing $\rho^2 = ah/b$. This leads to a real value of ρ provided that a and b have the same sign.
It follows that $\boldsymbol{v}^{\mathsf{T}} A \boldsymbol{v} = \boldsymbol{v}^{\mathsf{T}} L L^{\mathsf{T}} \boldsymbol{v} = (L^{\mathsf{T}}\boldsymbol{v})^{\mathsf{T}}(L^{\mathsf{T}}\boldsymbol{v}) = \boldsymbol{w}^{\mathsf{T}}\boldsymbol{w} \geq 0$ when $\boldsymbol{w} = L^{\mathsf{T}}\boldsymbol{v}$ (as in Lemma 6.1). This proves that A is positive semi-definite. We need to check that it is not possible to choose a nonzero vector $\boldsymbol{v}$ such that $\boldsymbol{v}^{\mathsf{T}} A \boldsymbol{v} = 0$. This can only occur if $\boldsymbol{w} = \boldsymbol{0}$ but it is easily verified that the only solution of $L^{\mathsf{T}}\boldsymbol{v} = \boldsymbol{0}$ is $\boldsymbol{v} = \boldsymbol{0}$.

6.33
Using (6.60) and the fact that $\delta^2 f_m = 12\delta^2 x_m = 0$ leads to

$$\begin{aligned} h^{-2}\big(-U_{m-1} + 2U_m - U_{m+1}\big) + \big(U_{m-1} + 10U_m + U_{m+1}\big) &= 12x_m, \\ \text{i.e.,} \quad -(M^2-1)U_{m-1} + (2M^2+10)U_m - (M^2-1)U_{m+1} &= 12x_m \end{aligned}$$

since $h^{-1} = M$. With BCs $u(0) = 3$ and $u(1) = -5$, these can be written when $M = 4$ as

$$\begin{bmatrix} 42 & -15 & 0 \\ -15 & 42 & -15 \\ 0 & -15 & 42 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 45 + 12x_1 \\ 12x_2 \\ -75 + 12x_3 \end{bmatrix},$$

where $x_m = m/4$.

6.35
For indices m for which $\mathscr{F}_{h,m} \neq 0$ we have the standard inequalities:

$$\mathscr{L}_h U_m = \mathscr{F}_{h,m} \leq \|\mathscr{F}_h\|_{h,\infty} \leq \|\mathscr{F}_{h,m}\|_{h,\infty} \mathscr{L}_h \Phi_m.$$

The same end result also holds for indices m such that $\mathscr{F}_{h,m} = 0$ since $\mathscr{L}_h U_m = 0 \leq \|\mathscr{F}_{h,m}\|_{h,\infty} \mathscr{L}_h \Phi_m$ (because the right hand side is automatically non-negative).
Thus, $\mathscr{L}_h U_m \leq \|\mathscr{F}_{h,m}\|_{h,\infty} \mathscr{L}_h \Phi_m$ for all m from which we have

$$\mathscr{L}_h\big(U_m - \|\mathscr{F}_{h,m}\|_{h,\infty} \Phi_m\big) \leq 0$$

and then $U_m - \|\mathscr{F}_{h,m}\|_{h,\infty} \Phi_m \leq 0$ by inverse monotonicity.

This result is particularly useful when dealing with the local truncation error since it is frequently identically zero at some grid points (see the following example).

6.37
Suppose that $\mathcal{L}_h U_m := -a_m U_{m-1} + b_m U_m - c_m U_{m+1}$ and the coefficients have to be chosen so that this is to be consistent with the operator defined by $\mathcal{L}v(x) := -v''(x) + r(x)v(x)$.
We require $\mathcal{L}_h v_m = (\mathcal{L}v(x))\big|_{x=x_m}$ for $v(x) = 1, (x - x_m)$ and $(x - x_m)^2$. Since $\mathcal{M}_h := \mathcal{L}_h - \mathcal{L}$ is a linear operator (i.e., $\mathcal{M}_h(u+v) = \mathcal{M}_h u + \mathcal{M}_h v$ and $\mathcal{M}_h(cu) = c\mathcal{M}_h u$ for any sufficiently smooth functions u and v and any constant c) it follows that

$$\mathcal{M}_h(A + B(x - x_m) + C(x - x_m)^2) = A\big(\mathcal{M}_h 1\big) + B\big(\mathcal{M}_h(x - x_m)\big) + C\big(\mathcal{M}_h(x - x_m)^2\big) = 0$$

for arbitrary constants A, B and C. Thus $\mathcal{M}_h p(x) = 0$ for any quadratic polynomial $p(x)$.
Applying the conditions $\mathcal{L}_h v_m = (\mathcal{L}v(x))\big|_{x=x_m}$ for $v(x) = 1, (x - x_m)$ and $(x - x_m)^2$ we find

$$\begin{array}{rcrcrl} \mathcal{L}_h 1 & = & -a_m + b_m - c_m & = & r_m & = (\mathcal{L}1)\big|_{x=x_m} \\ \mathcal{L}_h(x-x_m) & = & ha_m - hc_m & = & 0 & = (\mathcal{L}(x-x_m))\big|_{x=x_m} \\ \mathcal{L}_h(x-x_m)^2 & = & -h^2 a_m - h^2 c_m & = & -2 & = (\mathcal{L}(x-x_m)^2)\big|_{x=x_m} \end{array}$$

The first and third combine to give $b_m = -2h^{-2} + r_m$ and the second and third give $a_m = c_m = h^{-2}$. Thus $\mathcal{L}_h$ coincides with the approximation used in standard finite difference approximation (6.20).

6.39
Case (a) is entirely standard: $u(x) = 1 - 16(x - \frac{1}{2})^4$.
In case (b), $u''(x) = 0$ for $0 < x < \frac{1}{2}$ and $u(0) = 0$ which gives $u(x) = Ax$ for an arbitrary constant A.
For $\frac{1}{2} < x < 1$, $-u''(x) = 384(x - \frac{1}{2})^2$ which, on integrating twice gives $u(x) = ax + b - 32(x - \frac{1}{2})^4$. The three conditions $u(1) = a + b - 2 = 0$, continuity of u at $x = \frac{1}{2}$: $u(\frac{1}{2}-) = u(\frac{1}{2}+)$, i.e., $\frac{1}{2}A = \frac{1}{2}a + B$ and continuity of u' at $x = \frac{1}{2}$: $u'(\frac{1}{2}-) = u'(\frac{1}{2}+)$, i.e., $A = a$ lead to the solution

$$u(x) = 2x + \begin{cases} 0, & 0 \le x \le 1/2 \\ -32(x - \frac{1}{2})^4, & 1/2 < x \le 1 \end{cases}.$$

A similar procedure in case (c) leads to

$$u(x) = 2x + \begin{cases} 0, & 0 \le x \le 1/2 \\ -16(x - \frac{1}{2})^3, & 1/2 < x \le 1 \end{cases}.$$

These are shown in Fig. 5. These boundary value problems are solved numerically using the

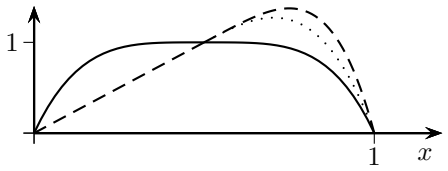


Figure 5: The solutions for Exercise 6.39: case (a) solid line, case (b) dashed line and case (c) dotted line.

standard "second order" scheme

$$-h^{-2}\delta^2 U_m = f_m, \quad m = 1, 2, \ldots, M-1$$

with BCs $U_0 = U_M = 0$. The values of M chosen for the experiments are 8^j, 9^j and 5^j ($j = 1, 2, \ldots, 7$) and the results are summarised in Fig. 6. On the left are shown graphs of $M^2 \times$ global error for $M = 9$ (dots) and $M = 16$ (crosses) in cases (a), (b) and (c). We also include the case labelled "Test" where $f(x) = 8$ and the exact solution $u(x) = 4x(1 - x)$ is a polynomial of degree two, for which the global error should be identically zero. The graphical results show an error of less than 10^{-14} attributable to roundoff error. Including such a test in numerical experiments is recommended to test the integrity of the code. In cases (a) and (b) the graphs with different values of M are indistinguishable, in keeping with the global error $E \propto 1/M^2$. In case (c) the graphs of $M^2 \times E$ appear to be two distinct continuous piecewise linear functions. Note that a grid point lies exactly on the discontinuity in $f'(x)$ at $x = 1/2$ when $M = 16$ but this is not the case when $M = 9$.

On the right we show loglog plots of the global error as a function of h. For cases (a) and (b) the results are almost identical and lie on a line having slope two, again in keeping with $E \propto h^2$ The global error in case (c) is more erratic, especially for the larger values of h. However, the error itself is never larger than in cases (a) and (b) and appears to behave more smoothly as $h \to 0$. The theory developed in this chapter requires that the first four derivatives of the exact solution to be bounded in order that the local truncation error be bounded by a multiple of h^2, and for the global error to converge to zero at a second order rate. The results of these experiments suggest that a second order convergence rate is attainable under less onerous constraints, but that goes beyond the scope of this book.

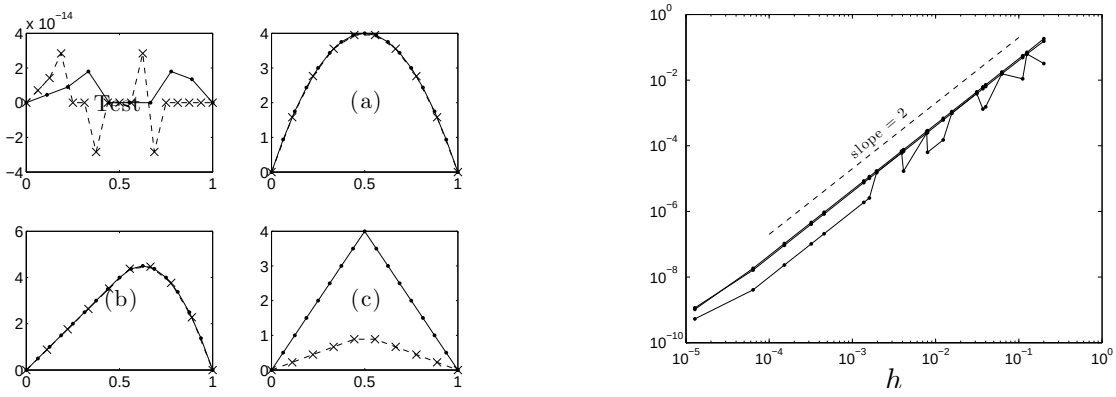


Figure 6: The graphs show $M^2 \times E$ (E is the global error) for $M = 9$ (dots) and $M = 16$ (crosses) in the test case and cases (a), (b) and (c).

Exercises 7 Maximum principles and energy methods

7.1
With $v(x,t) = u(x,t) + \varepsilon(\tau - t)$ in place of (7.2) and it follows that $-\kappa v_{xx} + v_t = -\kappa u_{xx} + u_t - \kappa\varepsilon$ so that $-\kappa v_{xx} + v_t < 0$ for all positive values of ε and the proof proceeds as before (but restricted to $0 \le t \le \tau$).

7.3
If $\mathscr{L}u(x,t) \ge 0$ then $-\kappa u_{xx} + u_t \ge 0$ and therefore, by Theorem 7.1 applied to $-u$ (or part (b) of the previous solution), $u(x,t)$ is either constant or else attains its minimum value on Γ_τ. But $\mathscr{L}u(x,t) \ge 0$ also implies that $u(x,t) \ge 0$ for $(x,t) \in \Gamma_\tau$. Hence $u(x,t) \ge 0$ for $(x,t) \in \Omega_\tau$.

7.5
The PDE $u_t = xu_{xx} + u_x$ may be written as $u_t = (xu_x)_x$ so, multiplying by u and integrating over the interval $(0,1)$ gives

$$\int_0^1 uu_t \,\mathrm{d}x = \int_0^1 u(xu_x)_x \,\mathrm{d}x = \big(uxu_x\big)\big|_{x=0}^1 - \int_0^1 x(u_x)^2 \,\mathrm{d}x$$
$$\tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 u^2 \,\mathrm{d}x = -\int_0^1 x(u_x)^2 \,\mathrm{d}x \le 0.$$

The boundary terms vanish by virtue of the BCs $u(0,t) = u(1,t) = 0$. Thus, the energy $E(t) := \int_0^1 u^2 \,\mathrm{d}x$ is a decreasing function of t so $E(t) \le E(0) = \int_0^1 \sin^2 \pi x \,\mathrm{d}x = 1/2$.

7.7
Differentiating under the integral sign we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 (u_x)^2 \,\mathrm{d}x = \quad 2\int_0^1 u_x u_{xt} \,\mathrm{d}x$$

but $u_{xt} = u_{xxx}$, and so

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 (u_x)^2 \,\mathrm{d}x = \quad 2\int_0^1 u_x u_{xxx} \,\mathrm{d}x = -2\int_1^2 (u_{xx})^2 \,\mathrm{d}r + 2\big(u_x u_{xx}\big)\big|_{x=0}^1.$$

The boundary terms clearly vanish if homogeneous Neumann conditions $u_x = 0$ are applied at $x = 0$ and $x = 1$. When homogeneous Dirichlet conditions $u = 0$ are applied we deduce, from the PDE, that $u_{xx} = u_t$ and boundary terms again vanish because $u_t(0,t) = u_t(1,t) = 0$. Thus $\int_0^1 (u_x)^2 \,\mathrm{d}x$ is a decreasing function of t and

$$\int_0^1 (u_x(x,t))^2 \,\mathrm{d}x \le \int_0^1 (u_x(x,0))^2 \,\mathrm{d}x = \int_0^1 (g'(x))^2 \,\mathrm{d}x.$$

7.9
With homogeneous Neumann BCs: $u_n := \vec{n} \cdot \operatorname{grad} u = 0$ on $\partial\Omega$. Thus, after application of the Divergence Theorem the boundary terms still vanish so v again satisfies

$$\int_\Omega \big((v_x)^2 + (v_y)^2\big) \,\mathrm{d}\Omega = 0$$

from the previous solution. However, the BCs no longer allow us to conclude that $v = 0$ from $v_x = v_y = 0$ in Ω. In fact the solution is only unique up to an arbitrary constant.

7.11
Differentiating $m(t)$ and using the PDE gives

$$\begin{aligned} m'(t) &= \int_{-\infty}^{\infty} u_t \,\mathrm{d}x = \int_{-\infty}^{\infty} \left(u_{xxx} - 6uu_x\right) \mathrm{d}x = \int_0^1 \left(u_{xxx} - 3(u^2)_x\right) \mathrm{d}x \\ &= \left(u_{xx} - 3u^2\right)\Big|_{x=-\infty}^{\infty} = 0 \end{aligned}$$

since u and its derivatives tend to zero at $\pm\infty$. Therefore $m(t)$ is constant in time.
Similarly for $M(t)$:

$$M'(t) = \int_{-\infty}^{\infty} 2uu_t \,\mathrm{d}x = 2\int_{-\infty}^{\infty} \left(uu_{xxx} - 6u^2u_x\right) \mathrm{d}x$$

and, integrating the term involving uu_{xxx} by parts,

$$\begin{aligned} M'(t) &= 2(uu_{xx})\Big|_{-\infty}^{\infty} - 2\int_{-\infty}^{\infty} \left(u_xu_{xx} + 2(u^3)_x\right) \mathrm{d}x = 2(uu_{xx})\Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} \left(\partial_x(u_x)^2 + 4(u^3)_x\right) \mathrm{d}x \\ &= \left(2uu_{xx} - (u_x)^2 - 4u^3\right)\Big|_{-\infty}^{\infty} = 0 \end{aligned}$$

since u and its derivatives tend to zero at $\pm\infty$. Therefore $M(t)$ is constant in time.

Exercises 8 Separation of variables

8.1
As in Example 8.1, $u(x,t) = X(x)T(t)$, where X satisfies $-X'' = \lambda X$ for $0 < x < 1$ but, in this case, the BCs are $X'(0) = X'(1) = 0$.

When $\lambda = -\mu^2 < 0$, the general solution is $X = Ae^{\mu x} + Be^{-\mu x}$ and the BC $X'(0) = 0$ gives $\mu(A - B) = 0$. Thus, since $\mu = 0$ is not possible (since $\lambda < 0$) we must have $A = B$. The second BC then gives $A\mu(e^{\mu} - e^{-\mu}) = 0$. But $e^{\mu} \neq e^{-\mu}$ for $\mu \neq 0$ so we are left with $A = 0$, leading to the trivial solution $X(x) = 0$.

When $\lambda = 0$, the general solution is $X = A + Bx$ and the BCs $X'(0) = X'(1) = 0$ both require $B = 0$ with no restriction on A. Thus $X(x) = A$ is a nontrivial solution corresponding to an eigenvalue $\lambda = 0$. Since $T'(t) = -\lambda T(t)$, we have $T(t) =$ constant and the corresponding solution of the heat equation is $u(x,t) = A$.

When $\lambda = \mu^2 > 0$, the general solution is $X = A\sin\mu x + B\cos\mu x$ and the BC $X'(0) = 0$ gives $\mu A = 0$. Thus, since $\mu = 0$ is not possible (since $\lambda > 0$) we must have $A = 0$. The second BC then gives $B\mu\sin\mu = 0$. To avoid the trivial solution, μ must be chosen so that $\sin\mu = 0$, thus $\mu = n\pi$ for $n = 1, 2, \ldots$. Correspondingly, $T'(t) = -\lambda T(t)$, so $T(t) = Ce^{-(n\pi)^2 t}$, leading to fundamental solutions $e^{-(n\pi)^2 t}\cos n\pi x$.

The general solution is a linear combination of all fundamental solutions and so takes the form

$$u(x,t) = u(x,t) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n e^{-\kappa n^2\pi^2 t}\cos n\pi x.$$

The factor $1/2$ in the leading term allows all the coefficients to be determined by the same formula

$$A_n = 2\int_0^1 g(x)\cos n\pi x\,\mathrm{d}x, \quad n = 0, 1, 2, \ldots.$$

(see Example 8.2). When $g(x) = x$ we find $A_0 = 2\int_0^1 x\,\mathrm{d}x = 1$ and, using integration by parts,

$$\begin{aligned} A_n &= 2\int_0^1 x\cos n\pi x\,\mathrm{d}x = 2x\frac{\sin n\pi x}{n\pi}\Big|_0^1 - \frac{2}{n\pi}\int_0^1 \sin n\pi x\,\mathrm{d}x \\ &= \frac{2}{n^2\pi^2}\cos n\pi x\Big|_0^1 = \frac{2}{n^2\pi^2}((-1)^n - 1). \end{aligned}$$

Thus, $A_n = 0$ when n is even and $A_n = 4/(n\pi)^2$ when n is odd.

8.3
The mean value theorem states that if g is continuous on an interval $[a, b]$ and differentiable on the open interval (a, b) then there is a point $c \in (a, b)$ such that

$$g'(c) = \frac{g(b) - g(a)}{b - a}.$$

If we choose $b = x$, $a = 1 - x$, then, for $x > \frac{1}{2}$,

$$g'(c) = \frac{g(x) - g(1 - x)}{2x - 1} = 0, \qquad 1 - x < c < x.$$

It follows that $g'(\frac{1}{2}) = 0$ by taking the limit $x \to \frac{1}{2}$.
With the change of variable $x = 1 - s$ and $v(s,t) = u(1-s,t)$,

$$v_t = u_t(1-s,t), \quad v_s = -u_x(1-s,t), \quad v_{ss} = u_{xx}(1-s,t)$$

and therefore

$$v_t - \kappa v_{ss} = u_t(1-s,t) - \kappa u_{xx}(1-s,t) = u_t(x,t) - \kappa u_{xx}(x,t) = 0.$$

Also $v(0,t) = u(1,t) = 0$, $v(1,t) = u(0,t) = 0$ and $v(s,0) = u(1-s,0) = g(1-s) = g(s)$ (since g is symmetric about $x = 1/2$).

8.5
From Example 8.1 the general solution of the heat equation with BCs $u = 0$ at both $x = 0$ and $x = 1$ is given by (8.6) and the coefficients by (8.8). When $g(1-x) = g(x)$ we have, making a change of variable $x = 1 - s$,

$$\begin{aligned}\langle g, X_n\rangle &= \int_0^1 g(x)\sin n\pi x \,\mathrm{d}x = \int_0^1 g(1-x)\sin n\pi x \,\mathrm{d}x = \int_0^1 g(s)\sin n\pi(1-s)\,\mathrm{d}s \\ &= -\int_0^1 g(s)\cos n\pi \sin n\pi s \,\mathrm{d}s = (-1)^{n+1}\int_0^1 g(s)\sin n\pi s \,\mathrm{d}s = (-1)^{n+1}\langle g, X_n\rangle.\end{aligned}$$

It follows that $\langle g, X_n\rangle = -\langle g, X_n\rangle$ when n is even. Consequently $\langle g, X_n\rangle = 0$, and so $A_n = 0$ when n is even.

8.7
From Exercise 4.20

$$u_x = \cos\theta\, u_r - \frac{1}{r}\sin\theta\, u_\theta = \cos\theta\, f'(r),$$

since $u = f(r)$. Allowing $x, y \to 0$ would lead to $u_x(x,y)$ being multivalued at the origin (because it would depend on the angle at which the origin was approached) unless $f'(0) = 0$.

8.9
When $g(r) = 1 - (r/b)^2$ for $0 \le r \le b$ and $g(r) = 0$ for $b < r \le a$

$$\int_0^a rg(r)J_0\left(\frac{r\xi_n}{a}\right)\,\mathrm{d}r = \int_0^b r(1-(r/b)^2)J_0\left(\frac{r\xi_n}{a}\right)\,\mathrm{d}r$$

From Exercise D.4 and writing $\omega = b\xi_n/a$,

$$\int_0^b rg(r)J_0\left(\frac{r\xi_n}{a}\right)\,\mathrm{d}r = \left(\frac{a}{\xi_n}\right)^2\int_0^\omega xJ_0(x)\,\mathrm{d}x = \frac{b^2}{\omega}J_1(\omega)$$

and, using $\int x^3J_0(x)\,\mathrm{d}x = 2x^2J_0(x) + x(x^2-4)J_1(x)$, we find

$$\int_0^b r\left(\frac{r}{b}\right)^2 J_0\left(\frac{r\xi_n}{a}\right)\,\mathrm{d}r = \frac{1}{b^2}\left(\frac{a}{\xi_n}\right)^4\int_0^\omega x^3J_0(x)\,\mathrm{d}x = \frac{b^2}{\omega^4}\big(2\omega^2J_0(\omega) + \omega(\omega^2-4)J_1(\omega)\big).$$

These combine to give the required result.

8.11
When $a = 1/2$, $b = 1/4$, and $g(r) = 1$ for $0 < r < b$ and is otherwise zero, we find on integration by parts,

$$A_n = 4\int_0^{\frac{1}{4}} r\sin(2n\pi r)\,\mathrm{d}r = \frac{1}{n^2\pi^2}(\sin\tfrac{1}{2}n\pi - \tfrac{1}{2}n\pi\cos\tfrac{1}{2}n\pi)$$

Hence

$$A_n = \begin{cases} -\dfrac{1}{4m\pi}\cos m\pi = \dfrac{1}{4m\pi}(-1)^{m-1}, & \text{when } n = 2m \\ \dfrac{1}{(2m-1)^2\pi^2}\sin\tfrac{1}{2}(2m-1)\pi = \dfrac{1}{(2m-1)^2\pi^2}(-1)^{m-1}, & \text{when } n = 2m-1. \end{cases}$$

8.13
The solution is given by (8.26) with the coefficients in Exercise 8.11 with $a = 1$ and $g(r) = 1$. Thus,

$$A_n = 2\int_0^1 r\sin(n\pi r)\,\mathrm{d}r = -\frac{2}{n\pi}\cos(n\pi) = \frac{2}{n\pi}(-1)^{n-1}$$

and the solution is

$$u(r,t) = 2\sum_{n=1}^{\infty}(-1)^{n-1}\mathrm{e}^{-n^2\pi^2 t}\frac{\sin n\pi r}{n\pi r}.$$

The result follows by taking the limit $r \to 0$ and using the fact that $\frac{\sin x}{x} \to 1$.
At large times the leading term dominates and $u(0,t) \approx 2\mathrm{e}^{-\pi^2 t}$.

8.15
When g_0 is given by (8.43)

$$g_0(x+ct) + g_0(x-ct) = \sum_{n=1}^{\infty} A_n(\sin n\pi(x+ct) + \sin n\pi(x-ct)) = 2\sum_{n=1}^{\infty} A_n \sin n\pi x\cos n\pi ct.$$

When g_1 is given by (8.44)

$$\begin{aligned}\int_{x-ct}^{x+ct} g_1(z)\,\mathrm{d}z &= \sum_{n=1}^{\infty} B_n cn\pi \int_{x-ct}^{x+ct} \sin n\pi z\,\mathrm{d}z \\ &= -c\sum_{n=1}^{\infty} B_n(\cos n\pi(x+ct) - \cos n\pi(x-ct)) = 2c\sum_{n=1}^{\infty} B_n \sin n\pi x\sin n\pi ct.\end{aligned}$$

Combining these in d'Alembert's solution (4.20) leads to (8.42).

8.17
Substituting $u(x,t) = X(x)T(t)$ into the PDE gives, on dividing by $X(x)T(t)$,

$$\frac{X''(x)}{X(x)} = \frac{T''(t)}{T(t)} = -\lambda$$

leading the the eigenvalue problem $-X''(x) = \lambda X(x)$, $-a < x < a$ with $X(-a) = X(a) = 0$. The standard arguments can be used to show that only trivial solutions are possible when $\lambda \le 0$.

When the origin is not one of the endpoints the calculations are simplified by writing the general solution for $\lambda = \mu^2 > 0$ as

$$X(x) = A \sin \mu(x + \beta)$$

for arbitrary constants A and β. The BC $X(-a) = 0$ immediately sets $\beta = a$ and $X(a) = 0$ leads to $A \sin 2\mu a = 0$. Thus, to avoid trivial solutions we must choose $\mu = \frac{1}{2}n\pi/a$ and the eigenvalues are $\lambda_n = (\frac{1}{2}n\pi)^2$ with corresponding eigenfunctions $X_n(x) = \sin(\frac{1}{2}n\pi(x + a)/a)$. These eigenfunctions are closely related to those in Example 8.1.
The ODE $T''(t) + \mu^2 T(t) = 0$ has the general solution $T(t) = C \sin \mu t + D \cos \mu t$, for arbitrary constants C and D. This leads to the general solution

$$u(x,t) = \sum_{n=1}^{\infty} \big(C_n \sin \frac{n\pi t}{2a} + D_n \cos \frac{n\pi t}{2a}\big) \sin \frac{n\pi(x+a)}{2a}$$

that satisfies the PDE and all BCs.
The initial conditions $u(x,0) = g_0(x)$ and $u_t(x,0) = 0$ give

$$g_0(x) = \sum_{n=1}^{\infty} D_n \ \sin \frac{n\pi(x+a)}{2a}, \quad 0 = \sum_{n=1}^{\infty} C_n \frac{n\pi}{2a} \sin \frac{n\pi(x+a)}{2a}.$$

Since the eigenfunctions are orthgonal on the interval $(-a, a)$,

$$D_n = \frac{\langle g_0, X_n \rangle}{\langle X_n, X_n \rangle}, \qquad C_n = 0.$$

and $\langle X_n, X_n \rangle = \frac{1}{2}a$.

When $g_0(x) = \max(0, 1 - (x/b)^2)$ $(b < a)$ a change of variable $x = bs$ and integrating by parts twice gives

$$D_n = \frac{2}{a} \int_{-b}^{b} (1 - (\frac{x}{b})^2) \sin \tfrac{1}{2}n\pi(x + a)/a \, \mathrm{d}x = \frac{2b}{a} \int_{-1}^{1} (1 - s^2) \sin \omega(s + \alpha) \, \mathrm{d}s, \quad \omega = \frac{n\pi b}{a}, \quad \alpha = \frac{a}{b}$$
$$= 8 \frac{1}{\omega} \frac{\sin \alpha\omega}{\alpha\omega} \big(\frac{\sin \omega}{\omega} - \cos \omega\big).$$

The solution is shown in Fig. 8.7 (dashed line) at time $t = .75$ with $a = 1/2$, $b = 1/4$ (see Example 8.11).

8.19
Suppose $u = \Phi(x, y, a, b)$ represents the solution (8.56) of subproblem P_1 in Example 8.12, so that

$$\Phi(x, 0, a, b) = \sum_{n=1}^{\infty} A_n \sin \Big(\frac{n\pi x}{a}\Big)$$

on E_1 and $\Phi = 0$ on the remaining edges. The PDE is unchanged under the linear change of variable $y \mapsto b - y$ (x unchanged) and maps the rectangle into itself with the edges E_1 being interchanged E_3. This gives the solution

$$u(x,y) = \Phi(x, b - y, a, b) = \sum_{n=1}^{\infty} C_n \frac{\sinh(n\pi y/(ab))}{\sinh(n\pi/a)} \sin \Big(\frac{n\pi x}{a}\Big), \quad u(x,b) = \sum_{n=1}^{\infty} C_n \sin \Big(\frac{n\pi x}{a}\Big)$$

and $u = 0$ on edges E_j, $j = 1, 2, 4$. This is the solution to P_3.

The PDE is unchanged under the interchange $x \leftrightarrow y$ and the domain is also unaltered if we make the interchange $a \leftrightarrow b$. Thus

$$u(x,y) = \Phi(y,x,b,a) = \sum_{n=1}^{\infty} D_n \frac{\sinh\left(n\pi(1-x/a)/b\right)}{\sinh\left(n\pi/b\right)} \sin\Big(\frac{n\pi y}{b}\Big), \quad u(0,y) = \sum_{n=1}^{\infty} D_n \sin\Big(\frac{n\pi y}{b}\Big)$$

is the solution to P_4.
Finally, applying the linear change of variable $x \mapsto a - x$ (y unchanged) gives the solution to P_2:

$$u(x,y) = \Phi(y,a-x,b,a) = \sum_{n=1}^{\infty} B_n \frac{\sinh\left(n\pi x/(ab)\right)}{\sinh\left(n\pi/b\right)} \sin\Big(\frac{n\pi y}{b}\Big), \quad u(a,y) = \sum_{n=1}^{\infty} B_n \sin\Big(\frac{n\pi y}{b}\Big).$$

8.21
By virtue of Exercise 8.19 the solution to problem P_2 is given by

$$u(x,y) = \sum_{n=1}^{\infty} B_n \frac{\sinh(n\pi x)}{\sinh n\pi} \sin(n\pi y), \quad u(1,y) = \sum_{n=1}^{\infty} B_n \sin(n\pi y)$$

and, with $u(1,y) = 1 - y$,

$$B_n = 2\int_0^1 (1-y)\sin(n\pi y)\,\mathrm{d}y = \frac{2}{n\pi} - 2\int_0^1 \frac{1}{n\pi}\cos n\pi y\,\mathrm{d}y = \frac{2}{n\pi} - 2\frac{\sin n\pi}{n^2\pi^2} = \frac{2}{n\pi}.$$

8.23
Substituting $u(x,y) = X(x)Y(y)$ into the PDE and dividing by $X(x)Y(y)$ gives

$$\frac{-X''(x) + 2X'(x)}{X(x)} = \frac{Y''(y)}{Y(y)}$$

and, the left hand side being a function of x only, while the right hand side is a function of y only, we deduce that both must be constant. The homogeneous BCs $u(x,0) = u(x,1) = 0$ imply that $Y(0) = Y(1) = 0$. We therefore look for eigenfunctions in the y-variable and set the separation constant to $-\lambda$ so that $-Y'' = \lambda Y$.

This is the eigenvalue problem solved in Example 8.1 (for $X(x)$) and so the eigenvalues are $\lambda_n = (n\pi)^2$ with corresponding eigenfunctions $Y_n = \sin(n\pi y)$, $n = 1, 2, \ldots$. The corresponding ODE for X is $-X'' + 2X' + \lambda X = 0$. The general solution is a linear combination of $\mathrm{e}^{x+\sigma x}$ and $\mathrm{e}^{x-\sigma x}$, where $\sigma = \sqrt{1+\lambda^2}$. This may be written in several equivalent ways but, in view of the BC $X'(1) = 0$ being applied at $x = 1$, the most convenient form is

$$X(x) = \mathrm{e}^x\big(C\cosh\sigma(1-x) + D\sinh\sigma(1-x)\big)$$

which satisfies $X'(1) = 0$ if $C = \sigma D$. Thus the fundamental solutions are

$$u_n = \mathrm{e}^x\big(\sigma\cosh\sigma(1-x) + \sinh\sigma(1-x)\big)\sin n\pi y, \quad \sigma = \sqrt{1+\lambda_n^2},$$

for $n = 1, 2, \ldots$ and the general solution is $u(x,y) = \sum_{n=1}^{\infty} D_n u_n(x,y)$.

To match the additional BC $u(0,y) = y(1-y)$, we have

$$y(1-y) = \sum_{n=1}^{\infty} A_n \sin n\pi y, \qquad A_n = D_n(\sigma\cosh\sigma + \sinh\sigma),$$

so that (see the discussion from equation (8.7) to (8.8))

$$A_n = 2\int_0^1 y(1-y)\sin n\pi y \,\mathrm{d}y = \frac{4}{(n\pi)^3}(1-(-1)^n).$$

Hence,

$$\begin{aligned} u(x,y) &= \mathrm{e}^x \sum_{n=1}^{\infty} \frac{4}{(n\pi)^3}(1-(-1)^n)\frac{\sigma\cosh\sigma(1-x)+\sinh\sigma(1-x)}{\sigma\cosh\sigma+\sinh\sigma}\sin n\pi y \\ &= \mathrm{e}^x \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{8}{(n\pi)^3}\frac{\sigma\cosh\sigma(1-x)+\sinh\sigma(1-x)}{\sigma\cosh\sigma+\sinh\sigma}\sin n\pi y. \end{aligned}$$

8.25
Using Exercise 4.20 we find that Laplace's equation in polar coordinates becomes

$$u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta} = 0$$

and substituting $u(r,\theta) = R(r)\Theta(\theta)$ into the PDE leads to

$$\frac{r^2R''(r) + rR'(r)}{R(r)} = -\frac{\Theta''(\theta)}{\Theta(\theta)}.$$

The left hand side is a function of r only, while the right hand side is a function of θ only so we deduce that both must be constant. The BCs $u(r,0) = u(r,\pi/4)$ suggest that we look for eigenfunctions in the θ-variable and set the separation constant to λ so that $-\Theta'' = \lambda\Theta$.

(a) $\lambda = -\mu^2 < 0$, then $-\Theta'' = -\mu^2\Theta$ has general solution $\Theta = A\mathrm{e}^{\mu\theta} + B\mathrm{e}^{-\mu\theta}$. This cannot satisfy $\Theta(0) = \Theta(\pi/4) = 0$ for any $\mu \neq 0$ unless $A = B = 0$.

(b) $\lambda = 0$, then $-\Theta'' = 0$ has general solution $\Theta = A + B\theta$ which cannot satisfy $\Theta(0) = \Theta(\pi/4) = 0$ unless $A = B = 0$.

(c) $\lambda = \mu^2 > 0$, then $-\Theta'' = \mu^2 X$ has general solution $\Theta = A\sin\mu\theta + B\cos\mu\theta$. The BC $\Theta(0) = 0$ requires $B = 0$. Then $\Theta(\pi/4) = 0$ requires $A\sin\mu\pi/4 = 0$. This will lead to the trivial solution $A = 0$ unless $\mu = 4n$, $n = 1, 2, \ldots$. The eigenvalues are therefore given by $\lambda_n = 16n^2$ and corresponding eigenfunctions $\Theta_n = \sin(4n\theta)$ Then R satisfies

$$r^2R''(r) + rR'(r) = 16n^2R.$$

Looking for solutions of the form $R = Ar^\alpha$ we find that $r^2R''(r) + rR'(r) = \alpha^2 Ar^\alpha$ and so $\alpha = \pm 4n$. There are two solutions and, by linearity of the ODE, a linear combination is also a solution and therefore

$$R(r) = Dr^{4n} + Er^{-4n}$$

for arbitrary constants D and E. The BC $u(1,\theta)$ implies that $R(1)=0$ so that $E=-D$ and the general solution satisfying the PDE and the homogeneous BCs is

$$u(r,\theta)=\sum_{n=1}^{\infty} D_n\big(r^{4n}-r^{-4n}\big)\sin 4n\theta.$$

The BC $u(2,\theta)=g(\theta)$ then means that the coefficients are given by

$$D_n(2^{4n}-2^{-4n})=\frac{8}{\pi}\int_0^{\pi/4} g(\theta)\sin 4n\theta\,\mathrm{d}\theta.$$

due to the mutual orthogonality of the eigenfunctions over $(0,\pi/4)$ and

$$\int_0^{\pi/4}\sin^2 4n\theta\,\mathrm{d}\theta=\tfrac{1}{8}\pi.$$

8.27

Written in polar coordinates, we require the eigenvalues λ and eigenfunctions u (not identically zero) such that

$$-\left(u_{rr}+\frac{1}{r}u_r+\frac{1}{r^2}u_{\theta\theta}\right)=\lambda u.$$

Substituting $u(r,\theta)=R(r)\Theta(\theta)$ into the PDE leads to

$$-\frac{r^2R''(r)+rR'(r)+\lambda r^2R(r)}{R(r)}=\frac{\Theta''(\theta)}{\Theta(\theta)}.$$

The left hand side is a function of r only, while the right hand side is a function of θ only so we deduce that both must be constant. The periodic BCs $u(r,\theta)=u(r,\theta+2k\pi)$ for any integer k suggests that we look for eigenfunctions in the θ-variable and set the separation constant to α so that $-\Theta''=\alpha\Theta$.

It is readily shown that there are no periodic solutions for $\alpha<0$ so we set $\alpha=\nu^2\geq 0$, so that $\Theta=A\sin\nu\theta+B\cos\nu\theta$. This will be periodic of period 2π if $\nu=n$, $n=0,1,\ldots$ (Note that $n=0$ is permissible). This being the case, then

$$r^2R''+rR'(\lambda r^2-n^2)R=0.$$

The simple change of variable $x=r\sqrt{\lambda}$ converts this into Bessel's equation (D.1) with $\nu=n$. The general solution that remains bounded at the centre of the circle is

$$R_n(r)=CJ_n(r\sqrt{\lambda}),$$

where J_n is the Bessel function of the first kind of order n (see Appendix D). It is necessary to have $R(a)=0$ in order for $u(a,\theta)=0$. Therefore, since $C=0$ leads to the trivial solution, we have to choose λ in such a way that $a\sqrt{\lambda}=\xi_{n,m}$, the mth nonnegative zero of J_n, $m=1,2,\ldots$. Hence the eigenvalues are $\lambda_n=(\xi_{n,m}/a)^2$ and the eigenfunctions are as given in the question.

Exercises 9 The method of characteristics

9.1

(a) $u_t + tu_x = u$:

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{t} = \frac{\mathrm{d}u}{u} \Rightarrow \frac{\mathrm{d}x}{\mathrm{d}t} = t, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = u$$

giving, in terms of (t,k): $x = k + \frac{1}{2}t^2$, $u = A(k)\mathrm{e}^t$.
In terms of (x,t): $u = A(x - \frac{1}{2}t^2)\mathrm{e}^t$.

(b) $tu_t - u_x = 1$:

$$\frac{\mathrm{d}t}{t} = \frac{\mathrm{d}x}{-1} = \frac{\mathrm{d}u}{1} \Rightarrow \frac{\mathrm{d}t}{\mathrm{d}x} = -t, \quad \frac{\mathrm{d}u}{\mathrm{d}x} = -1$$

giving, in terms of (x,k): $t = k\mathrm{e}^{-x}$ and $u = A(k) - x$.
In terms of (x,t): $u = A(t\mathrm{e}^x) - x$.

(c) $u_t + xu_x = -u$:

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{x} = \frac{\mathrm{d}u}{-u} \Rightarrow \frac{\mathrm{d}x}{\mathrm{d}t} = x, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = -u$$

giving, in terms of (t,k): $x = k\mathrm{e}^t$ and $u = A(k)\mathrm{e}^{-t}$.
In terms of (x,t): $u = A(x\mathrm{e}^{-t})\mathrm{e}^{-t}$.

(d) $xu_t - u_x = t$:

$$\frac{\mathrm{d}t}{x} = \frac{\mathrm{d}x}{-1} = \frac{\mathrm{d}u}{t} \Rightarrow \frac{\mathrm{d}t}{\mathrm{d}x} = -x, \quad \frac{\mathrm{d}u}{\mathrm{d}x} = -t$$

giving, $t = k - \frac{1}{2}x^2$ and so $\dfrac{\mathrm{d}u}{\mathrm{d}x} = -k + \dfrac{1}{2}x^2$.
Hence in terms of (x,k): $u = A(k) - kx + \frac{1}{6}x^3$.
In terms of (x,t): $u = A(t + \frac{1}{2}x^2) - (t + \frac{1}{2}x^2)x + \frac{1}{6}x^3$.

(e) $tu_t + xu_x = x$:

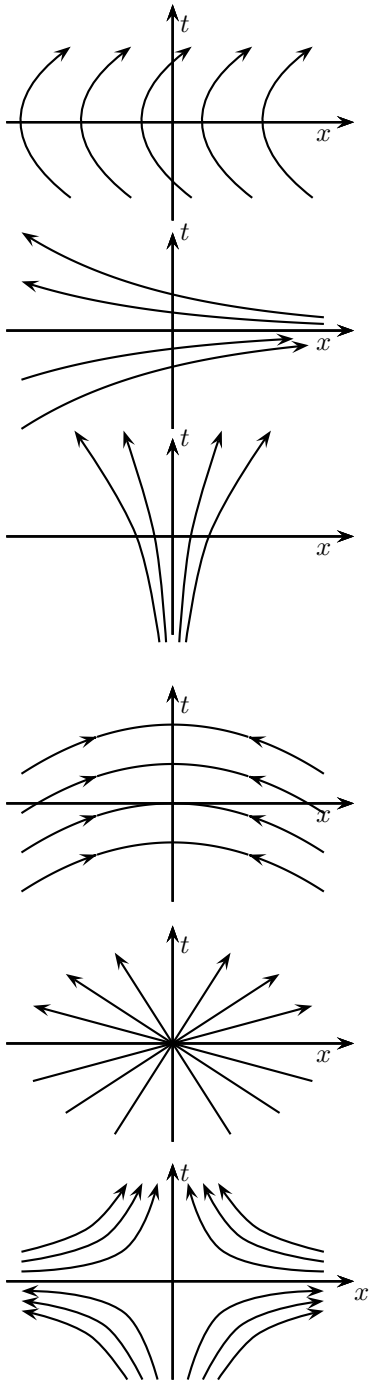
$$\frac{\mathrm{d}t}{t} = \frac{\mathrm{d}x}{x} = \frac{\mathrm{d}u}{x} \Rightarrow \int \frac{\mathrm{d}t}{t} = \int \frac{\mathrm{d}x}{x}, \quad \frac{\mathrm{d}u}{\mathrm{d}x} = 1$$

giving, in terms of (x,k): $t = kx$ and $u = A(k) + x$.
In terms of (x,t): $u = A(t/x) + x$.

(f) $tu_t - xu_x = t$:

$$\frac{\mathrm{d}t}{t} = \frac{\mathrm{d}x}{-x} = \frac{\mathrm{d}u}{t} \Rightarrow \int \frac{\mathrm{d}t}{t} = -\int \frac{\mathrm{d}x}{x}, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = 1$$

giving, in terms of (t,k): $x = k/t$ and $u = A(k) + t$.
In terms of (x,t): $u = A(tx) + t$.



(g) $xu_t - tu_x = xt$:

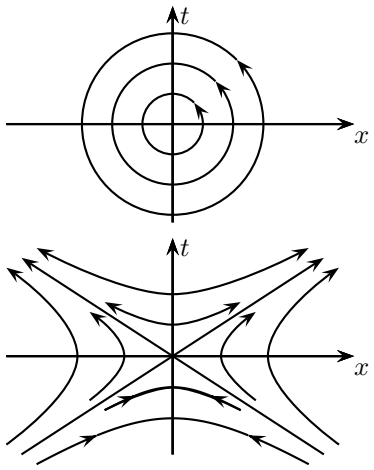
$$\frac{\mathrm{d}t}{x} = \frac{\mathrm{d}x}{-t} = \frac{\mathrm{d}u}{xt} \Rightarrow \int t\,\mathrm{d}t = -\int x\,\mathrm{d}x, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = t$$

giving, in terms of (t,k): $x^2 = k - t^2$ and $u = A(k) + \frac{1}{2}t^2$.
In terms of (x,t): $u = A(x^2+t^2) + \frac{1}{2}t^2$.

(h) $xu_t + tu_x = -xu$:

$$\frac{\mathrm{d}t}{x} = \frac{\mathrm{d}x}{t} = \frac{\mathrm{d}u}{-xu} \Rightarrow \int t\,\mathrm{d}t = \int x\,\mathrm{d}x, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = -u$$

giving, in terms of (t,k): $x^2 = k + t^2$ and $u = A(k)\mathrm{e}^{-t}$.
In terms of (x,t): $u = A(x^2 - t^2)\mathrm{e}^{-t}$.



9.3
The nonzero component of $\boldsymbol{f}_1$ is equal to the first component of $V^{\mathsf{T}}\boldsymbol{g}(X - \lambda_1 T)$. Thus $\boldsymbol{f}_1 = D_1 V^{\mathsf{T}}\boldsymbol{g}(X - \lambda_1 T)$, where D_1 is the $d \times d$ matrix zero matrix except that its $(1,1)$ entry is equal to one. Hence,

$$\boldsymbol{u}_1 = V^{-\mathsf{T}} D_1 V^{\mathsf{T}} \boldsymbol{g}(X - \lambda_1 T) \Rightarrow \|\boldsymbol{u}_1\|_2 \leq \|V^{-\mathsf{T}}\|_2 \|D_1\|_2 \|V^{\mathsf{T}}\|_2 \|\boldsymbol{g}(X - \lambda_1 T)\|_2.$$

However, $\|V^{-\mathsf{T}}\|_2 = \|V^{-1}\|_2$, $\|V^{\mathsf{T}}\|_2 = \|V\|_2$, $\|D_1\|_2 = 1$ and $\|\boldsymbol{g}(X - \lambda_1 T)\|_2 \leq M_2$ so the result follows.
A similar argument holds for $V^{\mathsf{T}}\boldsymbol{u}_j = \boldsymbol{f}_j$, where the only nonzero component of $\boldsymbol{f}_j$ is its jth component, which is $\boldsymbol{v}_j^{\mathsf{T}}\boldsymbol{g}(X - \lambda_j T)$. This time we can write $\boldsymbol{f}_j = D_j V^{\mathsf{T}}\boldsymbol{g}(X - \lambda_j T)$, where D_j is the $d \times d$ matrix zero matrix except that its (j,j) entry is equal to one.
The bound on the solution of (9.6) follows by applying the triangle inequality to $\boldsymbol{u} = \sum_{j=1}^{d} \boldsymbol{u}_j$: $\|\boldsymbol{u}\|_2 \leq \sum_{j=1}^{d} \|\boldsymbol{u}_j\|_2$.
When A is symmetric its eigenvalues are mutually orthogonal and $\kappa_2(V) = 1$.

9.5
At P_1, where $X_1 > 2T_1$, (see Fig. 7, Left)

(a) AP_1 is a Γ_1-characteristic: $x+t = X_1+T_1$ along which $u+w =$ constant. Thus $x_{\mathrm{A}} = X_1+T_1$ and

$$u(\mathrm{P}) + w(\mathrm{P}) = u(\mathrm{A}) + w(\mathrm{A}). \tag{9.5a}$$

(b) BP_1 is a Γ_2-characteristic: $x-t = X_1-T_1$ along which $v+w =$ constant. Thus $x_{\mathrm{B}} = X_1-T_1$ and

$$v(\mathrm{P}) + w(\mathrm{P}) = v(\mathrm{B}) + w(\mathrm{B}). \tag{9.5b}$$

(c) CP_3 is a Γ_3-characteristic: $x - 2t = X_1 - 2T_1$ along which $u + v =$ constant. Thus $x(\mathrm{C}) = X_1 - 2T_1$ and

$$u(\mathrm{P}) + v(\mathrm{P}) = u(\mathrm{C}) + v(\mathrm{A}). \tag{9.5c}$$

Equations (9.5a-c) are clearly of the form (9.13), where the columns of V are the eigenvectors of A^{T} and

$$\boldsymbol{f} = [\boldsymbol{v}_3^{\mathsf{T}}\boldsymbol{g}(\mathrm{A}), \boldsymbol{v}_2^{\mathsf{T}}\boldsymbol{g}(\mathrm{B}), \boldsymbol{v}_1^{\mathsf{T}}\boldsymbol{g}(\mathrm{C})]^{\mathsf{T}}.$$

The initial-conditions provide values for u, v and w at A, B, and C and these equations may be solved to give u, v and w at P_1.

At P_3, where $X_3 < T_3$, (see Fig. 7, Centre)

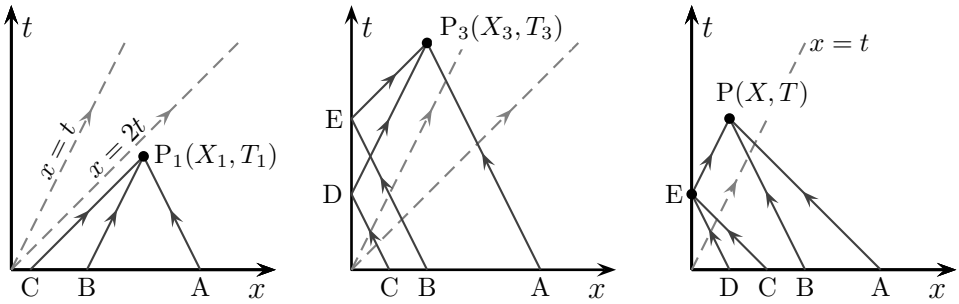


Figure 7: Left and Centre: The characteristics through points P_1, and P_3 for Example 9.2 drawn backwards in time, with reflections when they intersect the t-axis. The dashed lines show the characteristics $x - t = 0$ and $x - 2t = 0$ that pass through the origin. Right: The characteristics for Exercise 9.7.

(a) AP_3 is a Γ_1-characteristic: $x+t = X_3+T_3$ along which $u+w =$ constant. Thus $x_A = X_3+T_3$ and

$$u(P) + w(P) = u(A) + w(A). \tag{9.5A}$$

(b) The Γ_2-characteristic through P_3 intersects the t-axis at D: $x - t = X_3 - T_3$ along which $v + w =$ constant. Thus $t_D = T_3 - X_3$ and $v(P) + w(P) = v(D) + w(D)$. One of the BCs on $x = 0$ is $w(0, t) = 0$ and so $v(P) + w(P) = v(D)$. Now CD is a Γ_1-characteristic so $x_C = t_D$, $u + w$ is constant and, applying the BCs, gives $u(C) + w(C) = v(D)$. Combining these results leads to

$$v(P) + w(P) = u(C) + w(C). \tag{9.5B}$$

(c) The Γ_3-characteristic through P_3 intersects the t-axis at E: $x - 2t = X_3 - 2T_3$ along which $u + v =$ constant. Thus $t_E = T_3 - \frac{1}{2}X_3$ and $u(P) + v(P) = u(E) + v(E)$. On of the BCs on $x = 0$ is $u(0, t) = v(0, t)$ and so $u(P) + v(P) = 2u(E)$. Now BE is a Γ_1-characteristic so $x_B = t_E$, $u + w$ is constant and, applying the BCs, gives $u(B) + w(B) = u(E)$. Combining these results leads to

$$u(P) + v(P) = 2u(B) + 2w(B). \tag{9.5C}$$

Equations (9.5A-C) are clearly of the form (9.13), where the columns of V are the eigenvectors of A^T and the components of $\boldsymbol{f}$ are linear combinations of the values of $\boldsymbol{u}$ evaluated at the points where the characteristics intersect the x-axis.

9.7
The three families of characteristics are:

$$\begin{aligned}
\Gamma_1 : \lambda_1 = -2, &\quad x + 2t = \text{constant}, &\quad \boldsymbol{v}_1^T\boldsymbol{u} = u + w = \text{constant},\\
\Gamma_2 : \lambda_2 = -1, &\quad x + \ \ t = \text{constant}, &\quad \boldsymbol{v}_2^T\boldsymbol{u} = v + w = \text{constant},\\
\Gamma_3 : \lambda_3 = \ \ 1, &\quad x - \ \ t = \text{constant}, &\quad \boldsymbol{v}_3^T\boldsymbol{u} = u + v = \text{constant}.
\end{aligned}$$

In order to find the solution at $P(X, T)$ with$X < T$ (see Fig. 7, Right) we follow the characteristics through P backwards in time until they intersect the initial line $t = 0$ ($x \geq 0$). When a characteristic interests the t-axis (at E, for example), we also have to include the characteristics through E as shown.

AP: This is a Γ_1 characteristic $x + 2t = X + 2T$ along which $u + w$ is constant. Thus,

$$u(A) + w(A) = u(P) + w(P), \qquad x(A) = X + 2T. \tag{9.7a}$$

BP: This is a Γ_2 characteristic $x + t = X + T$ along which $v + w$ is constant. Thus,

$$v(\mathrm{B}) + w(\mathrm{B}) = v(\mathrm{P}) + w(\mathrm{P}), \qquad x(\mathrm{B}) = X + T. \tag{9.7b}$$

EP: This is a Γ_3 characteristic $x - t = X - T$ along which $u + v$ is constant. Thus,

$$u(\mathrm{E}) + v(\mathrm{E}) = u(\mathrm{P}) + v(\mathrm{P}), \qquad t(\mathrm{E}) = T - X. \tag{9.7c}$$

CE: This is a Γ_1 characteristic $x + 2t = 2(T - X)$ along which $u + w$ is constant. The BC specifies $w(0, t) = 0$, and so,

$$u(\mathrm{E}) = u(\mathrm{C}) + w(\mathrm{C}), \qquad x(\mathrm{C}) = 2(T - X). \tag{9.7d}$$

DE: This is a Γ_2 characteristic $x + t = (T - X)$ along which $v + w$ is constant. The BC specifies $w(0, t) = 0$, and so,

$$v(\mathrm{E}) = v(\mathrm{D}) + w(\mathrm{D}), \qquad x(\mathrm{D}) = (T - X). \tag{9.7e}$$

Hence, from (9.7c-e),

$$u(\mathrm{P}) + v(\mathrm{P}) = u(\mathrm{C}) + w(\mathrm{C}) + v(\mathrm{D}) + w(\mathrm{D}).$$

This, together with (9.7a-b) constitute three linearly independent equations to determine $u(\mathrm{P})$, $v(\mathrm{P})$ and $w(\mathrm{P})$. The coefficient matrix of this system is V^{T} whose rows are the eigenvectors of A^{T}.

9.9

With $\boldsymbol{u} = [u, v]^{\mathsf{T}}$ and $\boldsymbol{f} = [f, g]^{\mathsf{T}}$, the equations become

$$\boldsymbol{u}_t + A\boldsymbol{u}_x = \boldsymbol{f}, \quad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The matrix A has eigenvalue $\lambda = 1$ with eigenvector $\boldsymbol{v}_1 = [1, 1]^{\mathsf{T}}$ and eigenvalue $\lambda = -1$ with eigenvector $\boldsymbol{v}_1 = [1, -1]^{\mathsf{T}}$.
Multiplying $\boldsymbol{u}_t + A\boldsymbol{u}_x = \boldsymbol{f}$ by $\boldsymbol{v}_1^{\mathsf{T}} = [1, 1]$ and supposing that $f = 0$, $g = G_t$, we find that $U^+ = \boldsymbol{v}_1^{\mathsf{T}}\boldsymbol{u} = u + v$ satisfies $U_t^+ + U_x^+ = G_t$ for which the characteristic equation is (see (9.2))

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{1} = \frac{\mathrm{d}U^+}{G_t}$$

so that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 1, \frac{\mathrm{d}U^+}{\mathrm{d}t} = G_t \quad \Rightarrow x = t + k, \quad U^+ = G(x, t) + A(k), \Rightarrow u + v = G(x, t) + A(x - t),$$

where k is an arbitrary constant and $A(k)$ an arbitrary function. Similarly, multiplying $\boldsymbol{u}_t + A\boldsymbol{u}_x = \boldsymbol{f}$ by $\boldsymbol{v}_2^{\mathsf{T}} = [1, -1]$, we find that $U^- = \boldsymbol{v}_2^{\mathsf{T}}\boldsymbol{u} = u - v$ satisfies $U_t^- + U_x^- = -G_t$ for which the characteristic equation is

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{-1} = \frac{\mathrm{d}U^-}{-G_t}$$

so that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -1, \frac{\mathrm{d}U^-}{\mathrm{d}t} = G_t \quad \Rightarrow x = -t + k, \quad U^- = G(x, t) + B(k), \Rightarrow u - v = -G(x, t) + B(x + t)$$

where B is an arbitrary function.
Combining these we have

$$u = \tfrac{1}{2}(A(x-t) + B(x+t)), \quad v = G(x,t) + \tfrac{1}{2}(A(x-t) - B(x+t)).$$

The initial condition $u(x,0) = 0$ gives $B(x) = -A(x)$ and then $v(x,0) = 0$ gives $A(x) = -G(x,0)$.
Hence the solution is

$$u(x,t) = \tfrac{1}{2}(G(x+t,0) - G(x-t,0)), \quad v(x,t) = G(x,t) - \tfrac{1}{2}(G(x-t,0) + G(x+t,0)).$$

9.11
Suppose that P has coordinates (X, Y).

(a) For $X > 2Y > 0$ (see Fig. 8) both characteristics through P_1 intersect the boundary on the x-axis. Hence $u(X,Y)$ is determined by the initial conditions—the solution to Exercise 9.10 fulfils these conditions.

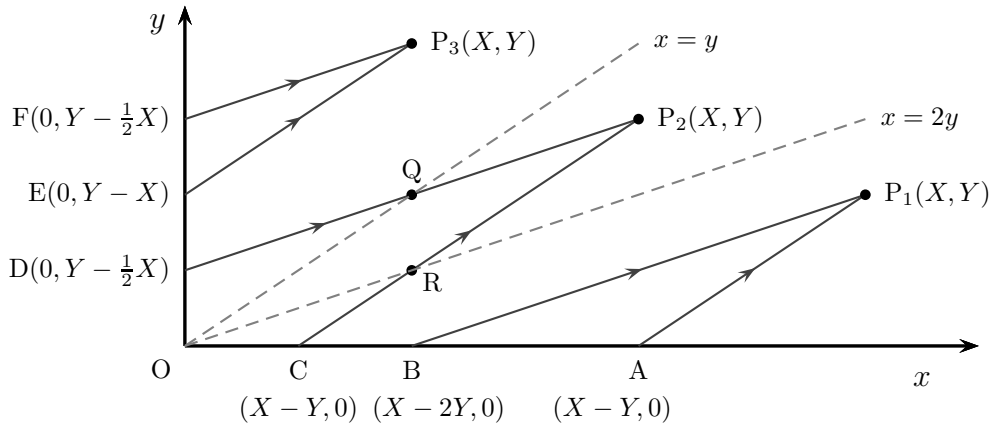


Figure 8: The characteristics through three general points P_j $(j = 1, 2, 3)$ for Exercise 9.11.

(b) For $0 < X < Y$ both characteristics through P_3 intersect the boundary on the y-axis. Hence $u(X,Y)$ is determined by the boundary conditions. We begin with the general solution $u(x,y) = F(x-y) + G(x-2y)$ of the PDE derived in the solution to Exercise 9.10. Applying the BCs leads to

$$F(-y)+G(-2y) = f_0(y), \quad F'(-y)+G'(-2y) = f_1(y) \quad \Rightarrow -F(-y)-\tfrac{1}{2}G(-2y) = \int^{y} f_1(s)\,\mathrm{d}s$$

from which we obtain

$$G(-2y) = 2f_0(y) + 2\int^{y} f_1(s)\,\mathrm{d}s \Rightarrow \qquad G(t) = \; 2f_0(-\tfrac{1}{2}t) + 2\int^{-t/2} f_1(s)\,\mathrm{d}s,$$

$$F(-y) = \; -f_0(y) - 2\int^{y} f_1(s)\,\mathrm{d}s \Rightarrow \qquad F(t) = -2f_0(-t) - 2\int^{-t} f_1(s)\,\mathrm{d}s.$$

The solution satisfying the BCs is, therefore,

$$u(x,y) = 2f_0(y - \tfrac{1}{2}x) - f_0(y-x) + 2\int_{y-x}^{y-\frac{1}{2}x} f_1(s)\,\mathrm{d}s.$$

(c) For $2Y > X > Y > 0$ (see Fig. 8) one characteristic through P_2 intersects the boundary on each axis. Hence $u(X,Y)$ is determined by both the initial and boundary conditions.

Using $u(x,y) = F(x-y) + G(x-2y)$ we find that, along $x = y$, $u(y,y) = F(0) + G(-y)$ while, along $x = 2y$, $u(2y,y) = F(y) + G(0)$. Hence

$$F(y) = u(2y,y) - G(0), \qquad G(y) = u(-y,-y) - F(0).$$

and

$$u(X,Y) = u(2(X-Y), X-Y) + u(2Y-X, 2Y-X) - F(0) - G(0).$$

Setting $X = Y = 0$ shows that $F(0) + G(0) = u(0,0)$. The characteristics $y = x + Y - X$ and $2y = x + 2Y - X$ through P_3 intersect the characteristics $y = x$ and $2y = x$ through the origin at Q and R whose coordinates are $\big(2Y-X, 2Y-X\big)$ and $\big(2(X-Y), X-Y\big)$, respectively. Hence, $u(\mathrm{P}) = u(\mathrm{Q}) + u(\mathrm{R}) - u(0,0)$.

9.13

The operator associated with the PDE $u_{tt} + u_{tx} - 2u_{xx} = t$ factorizes:

$$u_{tt} + u_{tx} - 2u_{xx} = (\partial_t - \partial_x)(\partial_t + 2\partial_x)u$$

Hence, with $v = (\partial_t + 2\partial_x)u$ we have $(\partial_t - \partial_x)v = t$ whose characteristic equations are

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{-1} = \frac{\mathrm{d}v}{t} \quad \Rightarrow \quad \frac{\mathrm{d}x}{\mathrm{d}t} = -1, \quad \frac{\mathrm{d}v}{\mathrm{d}t} = t$$

so the characteristics are $x + t = k_1$ along which $v = \frac{1}{2}t^2 + a(k_1)$, where $a(\cdot)$ is an arbitrary function. Then u is the solution of $(\partial_t + 2\partial_x)u = v$ whose characteristic equations are

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{2} = \frac{\mathrm{d}u}{v} \quad \Rightarrow \quad \frac{\mathrm{d}x}{\mathrm{d}t} = 2, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = v$$

so the characteristics are $x - 2t = k_2$ along which $u = \frac{1}{6}t^3 + \int^t a(x+s)\,\mathrm{d}s + B(k_2)$. This gives the general solution

$$u(x,t) = \tfrac{1}{6}t^3 + A(x+t) + B(x-2t),$$

where $A(x+t) = \int^t a(x+s)\,\mathrm{d}s$.
The initial conditions $u(x,0) = u_t(x,0) = 0$ give

$$A(x) + B(x) = 0, \quad A'(x) - 2B'(x) = 0 \quad \Rightarrow A(x) - 2B(x) = \text{constant} = C$$

and therefore $A(x) = \frac{1}{3}C = -B$. Hence $u(x,t) = \frac{1}{6}t^3$.

9.15

$$\begin{aligned}
\mathcal{L}_1\mathcal{L}_2 &= (a\partial_t + b\partial_x)(c\partial_t + d\partial_x) = a\partial_t(c\partial_t + d\partial_x) + b\partial_x(c\partial_t + d\partial_x) \\
&= a\big(c_t\partial_t + c\partial_t^2 + d_t\partial_x + d\partial_x\partial_t\big) + b\big(c_x\partial_t + c\partial_t\partial_x + d_x\partial_x + d\partial_x^2\partial_t\big) \\
&= \mathcal{L} + (ac_t + bc_x)\partial_t + (ad_t + bd_x)\partial_x
\end{aligned}$$

and so $\mathcal{L}_1\mathcal{L}_2 = \mathcal{L}$ if $ac_t + bc_x = 0$ and $ad_t + bd_x = 0$.
Thus, $\mathcal{L}u = 0$ is equivalent to $\mathcal{L}_1 v = 0$ where $\mathcal{L}_2 u = v$.

For the given PDE, comparing coefficients we find $ac = 1$, $ad+bc = t-1$ and $bd = -t$. A solution of these is $a = c = 1$, $b = t$, $d = -1$ so that $\mathcal{L}_1 = \partial_t + t\partial_x$, $\mathcal{L}_2 = \partial_t - \partial_x$. The characteristic equations of $\mathcal{L}_1 v = 0$ are

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{t} = \frac{\mathrm{d}v}{0} \;\Rightarrow\; \frac{\mathrm{d}x}{\mathrm{d}t} = \quad t,\; \frac{\mathrm{d}v}{\mathrm{d}t} = 0,\; \Rightarrow x = k_1 + \tfrac{1}{2}t^2,\; v = A(k_1)$$

and so $v = u_t - u_x = A(k_1)$ is constant along $x - \frac{1}{2}t^2 = k_1$. The characteristic through the point (x,t) cuts the x-axis at $(x - \frac{1}{2}t^2, 0)$, at which point $u_t = g_1(x - \frac{1}{2}t^2)$ and $u_x = g_0'(x - \frac{1}{2}t^2)$. Hence,

$$u_t - u_x = A(x - \tfrac{1}{2}t^2), \text{ where } A(z) = g_1(z) - g_0'(z).$$

This PDE $\mathcal{L}_2 u = v$ has characteristic equations

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{-1} = \frac{\mathrm{d}u}{v} \;\Rightarrow\; \frac{\mathrm{d}x}{\mathrm{d}t} = -1, \;\; \Rightarrow x = k_2 - t,$$
$$\frac{\mathrm{d}u}{\mathrm{d}t} = A(x - \tfrac{1}{2}t^2) = A(k_2 - t - \tfrac{1}{2}t^2) \;\Rightarrow u = \int_0^t A(x + t - s - \tfrac{1}{2}s^2)\,\mathrm{d}s + g_0(x+t)$$

using the initial condition $u(x,0) = g_0(x)$.

9.17
The PDE $u_t + uu_x = -2u$ has the characteristic equations

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{u} = \frac{\mathrm{d}u}{-2u}$$

leading to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = u, \quad \frac{\mathrm{d}u}{\mathrm{d}t} = -2u.$$

The second of these has general solution $u = C\mathrm{e}^{-2t}$ from which the first ODE gives $x = k - \frac{1}{2}C\mathrm{e}^{-2t}$, where k and C are related arbitrary constants—which we express by writing $C = A(k)$. Consider a characteristic emanating from a point on the boundary given by $x = 0$ and $t = t^*$, say. We then have $k = \frac{1}{2}A(k)\mathrm{e}^{-2t^*}$ and the BC $u(0,t^*) = \mathrm{e}^{-t} = A(k)\mathrm{e}^{-2t^*}$ so that $A(k) = \mathrm{e}^{t^*}$, $k = \frac{1}{2}\mathrm{e}^{-t^*}$ and therefore $kA(k) = \frac{1}{2}$. These results tie the arbitrary constants for a characteristic to its point of intersection on the t-axis. Furthermore, $u = \mathrm{e}^{t^*-2t}$ and $x = \frac{1}{2}(\mathrm{e}^{-t^*} - \mathrm{e}^{t^*-2t})$. Eliminating t^* shows that u is related to x and t via the quadratic equation

$$u^2 + 2xu - \mathrm{e}^{-2t} = 0$$

whose roots are

$$u = -x \pm \sqrt{x^2 + \mathrm{e}^{-2t}}.$$

Since $u > 0$ for $x = 0$ the positive root is appropriate and, by rationalising the result, we obtain

$$u = \frac{\mathrm{e}^{-2t}}{x + \sqrt{x^2 + \mathrm{e}^{-2t}}}.$$

9.19
Under the change of variables $s = -x$, $v(s,t) = -u(-s,t)$ the PDE becomes

$$-v_t + (-v)(-v)_{-s} = v \quad \Rightarrow v_t + vv_s = -u$$

so remains invariant. The initial condition becomes $v(s,0) = -u(-s,0) = -g(-s)$ for $s \in \mathbb{R}$. Thus $v(s,0) = g(s)$ when g is an odd function in which case v satisfies the same PDE and initial condition as u so $v(s,t) = u(s,t) = -u(-s,t)$ so that u must be an odd function of x.
With $g(x) = x/(1+|x|)$, which is an odd function, in (9.27) we have, for $k \geq 0$,

$$u = \frac{k}{1+k}\mathrm{e}^{-t}, \quad x = k + \frac{k}{1+k}(\mathrm{e}^{-t} - 1).$$

Defining $U = u\mathrm{e}^{t}$ and $a(t) = \mathrm{e}^{-t} - 1$ to simplify the notation, we find

$$U = \frac{k}{1+k}, \quad x = k + Ua(t) \quad \Rightarrow k = \frac{U}{1-U}, \quad x = \frac{U}{1-U} + Ua(t)$$

which gives the relationship between U, x and t. On rearranging we obtain the quadratic equation

$$a(t)U^2 - (1 + x + a(t))U + x = 0$$

whose roots are

$$U = \frac{1}{2a(t)}\big((1 + x + a(t)) \pm \sqrt{(1 + x + a(t))^2 - 4a(t)x}.$$

Since $a(t) \to 0$ as $t \to 0$, we have to choose the root for which the numerator also vanishes at $t = 0$—this is the negative root. Thus, since $u = U\mathrm{e}^{-t}$,

$$\begin{aligned} u(x,t) &= \frac{\mathrm{e}^{-t}}{2a(t)}\big((1 + x + a(t)) - \sqrt{(1 + x + a(t))^2 - 4a(t)x}\big) \\ &= \frac{2x\mathrm{e}^{-t}}{(1 + x + a(t)) + \sqrt{(1 + x + a(t))^2 - 4a(t)x}} \end{aligned}$$

which is valid for $x \geq 0$. The solution for $x < 0$ follows since u must be an odd function of x.

9.21
The general solution of Burgers' equation takes the form (9.30):

$$u = g(k), \quad x = g(k)t + k.$$

The characteristics pass through the point $x = k$ at $t = 0$ and $u(x,0) = g(x)$ and $u =$ constant along characteristics. When $g(x) = \max\{0, 1 - |x|\}$ the initial condition is $u = 0$ for $|x| \geq 1$ and the characteristics are therefore $x =$ constant if there are no shocks present.

(a) Consider characteristics emanating from the points $(x,t) = (k,0)$ for $0 \leq k \leq 1$. Here $g(x) = 1 - x$ and so

$$u = 1 - k, \quad x = (1-k)t + k \quad \Rightarrow u(x,t) = \frac{1-x}{1-t}$$

so this solution is valid only until $t = 1$ when the entire family of these characteristics intersect at $(1,1)$ and a shock forms. Note that $u(1,t) = 0$ for $0 \leq t < 1$.

(b) Consider characteristics emanating from the points $(x,t) = (k,0)$ for $-1 \leq k \leq 0$. Here $g(x) = 1 + x$ and so

$$u = 1 + k, \quad x = (1+k)t + k \quad \Rightarrow u(x,t) = \frac{1+x}{1+t}$$

a solution that is valid for $t > 0$. Note that $u(-1,t) = 0$ for $t \geq 0$.

(c) We have seen that the characteristics in case (a) intersect to form a shock at $x = t = 1$. However, the rightmost characteristic from case (b) (i.e., $k = 0^-$) also intersects the characteristic $x = 1$ and the speed and location of the shock is determined by the interaction of characteristics of case (b) with those for $x > 1$, which are vertical lines since $u = 0$ on each (see Fig. 9). The appropriate Rankine-Hugoniot condition (see Example 9.11) is

$$s'(t) = \tfrac{1}{2}(u^+ + u^-)$$

with $u^+ = 0$ and $u^- = 1 + k$. Along the (b)-characteristics we have $k = (x - t)/(1 + t)$ and therefore, when $x = s(t)$,

$$s'(t) = \tfrac{1}{2}\frac{1+s}{1+t}, \quad t > 1, \quad s(1) = 1.$$

This has solution $s(t) = \sqrt{2+2t} - 1$. The solution for $t > 1$ is therefore given by

$$u(x,t) = \begin{cases} 0, & x \le -1 \\ \frac{1+x}{1+t}, & -1 < x < s(t) \\ 0, & x > s(t). \end{cases}$$

The solution at selected times is shown in Fig. 9.

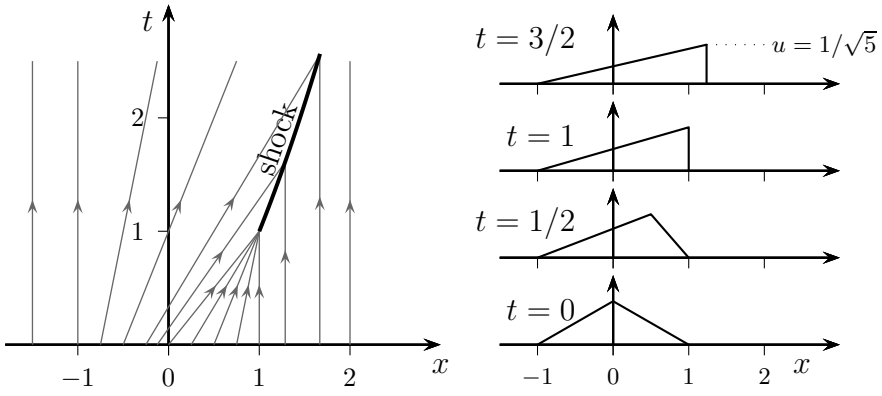


Figure 9: The characteristics (left) and solutions at selected times (right) for Exercise 9.21.

The triangular profile persists for all time. Prior to shock formation it has a fixed base of length 2 and a constant height $u = 1$ so the area of the triangle is 1 for $0 \le t \le 1$.
After shock formation the base of the triangle has length $(1 + s(t))$ and height $u(s(t), t) = (1 + s(t))/(1 + t)$ so the area is

$$\tfrac{1}{2}(1 + s(t))\frac{1 + s(t)}{1 + t} = 1$$

since $1 + s(t) = \sqrt{2 + 2t}$.

9.23

$$g(x) = 1 - \frac{1}{1 + |x|} = \begin{cases} \frac{1}{1+x}, & x \ge 0 \\ \frac{2x-1}{x-1} & x < 0. \end{cases}$$

Hence, for $k = k^+ > 0$, (9.41) leads to

$$s = k + \frac{t}{1+k} \quad \Rightarrow \quad k^2 - (s-1)k + t - s = 0$$

and, therefore,

$$k = \tfrac{1}{2}\big(s - 1 \pm \sqrt{(s-1)^2 - 4(t-s)}\big).$$

Since $k = k^+ > 0$ we have to choose the positive square root (this is readily checked at $t = 0$).
Similarly, for $k = k^- < 0$, (9.41) leads to

$$s = k + t\frac{2k-1}{k-1} \quad \Rightarrow \quad k^2 - (1+s-2t)k + s - t = 0$$

and, therefore,

$$k = \tfrac{1}{2}\big(1 + s - 2t \pm \sqrt{(1+s-2t)^2 - 4(s-t)}\big).$$

This time we have to choose the negative square root since $k = k^- < 0$.
Clearly, when $s(t) = t$ these give $k^+ = k^- = 0 = k^*$ for $t \geq 1$.

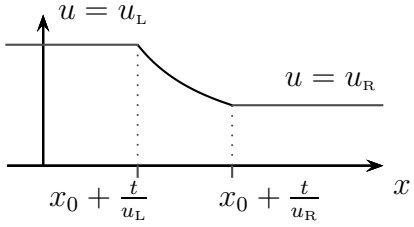
9.25
Differentiating the expression $q'(u) = \frac{x-x_0}{t}$ with respect to x and t gives

$$\partial_x q'(u) = q''(u)u_x = \frac{1}{t}, \qquad \partial_t q'(u) = q''(u)u_t = -\frac{x-x_0}{t^2}$$

so that $q''(u)u_t = -q'(u)q''(u)u_x$, that is, $u_t + q'(u)u_x = 0$ provided that $q''(u) \neq 0$.

9.27
$q(u) = \log u$ so $q'(u) = 1/u = (x - x_0)/t$ gives the expansion fan

$$u = \begin{cases} u_{\mathrm{L}}, & x - x_0 < t/u_{\mathrm{L}}, \\ t/(x-x_0), & t/u_{\mathrm{L}} \leq x - x_0 \leq t/u_{\mathrm{R}}, \\ u_{\mathrm{R}}, & t/u_{\mathrm{R}} < x - x_0. \end{cases}$$



9.29
The behaviour of the solution is broadly similar to that in the example—an expansion fan forms at $x = 0$ and a shock at $x = 1$ and they will collide at some point in time.

The expansion fan with $u_{\mathrm{L}} = 1$ and $u_{\mathrm{R}} = 2$ is given by

$$u = \begin{cases} 1, & x < t, \\ x/t, & t \leq x \leq 2t, \\ 2, & 2t < x. \end{cases}$$

A shock is already present at $x = 1$, $t = 0$ with $u_{\mathrm{L}} = 2$, $u_{\mathrm{R}} = 1$. The shock speed is $s'(t) = \frac{1}{2}(u_{\mathrm{L}} + u_{\mathrm{R}}) = \frac{3}{2}$ and its location is therefore $x = s(t) = 1 + \frac{3}{2}t$ (since $s(0) = 1$).
The rightmost characteristic from the expansion fan meets the shock wave when $x = 2t = 1 + \frac{3}{2}t$, i.e., when $t = 2$ and $x = 4$.
The solution is illustrated in Fig. 10 for $t = \frac{1}{3}$ (left) and $t = 2$ (right).

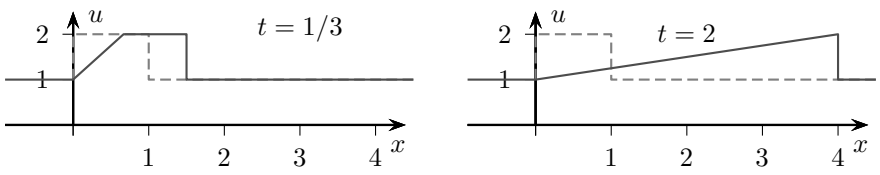


Figure 10: The solution of Exercise 9.29 when $t = \frac{1}{3}$ (left) and $t = 2$ (right). The dashed lines show the initial condition.

Exercises 10 Finite difference methods for elliptic PDEs

10.1
We use the replacements $u_{xx} = h_x^{-2}\delta_x^2 u - \frac{1}{12}h_x^2\partial_x^4 u + \mathcal{O}(h_x^4)$ and $u_{yy} = h_y^{-2}\delta_y^2 u - \frac{1}{12}h_y^2\partial_y^4 u + \mathcal{O}(h_y^4)$ and so

$$\nabla^2 u = u_{xx} + u_{yy} = h_x^{-2}\delta_x^2 u + h_y^{-2}\delta_y^2 u - \tfrac{1}{12}\left(h_x^2\partial_x^4 u + h_y^2\partial_y^4 u\right) + \mathcal{O}(h_x^4) + \mathcal{O}(h_y^4).$$

Hence the local truncation error for Poisson's equation: $-\nabla^2 u = f$ is

$$\begin{aligned}\mathcal{R}_{\ell,m} &= -\left(h_x^{-2}\delta_x^2 u + h_y^{-2}\delta_y^2 u\right) - f \\ &= -\nabla^2 u - \tfrac{1}{12}\left(h_x^2\partial_x^4 u + h_y^2\partial_y^4 u\right) + \mathcal{O}(h_x^4) + \mathcal{O}(h_y^4) - f \\ &= -\tfrac{1}{12}\left(h_x^2\partial_x^4 u + h_y^2\partial_y^4 u\right) + \mathcal{O}(h_x^4) + \mathcal{O}(h_y^4).\end{aligned}$$

Our finite difference replacement is

$$-\left(h_x^{-2}\delta_x^2 + h_y^{-2}\delta_y^2\right) U_{\ell,m} = f_{\ell,m}$$

and, because $\delta_x^2 U_{\ell,m} = U_{\ell-1,m} - 2U_{\ell,m} + U_{\ell+1,m}$, $\delta_y^2 U_{\ell,m} = U_{\ell,m-1} - 2U_{\ell,m} + U_{\ell,m+1}$ this becomes, on multiplying by $h_x h_y$

$$\begin{aligned}&\frac{h_y}{h_x}\left(U_{\ell-1,m} - 2U_{\ell,m} + U_{\ell+1,m}\right) + \frac{h_x}{h_y}\left(U_{\ell,m-1} - 2U_{\ell,m} + U_{\ell,m+1}\right) = h_x h_y f_{\ell,m} \\ &2(\theta + \theta^{-1})U_{\ell,m} - \theta\left(U_{\ell-1,m} + U_{\ell+1,m}\right) - \theta^{-1}\left(U_{\ell,m-1} + U_{\ell,m+1}\right) = h_x h_y f_{\ell,m}.\end{aligned}$$

If we use this scheme to solve Poisson's equation on the rectangle $\{0 < x < L_x, 0 < y < L_y\}$ with $h_x = L_x/M_x$ and $h_y = L_y/M_y$ (where M_x and M_y are specified integers indicating the number of grid points in each direction) the totality of equations can me written in matrix vector form by defining the $M_y \times M_y$ tridiagonal matrix

$$D = \begin{bmatrix} 2(\theta+\theta^{-1}) & -\theta^{-1} & & & \\ -\theta^{-1} & 2(\theta+\theta^{-1}) & -\theta^{-1} & & \\ & \ddots & \ddots & \ddots & \\ & -\theta^{-1} & 2(\theta+\theta^{-1}) & -\theta^{-1} & \\ & & -\theta^{-1} & 2(\theta+\theta^{-1}) & \end{bmatrix}$$

and the $M_x \times M_x$ block tridiagonal matrix

$$A = \begin{bmatrix} D & -\theta I & & & \\ -\theta I & D & -\theta I & & \\ & \ddots & \ddots & \ddots & \\ & & -\theta I & D & -\theta I \\ & & & -\theta I & D \end{bmatrix}$$

then, in the natural column ordering of points, $A\boldsymbol{u} = \boldsymbol{f}$. This reduces to the standard 5–point formula (10.6) when $\theta = 1$.

10.3

We use the standard finite difference operators

$$\delta_x^2 U_{\ell,m} = U_{\ell+1,m} - 2U_{\ell,m} + U_{\ell-1,m}, \quad \delta_y^2 U_{\ell,m} = U_{\ell,m+1} - 2U_{\ell,m} + U_{\ell,m-1}$$

so

$$\begin{aligned}\delta_x^2\delta_y^2 U_{\ell,m} = \delta_x^2\left(\delta_y^2 U\right)_{\ell,m} &= \left(\delta_y^2 U\right)_{\ell+1,m} - 2\left(\delta_y^2 U\right)_{\ell,m} + \left(\delta_y^2 U\right)_{\ell-1,m} \\ &= \left(U_{\ell+1,m+1} - 2U_{\ell+1,m} + U_{\ell+1,m-1}\right) - 2\left(U_{\ell,m+1} - 2U_{\ell,m} + U_{\ell,m-1}\right) \\ &\qquad + \left(U_{\ell-1,m+1} - 2U_{\ell-1,m} + U_{\ell-1,m-1}\right).\end{aligned}$$

The stencil of this result may be visualized as the "outer" product of a column and a row vector:

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} [1, -2, 1] = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}.$$

Then

$$2h^2\mathcal{L}_h^+ - \delta_x^2\delta_y^2 = \begin{bmatrix} & -2 & \\ -2 & 8 & -2 \\ & -2 & \end{bmatrix} - \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & & -1 \\ & 4 & \\ -1 & & -1 \end{bmatrix} = 2h^2\mathcal{L}_h^\times$$

We know from equation (10.13) that

$$\mathcal{L}_h^+ u_{\ell,m} = -\nabla^2 u_{\ell,m} + \mathcal{R}_{\ell,m}^+,$$

where

$$\mathcal{R}_{\ell,m}^+ = -\tfrac{1}{12}h^2\left(\partial_x^4 u + \partial_y^4 u\right) + \mathcal{O}(h^4)$$

so it then follows from $\mathcal{L}_h^\times U_{\ell,m} = \mathcal{L}_h^+ U_{\ell,m} - \frac{1}{2}h^{-2}\delta_x^2\delta_y^2 U_{\ell,m}$ that

$$\begin{aligned}\mathcal{L}_h^\times u_{\ell,m} = \mathcal{L}_h^+ u_{\ell,m} - \tfrac{1}{2}h^{-2}\delta_x^2\delta_y^2 u_{\ell,m} &= \mathcal{L}_h^+ u_{\ell,m} - \tfrac{1}{2}h^2\left(h^{-4}\delta_x^2\delta_y^2 u_{\ell,m}\right) \\ &= -\nabla^2 u - \tfrac{1}{12}h^2\left(\partial_x^4 u + \partial_y^4 u\right) + \mathcal{O}(h^4) - \tfrac{1}{2}h^2\left(\partial_x^2\partial_y^2 u_{\ell,m} + \mathcal{O}(h^2)\right) \\ &= -\nabla^2 u - \tfrac{1}{12}h^2\left(\partial_x^4 u + 6\partial_x^2\partial_y^2 u + \partial_y^4 u\right) + \mathcal{O}(h^4).\end{aligned}$$

10.5
In standard column-wise order, the equations $\mathcal{L}_h^\times U = F$ are

$$\left[\begin{array}{ccc|ccc|ccc} 4 & & & & -1 & & & & \\ & 4 & & -1 & & -1 & & & \\ & & 4 & & -1 & & & & \\ \hline & -1 & & 4 & & & & -1 & \\ -1 & & -1 & & 4 & & -1 & & -1 \\ & -1 & & & & 4 & & -1 & \\ \hline & & & & -1 & & 4 & & \\ & & & -1 & & -1 & & 4 & \\ & & & & -1 & & & & 4 \end{array}\right] \boldsymbol{u} = \left[\begin{array}{c} b_0 + b_2 + b_{14} \\ b_1 + b_3 \\ b_2 + b_4 + b_6 \\ \hline b_{13} + b_{15} \\ 0 \\ b_5 + b_7 \\ \hline b_{10} + b_{12} + b_{14} \\ b_9 + b_{11} \\ b_6 + b_8 + b_{10} \end{array}\right]$$

where b_j denotes the boundary value at the jth boundary node and these have been numbered clockwise from 0–15 starting at the origin. For general M, the coefficient matrix has the block tridiagonal structure:

$$A = \begin{bmatrix} 4I & T & & & \\ T & 4I & T & & \\ & \ddots & \ddots & \ddots & \\ & & T & 4I & T \\ & & & T & 4I \end{bmatrix}, \text{ where } T = \begin{bmatrix} & -1 & & & \\ -1 & & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & & -1 \\ & & & -1 & \end{bmatrix}.$$

The diagonal blocks of A are a multiple of the identity I and the nonzero off-diagonal blocks T are themselves tridiagonal.
Switching to an odd-even numbering system where a node (x_ℓ, y_m) is even if $\ell + m$ is even. The numbering on a grid with $h = 1/M$ and $M = 4$ is shown below. The numbers outside the box refer to the boundary nodes where, unlike Exercise 10.2, the corner nodes are now involved. There are $N_{\text{even}} = 5$ even nodes (numbered 1–5) and $N_{\text{odd}} = 4$ odd nodes (numbered 6–9).

$$\boldsymbol{u} = [U_{1,1}, U_{1,3}, U_{2,2}, U_{3,1}, U_{3,3}, U_{1,2}, U_{2,1}, U_{2,3}, U_{3,2}]^{\mathsf{T}}$$

In terms of this ordering, we define

4	5	6	7	8
3	**2**	8	**5**	9
2	6	**3**	9	10
1	**1**	7	**4**	11
0	15	14	13	12

$$\boldsymbol{u}_{\text{even}} = [U_1, U_2, U_3, U_4, U_5]^{\mathsf{T}}, \quad \boldsymbol{u}_{\text{odd}} = [U_6, U_7, U_8, U_9]^{\mathsf{T}}, , \quad \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_{\text{even}} \\ \boldsymbol{u}_{\text{odd}} \end{bmatrix}.$$

With this numbering system the equations $\mathcal{L}_h^{\times} U = F$ become, in matrix-vector form,

$$\begin{bmatrix} 4 & & -1 & & \\ & 4 & -1 & & \\ -1 & -1 & 4 & -1 & -1 \\ & & -1 & 4 & \\ & & -1 & & 4 \end{bmatrix} \boldsymbol{u}_{\text{odd}} = \begin{bmatrix} b_0 + b_2 + b_{14} \\ b_2 + b_4 + b_6 \\ \\ b_{10} + b_{12} + b_{14} \\ b_6 + b_8 + b_{10} \end{bmatrix}$$

$$\begin{bmatrix} 4 & -1 & -1 & \\ -1 & 4 & & -1 \\ -1 & & 4 & -1 \\ & -1 & -1 & 4 \end{bmatrix} \boldsymbol{u}_{\text{odd}} = \begin{bmatrix} b_1 + b_3 \\ b_{13} + b_{15} \\ b_5 + b_7 \\ b_9 + b_9 \end{bmatrix}.$$

The general pattern is difficult to detect with such a small value of M but, for even nodes (say), the structure is

$$\begin{bmatrix} 4I_1 & B & & & \\ B^{\mathsf{T}} & 4I_2 & B & & \\ & B^{\mathsf{T}} & 4I_1 & B & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & & & \\ -1 & -1 & & \\ & -1 & -1 & \\ & & \ddots & \ddots \end{bmatrix},$$

where I_1 and I_2 are identity matrices whose dimensions are the number of even nodes in odd and even numbered columns, respectively. The matrix B is square when there are the same number of nodes in consecutive columns, otherwise it is rectangular. The structure is similar for odd nodes with the roles of the pairs I_1, B and I_2, B^{T} being interchanged.
One reason for studying different orderings of unknowns is the profound effect these can have on the time taken to solve the finite difference equations.

10.7
If $Q(0) = Q(h) + Ch^2 + \mathcal{O}(h^3)$ for some constant C, then $Q(0) = Q(h/2) + C(h/2)^2 + \mathcal{O}(h^3)$. Subtracting these gives $Ch^2 = 4(Q(h/2) - Q(h))/3 + \mathcal{O}(h^3)$ and so

$$Q(h) = Q(0) + \tfrac{4}{3}(Q(h) - Q(h/2)) + \mathcal{O}(h^3).$$

Thus $4(Q(h) - Q(h/2))/3 \equiv 4(U_{\mathrm{P}}^h - U_{\mathrm{P}}^{h/2})/3$ gives an estimate of the leading term in the error in $Q(h) \equiv U_{\mathrm{P}}^h$.

10.9
Suppose that the Neumann BC $-u_x = g(0, y)$ is approximated at $y = y_m$ by the second-order central difference formula

$$-\triangle_x U_{0,m} = g(0, mh) \quad \Rightarrow -U_{1,m} + U_{-1,m} = 2hg(0, mh)$$

while (10.6) with $\ell = 0$ gives

$$4U_{0,m} - U_{0,m+1} - U_{1,m} - U_{0,m-1} - U_{-1,m} = h^2 f_{0,m}$$

and adding these two equations gives

$$4U_{0,m} - U_{0,m+1} - 2U_{1,m} - U_{0,m-1} - U_{-1,m} = 2hg_{0,m} + h^2 f_{0,m}$$

which is (10.22b), as required.

10.11
The local truncation error associated with (10.24) is

$$\mathscr{R}_h\big|_{0,M} = -h^{-1}\triangle_x^+ u_{0,M} + h^{-1}\triangle_y^- u_{0,M} - \tfrac{1}{2}hf_{0,M} - g(1^+, 0) - g(0^-, 1).$$

We have, from the previous exercise,

$$h^{-1}\triangle_x^+ u_{0,M} = u_x(0,1) + \tfrac{1}{2}hu_{xx}(0,1) + \tfrac{1}{6}h^2 u_{xxx}(\xi, 1)$$

and by making the change of variables $x \mapsto 1-y$, $y \mapsto 1-x$, $h \mapsto -h$ (so that $h^{-1}\triangle^+ \mapsto -h^{-1}\triangle^-$),

$$h^{-1}\triangle_y^- u_{0,M} = u_y(0,1) - \tfrac{1}{2}hu_{yy}(0,1) + \tfrac{1}{6}h^2 u_{yyy}(0, 1-\eta), \quad 0 < \eta < h.$$

Hence,

$$\begin{aligned}
\mathscr{R}_h\big|_{0,M} = &-\Big(u_x(0,1) + \tfrac{1}{2}hu_{xx}(0,1) + \tfrac{1}{6}h^2 u_{xxx}(\xi,1)\Big) \\
&+ \Big(u_y(0,1) - \tfrac{1}{2}hu_{yy}(0,1) + \tfrac{1}{6}h^2 u_{yyy}(0,1-\eta)\Big) \\
&- \tfrac{1}{2}hf_{0,M} - g(1^+,0) - g(0^-,1) \\
= &\big(-u_x(0,1) - g(1^+,0)\big) + \big(u_y(0,1) - g(0^-,1)\big) - \tfrac{1}{2}h\big(-\nabla^2 u(0,1) - f_{0,M}\big) \\
&- \tfrac{1}{6}h^2\big(u_{xxx}(\xi,1) + u_{yyy}(0,1-\eta)\big)
\end{aligned}$$

and therefore $\mathscr{R}_h\big|_{0,M} = \mathcal{O}(h^2)$ if the third partial derivatives of u are bounded in the neighbourhood of the corner $(0, 1)$.

10.13

The factors are:

Geometry The domain $0 < x, y < 1$ is unaffected by the interchange $x \leftrightarrow y$.

Boundary values $u(x, y) = x^2 + y^2$ is unaffected by the interchange $x \leftrightarrow y$.

PDE Neither the source term $f(x, y) = xy$ nor the differential operator $-\nabla^2 \equiv \partial_x^2 + \partial_y^2$ is affected by the interchange $x \leftrightarrow y$.

Thus $u(x, y) = u(y, x)$.
Since the finite difference grid is also unaffected by the interchange $x \leftrightarrow y$, the numerical solution U must also share the same symmetry, that is $U_{\ell,m} = U_{m,\ell}$. It follows that $U_{\ell,\ell+1} = U_{\ell+1,\ell}$ and $U_{\ell-1,\ell} = U_{\ell,\ell-1}$ so the 5–point formula (10.6) applied at the grid point $(\ell h, \ell h)$ gives

$$\begin{aligned}\mathcal{L}_h U_{\ell,\ell} &= h^{-2}\big(4U_{\ell,\ell} - U_{\ell+1,\ell} - U_{\ell,\ell+1} - U_{\ell-1,\ell} - U_{\ell,\ell-1}\big)\\ &= h^{-2}\big(4U_{\ell,\ell} - U_{\ell+1,\ell} - U_{\ell+1,\ell} - U_{\ell,\ell-1} - U_{\ell,\ell-1}\big)\\ &= h^{-2}\big(4U_{\ell,\ell} - 2U_{\ell+1,\ell} - 2U_{\ell,\ell-1}\big)\end{aligned}$$

and $f_{\ell,\ell} = x_\ell y_\ell = \ell^2 h^2$.
When $M = 4$ then symmetry reduces the number of unknowns to $N = \frac{1}{2}M(M-1) = 6$ and these, together with the boundary nodes, are numbered as shown.

4	5	6	7	8
3	4	5	**6**	7
2	2	**3**	**5**	6
1	**1**	**2**	**4**	5
0	1	2	3	4

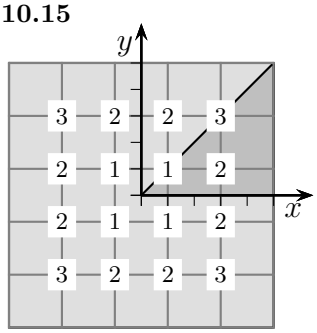
$$\begin{aligned}\boldsymbol{u} &= [U_{1,1}, U_{2,1}, U_{2,2}, U_{3,1}, U_{3,2}, U_{3,3}]^\mathsf{T}\\ \boldsymbol{f} &= \tfrac{1}{16}[1, 2, 4, 2, 6, 9]^\mathsf{T}\\ \boldsymbol{g} &= [2g_1, g_2, 0, g_3 + g_5, g_6, 2g_7]^\mathsf{T} = \tfrac{1}{16}[2, 4, 0, 27, 20, 50]^\mathsf{T},\end{aligned}$$

where g_j is the value of U at the jth boundary node.
The numerical solution is then found by solving $A\boldsymbol{u} = \boldsymbol{f} + \boldsymbol{g}$, where

$$A = 16\begin{bmatrix} 4 & -2 & & & & \\ -1 & 4 & -1 & -1 & & \\ & -2 & 4 & & -2 & \\ & -1 & & 4 & -1 & \\ & & -1 & -1 & 4 & -1 \\ & & & & -2 & 4 \end{bmatrix} = 16D\begin{bmatrix} 2 & -1 & & & & \\ -1 & 4 & -1 & -1 & & \\ & -1 & 2 & & -1 & \\ & -1 & & 4 & -1 & \\ & & -1 & -1 & 4 & -1 \\ & & & & -1 & 2 \end{bmatrix},$$

where the diagonal matrix D is equal to the identity matrix except for entries corresponding to nodes on the diagonal of the grid, where $D_{\ell,\ell} = 2$. The purpose of introducing D being that A is the product of a diagonal matrix and a symmetric matrix. When M is large the cost of solving $(D^{-1}A)\boldsymbol{u} = D^{-1}(\boldsymbol{g} + \boldsymbol{f})$ is approximately half the cost of solving the original system.

10.15

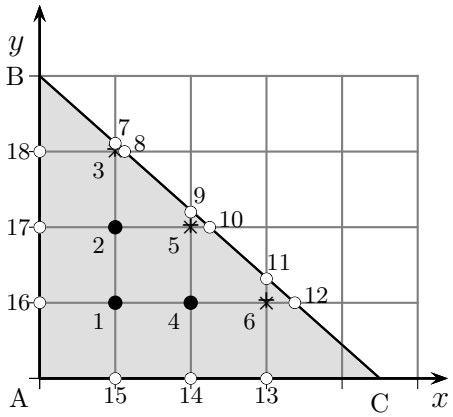


With $M = 5$ ($h = 2/5$) the implications of symmetry are shown in the figure and reveal that there are only 3 independent unknowns. Applying the 5–point approximation of the Poisson equation at nodes 1, 2 & 3 leads to the system:

$$\frac{25}{4}\begin{bmatrix} 2 & -2 & \\ -1 & 3 & -1 \\ & -2 & 4 \end{bmatrix}\begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

For general odd values of M, there are $(M-1)^2$ unknowns (1 per grid point) of which only about one eighth are independent. We therefore have to solve for $\frac{1}{8}(M-1)^2$ unknowns.

10.17



Node	h_+	k_+
3	1/32	1/36
5	2/32	2/36
7	3/32	3/36

The 6 internal grid points and the 12 active boundary nodes are numbered as shown. The standard 5–point finite difference equations

$$-h^{-2}\left(\delta_x^2+\delta_y^2\right)U_{\ell,m}+U_{\ell,m}=0.$$

can only be deployed at points 1, 2 and 4 where there is a uniform grid in both directions. These give:

$$\begin{aligned}
\mathrm{P}_1: &\quad 65U_1-16U_2-16U_4=16(U_{15}+U_{16})=16\left[4x(9-8x)\right]_{x=1/4}=112,\\
\mathrm{P}_2: &\quad -16U_1+65U_2-16U_3-16U_5=16U_{17}=0,\\
\mathrm{P}_4: &\quad -16U_1+65U_4-16U_5-16U_6=16U_{14}=16\left[4x(9-8x)\right]_{x=1/2}=160.
\end{aligned}$$

We focus first on the approximation of u_{xx} at the points 3, 5 and 6. At points 3, 5 and 6 we use equation (10.26) to give

$$\begin{aligned}
\mathrm{P}_3: &\quad h_+=\tfrac{1}{32},\quad h_-=h=\tfrac{1}{4}: \quad (u_{xx})_3\approx\tfrac{64}{9}\left(\tfrac{32}{1}(u_8-u_3)-4(u_3-u_{18})\right),\\
\mathrm{P}_5: &\quad h_+=\tfrac{2}{32},\quad h_-=h=\tfrac{1}{4}: \quad (u_{xx})_5\approx\tfrac{64}{10}\left(\tfrac{32}{2}(u_{10}-u_5)-4(u_5-u_2)\right),\\
\mathrm{P}_6: &\quad h_+=\tfrac{3}{32},\quad h_-=h=\tfrac{1}{4}: \quad (u_{xx})_6\approx\tfrac{64}{11}\left(\tfrac{32}{3}(u_{12}-u_6)-4(u_6-u_4)\right).
\end{aligned}$$

The u_{yy} derivatives are treated similarly.

$$\begin{aligned}
\mathrm{P}_3: &\quad k_+=\tfrac{1}{36},\quad k_-=k=\tfrac{1}{4}: \quad (u_{yy})_3\approx\tfrac{72}{10}\left(\tfrac{36}{1}(u_7-u_3)-4(u_3-u_2)\right),\\
\mathrm{P}_5: &\quad k_+=\tfrac{2}{36},\quad k_-=k=\tfrac{1}{4}: \quad (u_{yy})_5\approx\tfrac{72}{11}\left(\tfrac{36}{2}(u_9-u_5)-4(u_5-u_4)\right),\\
\mathrm{P}_6: &\quad k_+=\tfrac{3}{36},\quad k_-=k=\tfrac{1}{4}: \quad (u_{yy})_6\approx\tfrac{72}{12}\left(\tfrac{36}{1}(u_{11}-u_6)-4(u_6-u_{13})\right).
\end{aligned}$$

Putting these together we find the following finite difference expressions for the approximations to $-\nabla^2u+u=0$ at the 3 grid points closest to the hypotenuse.

$$\begin{aligned}
\mathrm{P}_3: &\quad -\tfrac{64}{9}\left(\tfrac{32}{1}(U_8-U_3)-4(U_3-U_{18})\right)-\tfrac{72}{10}\left(\tfrac{36}{1}(U_7-U_3)-4(U_3-U_2)\right)+U_3 &&=0\\
&\qquad\qquad 545U_3-\tfrac{144}{5}U_2 &&=0\\
\mathrm{P}_5: &\quad -\tfrac{64}{10}\left(\tfrac{32}{2}(U_{10}-U_5)-4(U_5-U_2)\right)-\tfrac{72}{11}\left(\tfrac{36}{2}(U_9-U_5)-4(U_5-U_4)\right)+U_5 &&=0\\
&\qquad\qquad 273U_5-\tfrac{128}{5}U_2-\tfrac{288}{11}U_4 &&=0\\
\mathrm{P}_6: &\quad -\tfrac{64}{11}\left(\tfrac{32}{3}(U_{12}-U_6)-4(U_6-U_4)\right)-\tfrac{72}{12}\left(\tfrac{36}{1}(U_{11}-U_6)-4(U_6-U_{13})\right) &&=0\\
&\qquad\qquad \tfrac{979}{3}U_6-\tfrac{256}{11}U_4 &&=24U_{13},
\end{aligned}$$

where $U_{13} = [4x(9-8x)]_{x=3/4} = 9$. These provide 6 equations in the 6 unknowns.

10.19
In view if the identity

$$2(U_{\ell,1} - U_{\ell,0}) = (U_{\ell,1} - 2U_{\ell,0} + U_{\ell,-1}) + (U_{\ell,1} - U_{\ell,-1}) = \delta_\theta^2 U_{\ell,0} + 2\triangle_\theta U_{\ell,0}$$

equation (10.33b) may be written

$$\mathcal{B}_h U_{\ell,0} := -\Delta\theta^{-1}\triangle_\theta U_{\ell,0} - \tfrac{1}{2}r_\ell^2 \Delta\theta \mathcal{L}_h U_{\ell,0},$$

where $\mathcal{L}_h$ (defined by (10.33)) is an approximation of $\mathcal{L}u = (1/r)(ru_r)_r + u_{\theta\theta}/r^2$. The local truncation error of the BC $\mathcal{B}_h U_{\ell,0} = 0$ is

$$\begin{aligned}\mathscr{R}_h\big|_{\ell,0} :=& \mathcal{B}_h u_{\ell,0} = \Delta\theta^{-1}\triangle_\theta u_{\ell,0} - \tfrac{1}{2}r_\ell^2\Delta\theta\mathcal{L}_h u_{\ell,0}\\ =& \big[-u_\theta + \mathcal{O}(\Delta\theta^2)\big]_{\ell,0} - \tfrac{1}{2}r_\ell^2\Delta\theta\big[\mathcal{L}u + \mathcal{O}(h^2)\big]_{\ell,0}\\ =& \big[-u_\theta - \tfrac{1}{2}r_\ell^2\Delta\theta\mathcal{L}u\big]_{\ell,0} + \mathcal{O}(\Delta\theta^2) + \mathcal{O}(h^2) = \mathcal{O}(\Delta\theta^2) + \mathcal{O}(h^2)\end{aligned}$$

as required (since $u_\theta = 0$ and $\mathcal{L}u = 0$ at $(\ell h, 0)$).

10.21
With

$$u(r,\theta) = r^2\Big(\frac{1}{\pi}\big[\log(\frac{1}{r})\sin 2\theta + (\frac{\pi}{4} - \theta)r^2\cos 2\theta\big] - \frac{\pi}{4}\Big)r^2$$

we find that $u(r,0) = 0$, $u(r,\frac{1}{2}\pi) = 0$ and $u(1,\theta) = \frac{1}{4}(1 - 4\theta/\pi)\cos 2\theta - \frac{1}{4}$.
Consider the functions $f(r,\theta) = r^2\log(\frac{1}{r})\sin 2\theta$, $g(r,\theta) = (\frac{\pi}{4} - \theta)r^2\cos 2\theta$ and $h(r,\theta) = \frac{1}{4}r^2$.

$$\begin{aligned} f_r &= -r(2r\log(r) + 1)\sin(2\theta), & f_\theta &= 2r^2\log(\frac{1}{r})\cos 2\theta\\ f_{rr} &= -(2r\log(r) + 3)\sin(2\theta), & f_{\theta\theta} &= 4r^2\log(r)\sin 2\theta\end{aligned}$$

so that $\nabla^2 f := f_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}f_{\theta\theta} = -4\sin 2\theta$. Similarly

$$\begin{aligned} g_r &= 2(\frac{\pi}{4} - \theta)r\cos 2\theta, & g_\theta &= -r^2\cos(2\theta) - 2\left(\tfrac{1}{4}\pi - \theta\right)r^2\sin 2\theta\\ g_{rr} &= 2(\frac{\pi}{4} - \theta)\cos 2\theta, & g_{\theta\theta} &= 4r^2\sin 2\theta - 4(\tfrac{1}{4}\pi - \theta)r^2\cos 2\theta\end{aligned}$$

and so $\nabla^2 g = 4\sin 2\theta$. Hence $\mathcal{L}(f+g) = 0$: $f+g$ is a harmonic function. Also $\nabla^2 h = -1$. Thus $-\nabla^2 u = 1$ since $u = f + g + h$.
Also,

$$u_{rr} = -2\log(r)\sin 2\theta + \text{terms involving } \theta \text{ only}$$

so that $|u_{rr}| \to 0$ as $r \to 0$ so long as $\sin 2\theta \neq 0$.

10.23
We rewrite (10.41) as

$$\begin{aligned}\mathcal{L}_h U_{\ell,m} = -h_x^{-1}h_y^{-1}\Big(&\frac{1}{\rho}a(U_{\ell-1,m} - 2U_{\ell,m} + U_{\ell+1,m}) + \tfrac{1}{2}b(U_{\ell+1,m+1} - U_{\ell-1,m+1} + U_{\ell-1,m-1} - U_{\ell+1,m-1})\\ &+ c\rho(U_{\ell,m-1} - 2U_{\ell,m} + U_{\ell,m+1})\\ -\tfrac{1}{2}\gamma(&4U_{\ell,m} - 2U_{\ell+1,m} + U_{\ell+1,m+1} - 2U_{\ell,m-1} + U_{\ell-1,m-1} - 2U_{\ell-1,m} + U_{\ell-1,m-1} - 2U_{\ell,m-1} + U_{\ell+1,m-1})\Big).\end{aligned}$$

This is of the form

$$\mathcal{L}_h u\big|_P = h_x^{-1}h_y^{-1}\Big(\alpha_0 U\big|_{\mathrm{P}} - \sum_{j=1}^{8}\alpha_j U\big|_{\mathrm{Q}_j}\Big)$$

with (these are easiest to calculate by adjusting the coefficients in Fig. 10.15)

$$\begin{aligned}
&\alpha_0 = 2a/\rho + 2c\rho - 2\gamma, && \alpha_1 = a/\rho - \gamma, && \alpha_2 = -\tfrac{1}{2}b + \tfrac{1}{2}\gamma\\
&\alpha_3 = c\rho - \gamma, && \alpha_4 = \tfrac{1}{2}b + \tfrac{1}{2}\gamma, && \alpha_5 = \alpha_1\\
&\alpha_6 = -\tfrac{1}{2}b + \tfrac{1}{2}\gamma, && \alpha_7 = \alpha_3, && \alpha_8 = \tfrac{1}{2}b + \tfrac{1}{2}\gamma.
\end{aligned}$$

These coefficients will be non-negative if γ is chosen such that $|b| \le \gamma$, $\gamma \le a/\rho$ and $\gamma \le c\rho$. When $\rho = \sqrt{a/c}$ these become

$$|b| \le \gamma \le \sqrt{ac}$$

which can always be achieved when $b^2 \le ac$.

10.25
Using $\triangle_x^- = \triangle - \frac{1}{2}\delta_x^2$, then the analogue of the operator $\mathcal{L}_h^-$ defined by (10.51) that is appropriate for the PDE $-\varepsilon\nabla^2 u + pu_x = 0$ with $p > 0$ may be written as

$$\begin{aligned}
\mathcal{L}_h^- U_{\ell,m} &= -\varepsilon h^{-2}(\delta_x^2 + \delta_y^2)U_{\ell,m} + ph^{-1}\triangle_x^- U_{\ell,m}\\
&= -(\varepsilon h^{-2} + \tfrac{1}{2}ph^{-1})\delta_x^2 - \varepsilon h^{-2}\delta_y^2 U_{\ell,m} + ph^{-1}\triangle_x U_{\ell,m}\\
&= -\varepsilon h^{-2}(1 + \mathrm{Pe}_h)\delta_x^2 - \varepsilon h^{-2}\delta_y^2 U_{\ell,m} + ph^{-1}\triangle_x U_{\ell,m},
\end{aligned}$$

where $\mathrm{Pe}_h = ph/(2\varepsilon)$.
When $p < 0$, the backward difference $\triangle_x^-$ should be replaced by a forward difference $\triangle_x^+$ and, with the identity $\triangle_x^+ = \triangle + \frac{1}{2}\delta_x^2$, the corresponding manipulations are

$$\begin{aligned}
\mathcal{L}_h^+ U_{\ell,m} &= -\varepsilon h^{-2}(\delta_x^2 + \delta_y^2)U_{\ell,m} + ph^{-1}\triangle_x^+ U_{\ell,m}\\
&= -(\varepsilon h^{-2} - \tfrac{1}{2}ph^{-1})\delta_x^2 - \varepsilon h^{-2}\delta_y^2 U_{\ell,m} + ph^{-1}\triangle_x U_{\ell,m}\\
&= -\varepsilon h^{-2}(1 - \mathrm{Pe}_h)\delta_x^2 - \varepsilon h^{-2}\delta_y^2 U_{\ell,m} + ph^{-1}\triangle_x U_{\ell,m}.
\end{aligned}$$

Here $\mathrm{Pe}_h < 0$ so that $\mathrm{Pe}_h = -|\mathrm{Pe}_h|$ and therefore both cases are accommodated in the single formula

$$\mathcal{L}_h^{\pm} U_{\ell,m} := -\varepsilon h^{-2}(1 + |\mathrm{Pe}_h|)\delta_x^2 U_{\ell,m} - \varepsilon h^{-2}\delta_y^2 U_{\ell,m} + h^{-1}p\triangle_x U_{\ell,m}.$$

The generalization to the PDE $-\varepsilon(u_{xx} + u_{yy}) + pu_x + qu_y = 0$, is

$$-\varepsilon h^{-2}(1 + |\mathrm{Pe}_{h,x}|)\delta_x^2 U_{\ell,m} - \varepsilon h^{-2}(1 + |\mathrm{Pe}_{h,y}|)\delta_y^2 U_{\ell,m} + h^{-1}p\triangle_x U_{\ell,m} + h^{-1}q\triangle_y U_{\ell,m} = 0,$$

where $\mathrm{Pe}_{h,x} := ph/(2\varepsilon)$ and $\mathrm{Pe}_{h,y} := qh/(2\varepsilon)$.

Exercises 11 Finite difference methods for parabolic PDEs

11.1
Differentiating $u_t = u_{xx}$ with respect to t we find $u_{tt} = u_{xxt}$ while differentiating twice with respect to x leads to $u_{xxxx} = u_{xxt}$. Hence (11.12) becomes

$$\begin{aligned}\mathcal{R}_m^n &= \tfrac{1}{2}ku_{tt}\big|_m^n - \tfrac{1}{12}h^2u_{xxxx}\big|_m^n + \mathcal{O}(k^2) + \mathcal{O}(h^4)\\ &= \tfrac{1}{2}h^2(r-\tfrac{1}{6})u_{xxt}\big|_m^n + \mathcal{O}(k^2) + \mathcal{O}(h^4).\end{aligned}$$

When $r = 1/6$ the order of consistency is $\mathcal{O}(k^2) + \mathcal{O}(h^4)$.

11.3
Equations (11.8b) give, for $m = 1, 2, \ldots, M-1$,

$$\begin{aligned}U_1^{n+1} &= rg_0(nk) + (1-2r)U_1^n + rU_2^n\\ U_2^{n+1} &= \quad rU_1^n + (1-2r)U_2^n + rU_3^n\\ &\vdots \qquad\quad \ddots \qquad\quad \ddots \qquad\quad \ddots\\ U_{M-2}^{n+1} &= \quad rU_{M-3}^n + (1-2r)U_{M-2}^n + rU_{M-1}^n\\ U_{M-1}^{n+1} &= \quad rU_{M-2}^n + (1-2r)U_{M-1}^n + rg_1(nk)\end{aligned}$$

which, in matrix-vector form, become

$$\boldsymbol{u}^{n+1} = \begin{bmatrix} 1-2r & r & 0 & \cdots & 0\\ r & 1-2r & r & & \\ 0 & r & \ddots & \ddots & \\ & & \ddots & & r\\ 0 & & & r & 1-2r \end{bmatrix}\boldsymbol{u}^n + r\begin{bmatrix} g_0(nk)\\ 0\\ \vdots\\ 0\\ 0\\ g_1(nk)\end{bmatrix}$$

where $\boldsymbol{u}^n = [U_1^n, U_2^n, \ldots, U_{M-1}^n]^\mathsf{T}$. The coefficient matrix is clearly the same as that in the BTCS scheme (11.25) with r replaced by $-r$.

11.5
The finite difference scheme $U_m^{n+1} = \frac{1}{3}rU_{m-2}^n + (1-r)U_m^n + \frac{2}{3}rU_{m+1}^n$ can be written in the form (11.10) with

$$\mathcal{L}_hU_m^n = -\tfrac{1}{3}h^{-2}(U_{m-2}^n - 3U_m^n + 2U_{m+1}^n).$$

Using the Taylor expansions of u_{m-2}^n and u_{m+1}^n about the point (x_m, t^n), we find that (all the terms on the right hand side are evaluated at (x_m, t^n))

$$\begin{aligned}\mathcal{L}_hu_m^n = -\tfrac{1}{3}h^{-2}\big(\quad & u - 2hu_x + \frac{1}{2!}(2h)^2u_{xx} - \frac{1}{3!}(2h)^3u_{xxx} + \mathcal{O}(h^4)\\ & - 3u\\ & + 2u + 2hu_x + 2\frac{1}{2!}h^2u_{xx} + \frac{1}{3!}h^3u_{xxx} + \mathcal{O}(h^4)\big)\\ & = -u_{xx} + \mathcal{O}(h).\end{aligned}$$

Since $\mathcal{L}_hu = \mathcal{L}u + \mathcal{O}(h)$, $\mathcal{L}_h$ is first order consistent with $\mathcal{L}$ and, from (11.11), the given method is therefore consistent with the heat equation. The stencil and domain of dependence are shown in Fig. 11.

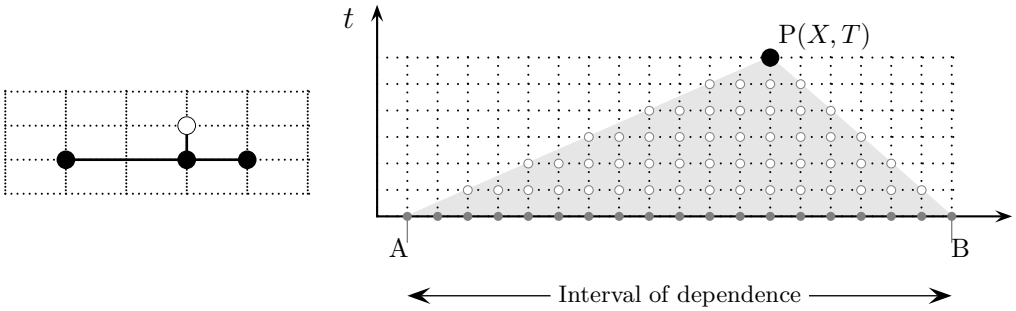


Figure 11: Stencil (left) and domain of dependence (right) for Exercise 11.5.

11.7
Replacing U by $-U$ in Theorem 11.5 we have

$$\begin{aligned} -h^{-2}\delta_x^2(-U_m^n) + k^{-1}\big((-U_m^{n+1}) - (-U_m^n)\big) &\leq 0 \\ \text{i.e., } -h^{-2}\delta_x^2 U_m^n + k^{-1}(U_m^{n+1} - U_m^n) &\geq 0 \end{aligned}$$

for $(x_m, t_{n+1}) \in \Omega_\tau$. Thus, if $r = k/h^2 \leq 1/2$ then U is either constant or else attains its *minimum* ($-U$ attains its maximum) value on Γ_τ.
If $-h^{-2}\delta_x^2 U_m^n + k^{-1}(U_m^{n+1} - U_m^n) = 0$ then U is either constant or else attains its maximum and minimum value on Γ_τ. The case when U is constant can only occur if U is the same constant on Γ_τ. So, when $U_0^n = U_M^n = 0$, the maximum absolute value must occur at $t = 0$.

11.9
The local truncation error of the BTCS approximation (11.22) is, by definition,

$$\mathcal{R}_h\big|_m^n := \frac{1}{k}\big(u_m^{n+1} - u_m^n - r\delta_x^2 u_m^{n+1}\big) = \frac{1}{k}(u_m^{n+1} - u_m^n) - h^{-2}\delta_x^2 u_m^{n+1}$$

and, using $u_m^n = u_m^{n+1} - ku_t|_m^{n+1} + \frac{1}{2}k^2 u_{tt}|_m^{n+1} + \mathcal{O}(k^3)$ together with $\delta_x^2 u_m^{n+1} = u_{xx}|_m^{n+1} + \frac{1}{12}h^2 u_{xxxx}|_m^{n+1} + \mathcal{O}(h^4)$, we find

$$\begin{aligned} \mathcal{R}_h\big|_m^n &= \big(u_t|_m^{n+1} - \tfrac{1}{2}ku_{tt}|_m^{n+1} + \mathcal{O}(k^2)\big) - \big(u_{xx}|_m^{n+1} + \tfrac{1}{12}h^2 u_{xxxx}|_m^{n+1} + \mathcal{O}(h^4)\big) \\ &= \big(u_t - u_{xx}\big)\big|_m^{n+1} - \tfrac{1}{2}h^2\big(ru_{tt} + \tfrac{1}{6}u_{xxxx}\big)\big|_m^{n+1}\big) + \mathcal{O}(k^2) + \mathcal{O}(h^4) \end{aligned}$$

and so $\mathcal{R}_h = \mathcal{O}(k) + \mathcal{O}(h^2)$. Since $u_{tt} = u_{txx} = \partial_x^2 u_t = u_{xxxx}$, this becomes

$$\mathcal{R}_h\big|_m^n = -\tfrac{1}{2}h^2\big(r + \tfrac{1}{6}\big)u_{txx}\big|_m^{n+1}\big) + \mathcal{O}(k^2) + \mathcal{O}(h^4).$$

11.11
Replacing U by $-U$ in Theorem 11.11 we deduce that

$$\begin{aligned} -h^{-2}\delta_x^2(-U_m^{n+1}) + k^{-1}\big((-U_m^{n+1}) - (-U_m^n)\big) &\leq 0 \\ \text{i.e., } -h^{-2}\delta_x^2 U_m^{n+1} + k^{-1}(U_m^{n+1} - U_m^n) &\geq 0 \end{aligned}$$

and so U is either constant or else attains its *minimum* value on Γ_τ when $r = k/h^2 \leq 1/2$. Thus, if $-h^{-2}\delta_x^2 U_m^{n+1} + k^{-1}(U_m^{n+1} - U_m^n) = 0$ then U is either constant or else attains its maximum and

minimum value on Γ_τ. The case when U is constant can only occur if U is the same constant on Γ_τ. So, when $U_0^n = U_M^n = 0$, the maximum absolute value must occur at $t = 0$.

11.13
With $\mathcal{L}_h U = -h^{-2}\delta_x^2 U$, the defining equation (11.30) of the θ–method becomes

$$U_m^{n+1} - \theta r\delta_x^2 U_m^{n+1} = U_m^n + (1-\theta)r\delta_x^2 U_m^n$$

i.e.,

$$-r\theta U_{m-1}^{n+1} + (1+2r\theta)U_m^{n+1} - r\theta U_{m+1}^{n+1} = r(1-\theta)U_{m-1}^n + (1-2(1-\theta)r)U_m^n + r(1-\theta)U_{m+1}^n$$

so, with the given BCs, these give, for $m = 1$ and $m = M-1$,

$$(1+2r\theta)U_1^{n+1} - r\theta U_2^{n+1} = (1-(1-2\theta)r)U_1^n + r(1-\theta)U_2^n + r\big((1-\theta)g_0(t_n) + \theta g_0(t_{n+1})\big)$$
$$-r\theta U_{M-2}^{n+1} + (1+2r\theta)U_{M-1}^{n+1} = r(1-\theta)U_{M-2}^n + (1-(1-2\theta)r)U_{M-1}^n + r\big((1-\theta)g_1(t_n) + \theta g_1(t_{n+1})\big).$$

Thus, for $(M-1)\times(M-1)$ matrices B and C:

$$B\boldsymbol{u}^{n+1} = C\boldsymbol{u}^n + \boldsymbol{f}^n$$

where

$$B = \begin{bmatrix} 1+2r\theta & -r\theta & 0 & \cdots & 0 \\ -r\theta & 1+2r\theta & -r\theta & & \\ 0 & -r\theta & \ddots & \ddots & \\ & & \ddots & & -r\theta \\ 0 & & & -r\theta & 1+2r\theta \end{bmatrix} = A(\theta r)$$

$$C = \begin{bmatrix} 1-2r(1-\theta) & r(1-\theta) & 0 & \cdots & 0 \\ r(1-\theta) & 1-2r(1-\theta) & r(1-\theta) & & \\ 0 & r(1-\theta) & \ddots & \ddots & \\ & & \ddots & & r\theta \\ 0 & & & r(1-\theta) & 1-2r(1-\theta) \end{bmatrix} = A\big((\theta-1)r\big),$$

where $A \equiv A(r)$ is the coefficient matrix appearing in (11.25). Also,

$$\boldsymbol{f}^n = r[(1-\theta)g_0(t_n) + \theta g_0(t_{n+1}), 0, \ldots, 0, (1-\theta)g_1(t_n) + \theta g_1(t_{n+1})]^{\mathsf{T}}.$$

11.15
Replacing U by $-U$ in Theorem 11.16 we deduce that

$$-\tfrac{1}{2}h^{-2}\delta_x^2\big((-U_m^{n+1}) + (-U_m^n)\big) + k^{-1}\big((-U_m^{n+1}) - (-U_m^n)\big) \le 0$$
$$\text{i.e., } -\tfrac{1}{2}h^{-2}\delta_x^2(U_m^{n+1} + U_m^n) + k^{-1}(U_m^{n+1} - U_m^n) \ge 0$$

for $(x_m, t_{n+1}) \in \Omega_\tau$. If $r = k/h^2 \le 1$ then, from Theorem 11.16, $-U$ is either constant or else attains its maximum value—therefore U attains its *minimum* value, on Γ_τ.
Thus, if $-\frac{1}{2}h^{-2}\delta_x^2(U_m^{n+1} + U_m^n) + k^{-1}(U_m^{n+1} - U_m^n) = 0$ then U is either constant or else attains both its maximum and minimum value on Γ_τ. The case when U is constant can only occur if U

is the same constant on Γ_τ. So, when $U_0^n = U_M^n = 0$, the maximum absolute value must occur at $t = 0$.

11.17
We use two properties of a sum of non-negative terms: it is always smaller than the (largest term) $\times$ (the number of terms) and it is always larger than the largest single term. The norm

$$\|U.\|_{h,2} = \left(h \sum_{m=0}^{M}{}'' |U_m|^2 \right)^{1/2} \leq \left(h \sum_{m=0}^{M}{}'' 1 \right)^{1/2} \max_{0 \leq m \leq M} |U_m| = \|U.\|_{h,\infty}.$$

If the largest term among $\{U_m^n\}_{m=0}^M$ occurs at $m = 0$ or $m = M$, then $\|U.\|_{h,\infty} = |U_0|$ or $|U_M|$ and $\|U.\|_{h,2} \geq \frac{1}{2}\sqrt{h}\|U.\|_{h,\infty}$.
If $\|U.\|_{h,\infty} = |U_m|$ with $0 < m < M$, then $\|U.\|_{h,2} \geq \sqrt{h}\|U.\|_{h,\infty}$. Thus, in general,

$$\tfrac{1}{2}\sqrt{h}\|U.\|_{h,\infty} \leq \|U.\|_{h,2} \leq \|U.\|_{h,2}.$$

For the two possible grid functions a simple calculation reveals

Case	$\|U.\|_{h,\infty}$	$\|U.\|_{h,2}$	
(a)	1	1	Right hand bound attained
(b)	1	$\frac{1}{2}\sqrt{h}$	Left hand bound attained

These examples show that the bounds are attained so they cannot be improved upon.

11.19
From the solution to Exercise 11.13, the θ-method applied to the heat equation gives

$$U_m^{n+1} - \theta r \delta_x^2 U_m^{n+1} = U_m^n + (1-\theta) r \delta_x^2 U_m^n.$$

We now substitute $U_m^n = \xi^n \mathrm{e}^{\mathrm{i}\kappa m h}$ and use the relationships $U_m^{n+1} = \xi U_m^n$ and (11.47b) to give (we define $s = \sin^2(\frac{1}{2}\kappa h)$ as a convenient abbreviation)

$$[1 + 4r\theta s]\xi U_m^n = [1 - 4r(1-\theta)s] U_m^n \quad \Rightarrow \quad \xi = \frac{1 - 4r(1-\theta)s}{1 + 4r\theta s}.$$

We can rewrite this as

$$\xi = 1 - \frac{4rs}{1 + 4r\theta s}$$

so $\xi \leq 1$ for all $r > 0$ and $\theta \geq 0$. In order that $\xi \geq -1$ we require

$$1 - 4r(1-\theta)s \geq -1 - 4r\theta s \quad \Rightarrow \quad 2r(1-2\theta)s \leq 1.$$

This holds for all $r > 0$ when $\theta \geq \frac{1}{2}$ otherwise, it will hold for all $s \in [0,1]$ only when r is restricted by

$$r \leq \frac{1}{2(1-2\theta)}.$$

This reduces, as it should, to $r \leq \frac{1}{2}$ when $\theta = 0$ and the method becomes the FTCS scheme.

11.21
The FTCS scheme (11.53)

$$\begin{aligned} U_m^{n+1} &= U_m^n + r\delta_x^2 U_m^n + c\triangle_x U_m^n \\ &= (r - \tfrac{1}{2}\rho) U_{m-1}^n + (1-2r) U_m^n + (r + \tfrac{1}{2}\rho) U_{m+1}^n \end{aligned}$$

is of the form

$$U_m^{n+1} = \alpha_{-1}U_{m-1}^n + \alpha_0 U_m^n + \alpha_1 U_{m+1}^n$$

with $\alpha_{-1} = r - \frac{1}{2}\rho, \alpha_0 = 1 - 2r, \alpha_1 = r + \frac{1}{2}\rho$. Non-negativity of the coefficients requires $\frac{1}{2}\rho \le r \le \frac{1}{2}$. The left inequality simplifies to $h \le 2\varepsilon$ (so that stability is only possible if the spatial grid size is sufficiently small compared to ε), that is, $\mathrm{Pe}_h \le 1$, where $\mathrm{Pe}_h := h/(2\varepsilon)$ is the mesh Peclet number. The right inequality leads to $k \le h^2/2\varepsilon$—the stability region is shown in Fig. 12 (Left).
The conditions $\frac{1}{2}\rho^2 \le r \le \frac{1}{2}$ for ℓ_2-stability require $k \le 2\varepsilon$ and $k \le h^2/2\varepsilon$, both coinciding when $\mathrm{Pe}_h = 1$. The corresponding stability region is shown in Fig. 12 (Right). Thus the scheme is ℓ_2-stable for any spatial grid size.

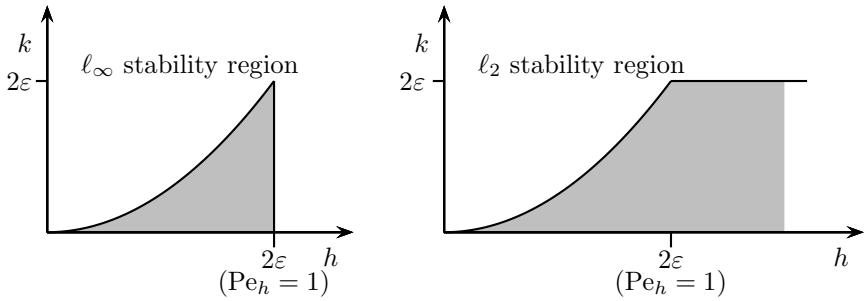


Figure 12: The stability regions for Exercise 11.21.

The proof of Theorem 11.5 may be used to establish a maximum principle under the conditions $\frac{1}{2}\rho \le r \le \frac{1}{2}$. The only change necessary is the derivation of an upper bound on U_m^{j+1}: This now reads

$$U_m^{j+1} \le \alpha_{-1}U_{m-1}^j + \alpha_0 U_m^j + \alpha_1 U_{m+1}^j$$

and, because $U_{m-1}^j, U_m^j, U_{m+1}^j \le K_\tau$ and the non-negativity of the coefficients

$$U_m^{j+1} \le \big(\alpha_{-1} + \alpha_0 + \alpha_1\big)K_\tau = K_\tau$$

since $\alpha_{-1} + \alpha_0 + \alpha_1 = 1$.

11.23
We substitute $U_m^n = \xi^n \mathrm{e}^{\mathrm{i}\kappa mh}$ into the finite difference scheme

$$U_m^{n+1} = \tfrac{1}{3}rU_{m-2}^n + (1-r)U_m^n + \tfrac{2}{3}rU_{m+1}^n$$

and use the relationships $U_m^{n+1} = \xi U_m^n$, $U_{m-2}^n = \mathrm{e}^{-2\mathrm{i}\kappa mh}U_m^n$ and $U_{m+1}^n = \mathrm{e}^{\mathrm{i}\kappa mh}U_m^n$ to give

$$\xi U_m^n = \tfrac{1}{3}r\mathrm{e}^{-2\mathrm{i}\kappa h}U_m^n + (1-r)U_m^n + \tfrac{2}{3}r\mathrm{e}^{\mathrm{i}\kappa h}U_m^n \quad \Rightarrow \quad \xi = 1 - r + \tfrac{1}{3}r(2\mathrm{e}^{\mathrm{i}\kappa h} + \mathrm{e}^{-2\mathrm{i}\kappa h}).$$

Extracting real and imaginary parts of the right hand side we have

$$\xi = (1 - r - r\cos\kappa h) + \tfrac{1}{3}\mathrm{i}r\sin\kappa h = (1 - 2r\sin^2\tfrac{1}{2}\kappa h) + \tfrac{1}{3}\mathrm{i}r\sin\kappa h$$

so that, after expressing $\sin\kappa h$ in terms of half-angles,

$$|\xi|^2 - 1 = -4rs\big[1 - r + \tfrac{1}{9}r(1-s)\big], \qquad s = \sin^2\tfrac{1}{2}\kappa h.$$

Hence $|\xi|^2 - 1 \le 0$ if the bracketed term on the right hand side is positive for $s \in [0,1]$, Since it is a linear expression in s, only its end-points need to be examined. It is non-negative at $s = 0$ if $r \le 9/8$ and at $s = 1$ if $r \le 1$. The method is therefore ℓ_2 stable for $r \le 1$.

11.25
With the Robin end condition $-u_x(0,t) + \sigma u(0,t) = g_0(t)$, the numerical boundary condition (11.57) is replaced by

$$-\tfrac{1}{2}h^{-1}(-U_{-1}^n + U_1^n) + \sigma U_0^n = g_0(nk).$$

Using (11.59) to eliminate U_{-1}^n leads to

$$\frac{1}{2hr}\big(U_0^{n+1} - (1 - 2r - 2hr\sigma)U_0^n - 2rU_1^n\big) = g_0(nk).$$

or, for computational purposes,

$$U_0^{n+1} = (1 - 2r - 2hr\sigma)U_0^n + 2rU_1^n + 2rhg_0(nk).$$

However, it is the former version that is correctly scaled for determining the local truncation error (all the terms involving u on the right hand side in the following expansions are evaluated at $(0, nk)$):

$$\begin{aligned}
\mathcal{R}_h\big|_0^n &= \frac{1}{2hr}\big(u_0^{n+1} - (1 - 2r - 2hr\sigma)u_0^n - 2ru_1^n\big) - g_0(nk) \\
&= \frac{1}{2hr}\big(u + ku_t + \tfrac{1}{2}k^2u_{tt} + \mathcal{O}(k^3) \\
&\qquad - (1 - 2r - 2hr\sigma)u \\
&\qquad - 2r(u + hu_x + \tfrac{1}{2}h^2u_{xx} + \tfrac{1}{6}h^3u_{xxx} + \mathcal{O}(h^4))\Big) - g_0(nk) \\
&= \frac{1}{2hr}\big((1 - (1 - 2r - 2hr\sigma) - 2r)u + ku_t - 2hru_x + \tfrac{1}{2}k^2u_{tt} - ku_{xx} + \cdots\big) - g_0(nk) \\
&= (\sigma u - u_x - g_0(nk)) + \tfrac{1}{2}h\big(u_t - u_{xx}\big) + \tfrac{1}{6}h(3ku_{tt} - 2hu_{xxx}) + \cdots .
\end{aligned}$$

Hence $\mathcal{R}_h = \mathcal{O}(hk) + \mathcal{O}(h^2)$ is consistent of second order in h since $k = \mathcal{O}(h)$, $u(0,t) - u_x(0,t) = g_0(t)$ and $u_t = u_{xx}$.

11.27
The Crank–Nicolson scheme (11.34b) for the heat equation at $m = 0$ is

$$-\tfrac{1}{2}rU_{-1}^{n+1} + (1 + r)\,U_0^{n+1} - \tfrac{1}{2}rU_1^{n+1} = \tfrac{1}{2}rU_{-1}^n + (1 - r)\,U_0^n + \tfrac{1}{2}rU_1^n.$$

and using (11.57) at $t = nk$ and also with n replaced by $n + 1$ gives

$$\begin{aligned}
-\tfrac{1}{2}h^{-1}(-U_{-1}^n + U_1^n) = g_0(nk) \qquad &\Rightarrow \quad U_{-1}^n \;\; = U_1^n + 2hg_0(nk) \\
-\tfrac{1}{2}h^{-1}(-U_{-1}^{n+1} + U_1^{n+1}) = g_0((n+1)k) \qquad &\Rightarrow \quad U_{-1}^{n+1} = U_1^{n+1} + 2hg_0((n+1)k).
\end{aligned}$$

On substituting for U_{-1}^n and U_{-1}^{n+1} the Crank–Nicolson scheme then becomes

$$(1 + r)\,U_0^{n+1} - rU_1^{n+1} = (1 - r)\,U_0^n + rU_1^n + hr\big(g_0(nk) + g_0((n+1)k)\big).$$

11.29
We prove the result by induction on n. Let $K_0 = \|U^0\|_{h,\infty}$, then the induction hypothesis $\|U^n\|_{h,\infty} \le \|U^0\|_{h,\infty}$ is satisfied at $n = 0$. Let us suppose that it holds up until $n = j$ so that $\|U^j\|_{h,\infty} \le K_0$. We need to prove that $\|U^{j+1}\|_{h,\infty} \le K_0$ when $r \le \frac{1}{2}$. This is done in a two step process, exploiting the one-dimensional nature of the factors constituting the scheme. Let $V^{j+1}_{\ell,m} = (1 + r\delta_y^2)U^j_{\ell,m}$ then

$$\begin{aligned} |V^{j+1}_{\ell,m}| &= \left|rU^j_{\ell,m-1} + (1-2r)U^j_{\ell,m} + rU^j_{\ell,m+1}\right| \\ &\le \big(r + |1-2r| + r\big)K_0 = K_0 \end{aligned}$$

since $1 - 2r \ge 0$. A similar argument with $U^{j+1}_{\ell,m} = (1 + r\delta_y^2)V^{j+1}_{\ell,m}$ gives

$$\begin{aligned} |U^{j+1}_{\ell,m}| &= \left|rV^j_{\ell-1,m} + (1-2r)V^j_{\ell,m} + rV^j_{\ell+1,m}\right| \\ &\le \big(r + |1-2r| + r\big)K_0 = K_0 \end{aligned}$$

and so the induction hypothesis holds with $n = j+1$. Thus $\|U^n\|_{h,\infty} \le K_0$ holds for $n = 0, 1, \ldots$ provided that $r \le \frac{1}{2}$.

11.31
(a) The given finite difference scheme may be written

$$U^{n+1}_m = U^n_m + \frac{k}{h^2 r_m}\left(r_{m+1/2}(U^n_{m+1} - U^n_m) - r_{m-1/2}(U^n_m - U^n_{m-1})\right),$$

Since $r_{m\pm1/2} = r_m \pm h/2$, this may be written as

$$\begin{aligned} U^{n+1}_m &= U^n_m + \frac{k}{h^2 r_m}\left(r_m(U^n_{m+1} - 2U^n_m + U^n_{m-1}) + \tfrac{1}{2}h(U^n_{m+1} - U^n_{m-1})\right) \\ k^{-1}(U^{n+1}_m - U^n_m) &= h^{-2}\delta_r^2 U^n_m + \frac{1}{r_m}h^{-1}\triangle_r U^n_m \end{aligned}$$

which is the same as the scheme obtained by replacing the spatial dervatives in $u_{rr} + \frac{1}{r}u_r$ with second order accurate differences and the time derivative by a forward difference $k^{-1}\triangle_t^+$.
(b) We may also write the scheme as

$$U^{n+1}_m = \frac{k}{h^2}(1 - \frac{h}{2r_m})U^n_{m-1} + (1 - 2\frac{k}{h^2})U^n_m + \frac{k}{h^2}(1 + \frac{h}{2r_m})U^n_{m+1}$$

which is of positive type provided that $k \le h^2/2$ (the formula is valid only for $m \ge 1$ so $h/(2r_m) = 1/(2m) \le \frac{1}{2}$).
It was shown in Exercise 8.7 that $u_r \to 0$ as $r \to 0$ when u possesses circular symmetry. Hence u_r/r is of the form $0/0$ at the origin and, by l'Hôpital's rule:

$$\lim_{r\to 0}\frac{u_r}{r} = \lim_{r\to 0}\frac{\partial_r u_r}{\partial_r r} = u_{rr}(0,t)$$

so that the PDE becomes $u_t = 2u_{rr}$ at $r = 0$. The standard FTCS approximation of this equation is

$$U^{n+1}_0 = U^n_0 + 2\frac{k}{h^2}\delta_r^2 U^n_0 = U^n_0 + 2\frac{k}{h^2}\big(U^n_{-1} - 2U^n_0 + U^n_1\big)$$

but, because of symmetry, $U_{-1}^n = U_1^n$ and so

$$U_0^{n+1} = U_0^n + 4\frac{k}{h^2}(U_1^n - U_0^n) = (1 - 4\frac{k}{h^2})U_0^n + 4\frac{k}{h^2}U_1^n$$

which is of positive type for $k \leq h^2/4$.

11.33
We use the expressions derived for $\mathcal{L}_h U_{\ell,m}$ derived in the solution to Exercise 10.17—the notation, and numbering of nodes, is taken from that solution. Then

$$\begin{aligned}
\mathrm{P}_1 &: U_1^{n+1} = U_1^n - k\big(\ \ 65U_1^n - 16U_2^n - 16U_4^n - 112\big),\\
\mathrm{P}_2 &: U_2^{n+1} = U_2^n - k\big(-16U_1^n + 65U_2^n - 16U_3^n - 16U_5^n\big),\\
\mathrm{P}_3 &: U_3^{n+1} = U_3^n - k\big(\ \ 545U_3^n - \frac{144}{5}U_2^n\big),\\
\mathrm{P}_4 &: U_4^{n+1} = U_4^n - k\big(-16U_1^n + 65U_4^n - 16U_5^n - 16U_6^n - 160\big),\\
\mathrm{P}_5 &: U_5^{n+1} = U_5^n - k\big(\ \ 273U_5^n - \frac{128}{5}U_2^n - \frac{288}{11}U_4^n\big),\\
\mathrm{P}_6 &: U_6^{n+1} = U_6^n - k\big(\ \ \frac{979}{3}U_6^n - \frac{256}{11}U_4^n - 9\big).
\end{aligned}$$

These equations will be of positive type if the coefficient of U_j^n is non-negative at the point P_j. The most restrictive condition occurs at P_3 where it is required that $k \leq 1/545$. The corresponding limit for a regular grid is $k \leq 1/64$ (as, for instance, at P_j, $j = 1, 2, 4$).

Exercises 12 Finite difference methods for hyperbolic PDEs

12.1
Differentiating the PDE $u_t + au_x = 0$ with respect to t gives:

$$u_{tt} + au_{tx} = 0 \quad \Rightarrow \quad u_{tt} + a\partial_x u_t = 0 \quad \Rightarrow \quad u_{tt} + a\partial_x(-au_x) = 0$$

and so $u_{tt} = a^2 u_{xx}$. Using this in (12.21) leads to

$$\mathcal{R}_h\big|_m^n = \tfrac{1}{2}ah(c-1)u_{xx}\big|_m^n + \mathcal{O}(h^2) + \mathcal{O}(k^2),$$

as required.

12.3
The product of the coefficients of $U_{m\pm1}^n$ have opposite sign so it is not possible for both to be non-negative.
The amplification factor of the scheme is

$$\xi = 1 - \mathrm{i}c\sin\kappa h \quad \Rightarrow \quad |\xi|^2 = 1 + c^2\sin^2\kappa h$$

thus $\xi(\frac{1}{2}\pi) = 1 + c^2$ and it is not possible to find a constant C, independent of h and k such that $|\xi| \le 1 + Ck$ (see Definition 11.20).

12.5

(a) The Lax–Friedrichs scheme can be written as[2]

$$U_m^n = \tfrac{1}{2}(1+c)U_{m-1}^n + U_m^n + \tfrac{1}{2}(1-c)U_{m-1}^n$$

so its stencil is the same as that of the Lax-Wendroff method (see Fig. 12.6 (left)).

(b) The local truncation error is (note the scaling), using the results from Table 6.1,

$$\begin{aligned}\mathcal{R}_h\big|_m^n :&= k^{-1}\big(u_m^{n+1} - u_m^n + c\triangle_x u_m^n - \tfrac{1}{2}\delta_x^2 u_m^n\big)\\ &= \big(u_t + \tfrac{1}{2}ku_{tt} + \mathcal{O}(k^2)\big) + a\big(u_x + \tfrac{1}{6}h^2u_{xxx} + \mathcal{O}(h^2)\big) - \tfrac{1}{2}(h^2/k)u_{xx} + \mathcal{O}(h^4)\\ &= (u_t + au_x) + \tfrac{1}{2}ku_{tt} - \tfrac{1}{2}(h^2/k)u_{xx} + \mathcal{O}(k^2) + \mathcal{O}(h^2) + \mathcal{O}(h^4/k)\end{aligned}$$

so that $\mathcal{R}_h = \mathcal{O}(h) + \mathcal{O}(k)$ if c is fixed—the method is consistent of first order.

(c) We see from part (a) that the coefficients on the right hand side are non-negative for $-1 \le c \le 1$ in whch case the scheme is of positive type.

(d) The amplification factor of the scheme is, writing $\theta = \kappa h$,

$$\xi = 1 - \mathrm{i}c\sin\theta - 2\sin^2\tfrac{1}{2}\theta = \cos\theta - \mathrm{i}c\sin\theta.$$

Therefore,

$$|\xi|^2 - 1 = \cos^2\theta - 1 + c^2\sin^2\theta = -(1-c^2)\sin^2\theta$$

and $|\xi|^2 \le 1$ for all $\theta \in [-\pi, \pi]$ if, and only if, $-1 \le c \le 1$.

[2]Note that this can also be deduced from (12.26) by replacing U_m^n by the average $\frac{1}{2}(U_{m-1}^n + U_{m+1}^n)$.

(e) The Lax-Friedrichs method has the same order of consistency as the FTBS and FTFS methods so holds no advantage in that department. One major advantage that is does hold is that it is stable for $|c| \leq 1$ regardless of the sign of the advection speed so that it can be applied to systems of hyperbolic equations that have both right and left moving characteristics.

(f) The end product of the calculation in part (b) may be reorganized to give

$$\mathcal{R}_h\big|_m^n = \big(u_t + au_x - \tfrac{1}{2}\frac{h^2}{k}(1+c^2)u_{xx}\big) + \mathcal{O}(k^2) + \mathcal{O}(h^2) + \mathcal{O}(h^4/k)$$

where we have used $u_{tt} = a^2 u_{xx}$. Thus the local truncation error is consistent of order $\mathcal{O}(k^2) + \mathcal{O}(h^2)$ (provided that $c = ak/h$ is fixed) with the PDE $u_t + au_x = \varepsilon u_{xx}$, where $\varepsilon = \frac{1}{2}h^2(1+c^2)/k$.

The scheme is consistent with an advection-diffusion equation which is of parabolic type. The numerical solutions will therefore become smoother (damped) as time proceeds.

12.7

We substitute $U_m^n = \xi^n \mathrm{e}^{\mathrm{i}\kappa mh}$ into the Lax-Wendroff method scheme

$$U_m^{n+1} = [1 - c\triangle_x + \tfrac{1}{2}c^2\delta_x^2]\, U_m^n$$

and use the relationships $U_m^{n+1} = \xi U_m^n$, (11.47b) and (11.47c) to give

$$\xi U_m^n = U_m^n - \mathrm{i}c\sin(\kappa h)U_m^n - 2c^2\sin^2(\tfrac{1}{2}\kappa h)U_m^n \quad \Rightarrow \quad \xi = 1 - \mathrm{i}c\sin\kappa h - 2c^2\sin^2\tfrac{1}{2}\kappa h.$$

Then, writing $s = \sin^2\frac{1}{2}\kappa h$

$$\begin{aligned}|\xi|^2 - 1 &= (1-2c^2 s)^2 - 1 + c^2\sin^2(\kappa h) = -4c^2 s + 4c^4 s^2 + 4c^2 s(1-s)\\ &= -4c^2 s^2(1-c^2).\end{aligned}$$

Clearly $|\xi|^2 - 1 \leq 0$ if $c^2 \leq 1$.

12.9

With $\mu = 2, \nu = 0$ the formula (12.17) leads to the coefficients

$$\alpha_{-2} = \prod_{\substack{\ell=-2\\ \ell\neq -2}}^{0} \frac{\ell + c}{\ell + 2} = -\tfrac{1}{2}c(1-c), \quad \alpha_{-1} = \prod_{\substack{\ell=-2\\ \ell\neq -1}}^{0} \frac{\ell + c}{\ell + 1} = c(2-c), \quad \alpha_0 = \prod_{\substack{\ell=-2\\ \ell\neq 0}}^{0} \frac{\ell + c}{\ell} = \tfrac{1}{2}(1-c)(2-c)$$

$U_m^{n+1} = \alpha_{-2}U_{m-2}^n + \alpha_{-1}U_{m-1}^n + \alpha_0 U_m^n$ gives the Warming–Beam scheme (12.64).

12.11

(a) The local truncation error of Leith's method is

$$\begin{aligned}\mathcal{R}_h\big|_m^n &:= k^{-1}(u_m^{n+1} - u_m^n + c\triangle_x U_m^n - (r + \tfrac{1}{2}c^2)\delta_x^2 u_m^n),\\ &= u_t + \tfrac{1}{2}ku_{tt} + \mathcal{O}(k^2) + a\big(u_x + \tfrac{1}{6}h^2 u_{xxx} + \mathcal{O}(h^2)\big) - (\varepsilon + \tfrac{1}{2}a^2 k)\big(u_{xx} + \mathcal{O}(h^2)\big)\\ &= \big(u_t + au_x - \varepsilon u_{xx}\big) + \tfrac{1}{2}k(u_{tt} - a^2 u_{xx}) + \mathcal{O}(k^2) + \mathcal{O}(h^2)\end{aligned}$$

and so is consistent of order $\mathcal{O}(k) + \mathcal{O}(h^2)$ with the advection-diffusion equation $u_t + au_x = \varepsilon u_{xx}$.

(b) Differentiating the PDE with respect to t and x gives

$$\begin{aligned} u_{tt} + a u_{xt} &= \varepsilon u_{xxt} \\ u_{xt} + a u_{xx} &= \varepsilon u_{xxx} \end{aligned}$$

so, subtracting $a \times$ second from the first equation leads to $u_{tt} = a^2 u_{xx} + \varepsilon(u_{xxt} - a u_{xxx})$ and, therefore,

$$\mathcal{R}_h\big|_m^n = \tfrac{1}{2} k\varepsilon(u_{xxt} - a u_{xxx}) + \mathcal{O}(k^2) + \mathcal{O}(h^2).$$

From the point of view of convergence as $h, k \to 0$ the scheme is clearly consistent of order $\mathcal{O}(k) + \mathcal{O}(h^2)$ with the advection-diffusion equation. However, the factor ε multiplying the leading term in the local truncation error means that in computations where $\varepsilon \ll 1$, the method will, on coarser grids, effectively perform as a second order scheme.

(c) We observe that the amplification factor for Leith's scheme

$$\xi = 1 - 4(r + \tfrac{1}{2}c^2)\sin^2(\tfrac{1}{2}\kappa h) + \mathrm{i}\,\rho \sin(\kappa h)$$

is the same as the FTCS scheme (11.53) with r replaced by $(r + \frac{1}{2}c^2)$ and ρ replaced by $-c = -ak/h$. The stability conditions (11.56) become

$$\tfrac{1}{2}c^2 \le r + \tfrac{1}{2}c^2 \le \tfrac{1}{2},$$

so the left inequality is always satisfied leaving $2r + c^2 \le 1$. Written in terms of h and k, this requires

$$2\varepsilon k + a^2 k^2 \le h^2.$$

If both sides are multiplied by a^2/ε^2 we find

$$2\frac{ka^2}{\varepsilon} + \big(\frac{ka^2}{\varepsilon}\big)^2 \le \big(\frac{ah}{\varepsilon}\big)^2 \quad \Rightarrow \quad 2\widehat{k} + \widehat{k}^2 \le \widehat{h}^2,$$

where the stability region is independent of any parameters when expressed in terms of $\widehat{k} := ka^2/\varepsilon$ and $\widehat{h} := ha/\varepsilon$ (known as non-dimensional variables because their values are independent of the units used to measure k, h, a and ε). Thus, in the $\widehat{h}$-$\widehat{k}$ plane the boundary is a branch of the hyperbola

$$(1 + \widehat{k})^2 - \widehat{h}^2 = 1$$

whose centre is at $\widehat{k} = -1$, $\widehat{h} = 0$ and has asymptotes $\widehat{k} = -1 \pm \widehat{h}$. The region bounded by this curve and the $\widehat{h}$-axis is shown shaded in Fig. 13 but the scales shown on the axes are for the grid sizes h and k. Also shown for reference is the curve is $k = \varepsilon h^2/2$ (dashed), which is the corresponding stability limit for the FTCS approximation of the heat equation $u_t = \varepsilon u_{xx}$. We can see that this is the relevant limit when ε/a is large (when we should focus on the stability region near the origin where h and k are small). On the other hand, when ε/a is small—the advection dominated case—we enjoy a limit which is approximately that for the FTCS approximation of the advection equation $u_t + a u_x = 0$, i.e., $k \le ah$.

12.13

(a) We evaluate the coefficients supplied in (12.33a) at the Courant numbers $c = -1, 0, 1, 2$ and display the results in the following table.

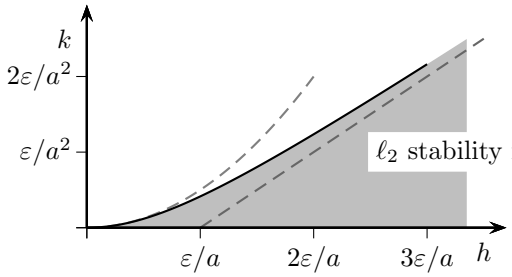


Figure 13: The ℓ_2-stability region for Leith's method. The dashed curve shows $k = \varepsilon h^2/2$ and the dashed line the asymptote $k = (ah - \varepsilon)/a^2$.

Coeffs.	α_{-2}	α_{-1}	α_0	α_1
c	$-\frac{1}{6}c(1-c^2)$	$\frac{1}{2}c(1+c)(2-c)$	$\frac{1}{2}(1-c^2)(2-c)$	$\frac{1}{6}c(c-1)(2-c)$
2	1	0	0	0
1	0	1	0	0
0	0	0	1	0
-1	0	0	0	1

and so $U_m^{n+1} = U_{m-c}^n$ for $c = -1, 0, 1, 2$.

(b) Using (12.33a), the LTE is given by

$$\mathcal{R}_h\big|_m^n := k^{-1}\big(u_m^{n+1} - u_m^n + c\triangle_x u_m^n - \tfrac{1}{2}c^2\delta_x^2 u_m^n - \tfrac{1}{6}c(1-c^2)\triangle_x^-\delta_x^2 u_m^n\big)$$

The expansion of all terms except the last on the right hand side are available in Table 6.1. For the final term we proceed as follows (many other routes to the same destination are possible): Since $\triangle_x^- v_m^n = hv_x - \frac{1}{2}h^2 v_{xx} + \mathcal{O}(h^3)$ (Table 6.1) then, with $v = \delta_x^2 u_m^n = h^2 u_{xx} + \mathcal{O}(h^4)$,

$$\triangle_x^-\delta_x^2 u_m^n = h\partial_x\big(h^2 u_{xx} + \mathcal{O}(h^4)\big) + \mathcal{O}(h^4) = h^3 u_{xxx} - h^4 u_{xxxx} + \mathcal{O}(h^5).$$

Now

$$\begin{aligned}\mathcal{R}_h\big|_m^n = \quad & u_t + \tfrac{1}{2}ku_{tt} + \tfrac{1}{6}k^2 u_{ttt} + \tfrac{1}{24}k^3 u_{tttt} + \mathcal{O}(k^4)\\ & + a\big(u_x \quad + \quad \tfrac{1}{6}h^2 u_{xxx} + \mathcal{O}(h^4)\big)\\ & - \tfrac{1}{2}ka^2\big(u_{xx} \qquad + \qquad \tfrac{1}{12}h^2 u_{xxxx} + \mathcal{O}(h^4)\big)\\ & - \tfrac{1}{6}a(1-c^2)\big(h^2 u_{xxx} - h^3 u_{xxxx} + \mathcal{O}(h^4)\big)\end{aligned}$$

so that

$$\begin{aligned}\mathcal{R}_h\big|_m^n = \quad & u_t + au_x + \tfrac{1}{2}k\big(u_{tt} - a^2 u_{xx}\big) + \tfrac{1}{6}k^2\big(u_{ttt} + a^2 u_{xxx}\big)\\ & + \tfrac{1}{24}k^3\big(u_{tttt} + a^2 u_{xxxx}\big) + \tfrac{1}{24}a^4h^3(1+c)(1-c)(2-c)u_{xxxx} + \mathcal{O}(h^4).\end{aligned}$$

But $\partial_t^j u = (-a\partial_x)^j u$ so that this reduces to

$$\mathcal{R}_h\big|_m^n = \tfrac{1}{24}a^4h^3(1+c)(1-c)(2-c)u_{xxxx} + \mathcal{O}(h^4).$$

Notice how the leading coefficient vanishes at $c = -1, 1, 2$ (as do all subsequent coefficients in deference to part (a)).

(c) The CFL condition follows immediately from Example 12.3.

(d) We substitute $U_m^n = \xi^n \mathrm{e}^{\mathrm{i}\kappa m h}$ into the finite difference scheme and use the relationships $U_m^{n+1} = \xi U_m^n$, $\triangle_x^- U_m^n = (1 - \mathrm{e}^{-\mathrm{i}\kappa h}$, (11.47b) and (11.47c) to give

$$\xi = 1 - \mathrm{i}c\sin\theta - 2c^2\sin^2\tfrac{1}{2}\theta - \tfrac{2}{3}c(1-c^2)(1-\mathrm{e}^{-\mathrm{i}\theta})\sin^2\tfrac{1}{2}\theta,$$

where $\theta = \kappa h$. We now write $s = \sin^2\frac{1}{2}\theta$ and then $1-\mathrm{e}^{-\mathrm{i}\theta} = (1-\cos\theta)+\mathrm{i}\sin\theta = 2s+\mathrm{i}\sin\theta$ so that, collecting real and imaginary parts

$$\xi = 1 - 2c^2 s - \tfrac{4}{3}c(1-c^2)s^2 - \mathrm{i}c\big(1 + \tfrac{2}{3}c(1-c^2)s\big)\sin\theta.$$

Therefore, since $\sin^2\theta = 4s(1-s)$ and defining $A = \frac{2}{3}c(1-c^2)s$ in order to lighten the complexity,

$$\begin{aligned}
\xi &= 1 - 2c^2 s - 2As - \mathrm{i}(A+c)\sin\theta.\\
|\xi|^2 - 1 &= \big(1 - 2s(A+c^2)\big)^2 - 1 + 4(A+c)^2 s(1-s)\\
&= -4s(A+c^2) + 4s^2(A+c^2)^2 + 4(A+c)^2 s - 4(A+c)^2 s^2\\
&= -4s\big((A+c^2) - (A+c)^2\big) + 4s^2\big((A+c^2)^2 - (A+c)^2\big)\\
&= -4sA(1-2c-A) + 4s^2(c^2-c)(c^2+c+2A)\\
&= -4sA(1-2c-A)\underbrace{-4s^2c^2(1-c^2)}_{-6scA} - 8s^2Ac(1-c)\\
&= -4sA(1-\tfrac{1}{2}c\underbrace{-A+2sc(1-c))}_{\frac{2}{3}sc(1-c)(c-2)} = -\tfrac{8}{3}c(1-c^2)(2-c)s^2\big[1+\tfrac{2}{3}sc(1-c)\big].
\end{aligned}$$

(e) For ℓ_2-stability it is necessary to have $|\xi|^2 - 1 \le 0$. Since $F(s) := (|\xi|^2-1)/s^2$ is linear in s and, in order for it to be non-positive for $0 \le s \le 1$, it must be non-positive at $s = 0$ and $s = 1$. Now

$$\begin{aligned}
F(0) &= -\tfrac{8}{3}c(1-c^2)(2-c) \le 0 && \Rightarrow c \in [-1, -\tfrac{1}{2}] \cup [0,1] \cup [\tfrac{3}{2}, 2]\\
F(1) &= -\tfrac{8}{9}c(1-c^2)(2-c)(1+2c)(3-2c) && \Rightarrow c \in (-\infty, -1] \cup [0,1] \cup [2, \infty)
\end{aligned}$$

and the only interval held in common is $c \in [0,1]$.

12.15

The mth component of $C\boldsymbol{v}_j$ is

$$\begin{aligned}
\big(C\boldsymbol{v}_j\big)_m &= -c\big(\boldsymbol{v}_j\big)_{m-1} + (1+c)\big(\boldsymbol{v}_j\big)_m = -c\mathrm{e}^{2\pi\mathrm{i}(m-1)j/M} + (1+c)\mathrm{e}^{2\pi\mathrm{i}m\,j/M}\\
&= \big(1 + c - c\mathrm{e}^{-2\pi\mathrm{i}\,j/M}\big(\boldsymbol{v}_j\big)_m.
\end{aligned}$$

this holds for all $m = 1, 2, \ldots, M$ (at $M = 1$ we need to recognise that $\big(\boldsymbol{v}_j\big)_M = \big(\boldsymbol{v}_j\big)_{M-1}$ because of the periodic nature of its components). Thus $C\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$, where $\lambda_j = 1 + c - c\mathrm{e}^{-2\pi\mathrm{i}\,j/M}$. Comparing this with the amplification factor $\xi(\kappa h)$ given by (12.36), it is seen that

$$\lambda_j = 1/\xi(2\pi j h)$$

corresponding to the wavenumber $\kappa = 2\pi j$. This result generalises readily to all *constant coefficient* finite difference approximations of parabolic and hyperbolic PDEs with periodic BCs because all circulant matrices of a given dimension share the same set of eigenvectors.

12.17

From (12.22) and Exercise 12.16 we find, respectively,

$$\xi_{\mathrm{F}}(\theta) = 1 - c + c\mathrm{e}^{-\mathrm{i}\theta} \text{ and } \xi_{\mathrm{B}}(\theta) = \frac{1}{1 - c + c\mathrm{e}^{\mathrm{i}\theta}} = \frac{1}{\xi_{\mathrm{F}}^*(\theta)}$$

(where $\xi_{\mathrm{F}}^*(\theta)$ is the complex conjugate of $\xi_{\mathrm{F}}(\theta)$) so that $\xi_{\mathrm{F}}^*(\theta)\xi_{\mathrm{B}}(\theta) = 1$.
Since $|\xi_{\mathrm{F}}(\theta)|\,|\xi_{\mathrm{B}}(\theta)| = 1$ it follows that if $|\xi_{\mathrm{F}}(\theta)| < 1$ (the FTBS scheme is stable) for some scaled wavenumber θ, then $|\xi_{\mathrm{B}}(\theta)| > 1$ (the BTFS scheme is unstable) at that wavenumber, and *vice versa*. Their stability regions are therefore complements of each other. Since the FTBS scheme is stable for $0 \le c \le 1$ we deduce that the BTFS scheme is stable for $c \ge 1$ and $c \le 0$.

12.19
With $\eta = (1 - c) + (1 + c)\mathrm{e}^{-\mathrm{i}\kappa h}$, then

$$\eta^* = (1 - c) + (1 + c)\mathrm{e}^{\mathrm{i}\kappa h} = \big((1 - c)\mathrm{e}^{-\mathrm{i}\kappa h} + (1 + c)\big)\mathrm{e}^{\mathrm{i}\kappa h}$$

so that (12.39) becomes $\xi = \mathrm{e}^{-\mathrm{i}\kappa h}\eta/\eta^*$. It follows that $|\xi| = 1$ (for all c) since $|\eta| = |\eta^*|$ and $|\mathrm{e}^{-\mathrm{i}\kappa h}| = 1$.

12.21
The amplification factor of the method (12.5) is

$$\xi = \sum_{j=-\mu}^{\nu} \alpha_j \mathrm{e}^{\mathrm{i}j\kappa h}$$

and its LTE is

$$\mathcal{R}_h\big|_m^n := k^{-1}\big(u_m^{n+1} - \sum_{j=-\mu}^{\nu} \alpha_j u_{m+j}^n\big).$$

The solution of the advection equalion $u_t + au_x = 0$ with initial condition $u(x, 0) = \exp(\mathrm{i}\kappa x)$ is $u(x, t) = \exp(\mathrm{i}\kappa(x - at)$ so the LTE becomes

$$\begin{aligned}\mathcal{R}_h\big|_m^n &= k^{-1}\big(\mathrm{e}^{\mathrm{i}\kappa(x_m - a(n+1)k)} - \sum_{j=-\mu}^{\nu} \alpha_j \mathrm{e}^{\mathrm{i}\kappa(x_{m+j} - ank)}\big) \\ &= k^{-1}\Big(\mathrm{e}^{-a\mathrm{i}\kappa k} - \sum_{j=-\mu}^{\nu} \alpha_j \mathrm{e}^{\mathrm{i}j\kappa h}\Big)\mathrm{e}^{\mathrm{i}\kappa(x_m - ank)} = k^{-1}\Big(\mathrm{e}^{-\mathrm{c}\mathrm{i}\theta} - \xi(\theta)\Big)u_m^n,\end{aligned}$$

where $\theta = \kappa h$. This shows that equation (12.50) holds more generally than just for the Lax-Wendroff method.
Now $h^{p+1}\partial_x^{p+1}u = (\mathrm{i}\kappa h)^{p+1}u = (\mathrm{i}\theta)^{p+1}u$ with a similar expression for $h^{p+2}\partial_x^{p+2}u$ and the given error expansion gives

$$k\mathcal{R}_h\big|_m^n = C_{p+1}(\mathrm{i}\theta)^{p+1}u + C_{p+2}(\mathrm{i}\theta)^{p+2}u + \mathcal{O}(h^{p+3})$$

and, using (12.50), we obtain

$$\xi(\theta) = \mathrm{e}^{-\mathrm{i}c\theta} - C_{p+1}(\mathrm{i}\theta)^{p+1} - C_{p+2}(\mathrm{i}\theta)^{p+2}u + \mathcal{O}(h^{p+3})$$

showing that the amplification factor is an order $(p + 1)$ approximation of $\mathrm{e}^{-\mathrm{i}c\theta}$.

(a) p is odd: then $(\mathrm{i}\theta)^{p+1}$ is real and $(\mathrm{i}\theta)^{p+2}$ is imaginary. In fact, supposing that $p = 2q-1$ (where q is a positive integer),

$$(\mathrm{i}\theta)^{p+1} = \mathrm{i}^{2q}\theta^{p+1} = (-1)^q\theta^{p+1} = (-1)^{(p+1)/2}\theta^{p+1}$$

and $(\mathrm{i}\theta)^{p+2} = (-1)^{(p+1)/2}\mathrm{i}\theta^{p+2}$. Hence

$$\xi(\theta) = \big(\cos c\theta - (-1)^{(p+1)/2}C_{p+1}\theta^{p+1}\big) - \mathrm{i}\big(\sin c\theta + (-1)^{(p+1)/2}\mathrm{i}\theta^{p+2}\big) + \mathcal{O}(\theta^{p+3})$$

leading to

$$\begin{aligned}|\xi(\theta)|^2 &= \big(\cos c\theta - (-1)^{(p+1)/2}C_{p+1}\theta^{p+1}\big)^2 + \big(\sin c\theta + (-1)^{(p+1)/2}C_{p+2}\theta^{p+2}\big)^2 + \mathcal{O}(\theta^{p+3})\\ &= \cos^2 c\theta + \sin^2 c\theta - 2(-1)^{(p+1)/2}C_{p+1}\theta^{p+1}\cos c\theta + \mathcal{O}(\theta^{p+3})\\ &= 1 - 2(-1)^{(p+1)/2}C_{p+1}\theta^{p+1} + \mathcal{O}(\theta^{p+3}),\end{aligned}$$

where we have used $\theta^{p+1}\cos c\theta = \theta^{p+1} + \mathcal{O}(\theta^{p+3})$ and $\theta^{p+2}\sin c\theta = \mathcal{O}(\theta^{p+3})$.

(b) p is even: then $(\mathrm{i}\theta)^{p+1}$ is imaginary and $(\mathrm{i}\theta)^{p+2}$ is real. In fact, supposing that $p = 2q$ (where q is a positive integer),

$$(\mathrm{i}\theta)^{p+1} = \mathrm{i}(\mathrm{i}^{2q})\theta^{p+1} = (-1)^q\mathrm{i}\theta^{p+1} = (-1)^{p/2}\mathrm{i}\theta^{p+1}$$

and $(\mathrm{i}\theta)^{p+2} = (-1)^{p/2+1}\theta^{p+2}$. Hence

$$\xi(\theta) = \big(\cos c\theta - (-1)^{p/2+1}C_{p+2}\theta^{p+2}\big) - \mathrm{i}\big(\sin c\theta + (-1)^{p/2}C_{p+1}\theta^{p+1}\big) + \mathcal{O}(\theta^{p+3})$$

leading to

$$\begin{aligned}|\xi(\theta)|^2 &= \big(\cos c\theta - (-1)^{p/2+1}C_{p+2}\theta^{p+2}\big)^2 + \big(\sin c\theta + (-1)^{p/2}C_{p+1}\theta^{p+1}\big)^2 + \mathcal{O}(\theta^{p+3})\\ &= \cos^2 c\theta + \sin^2 c\theta - 2(-1)^{p/2+1}C_{p+2}\theta^{p+2}\cos c\theta + 2(-1)^{p/2}C_{p+1}\theta^{p+1}\sin c\theta + \mathcal{O}(\theta^{p+3})\\ &= 1 - 2(-1)^{p/2+1}C_{p+2}\theta^{p+2} + 2(-1)^{p/2}C_{p+1}\theta^{p+2} + \mathcal{O}(\theta^{p+3})\\ &= 1 - 2(-1)^{p/2+1}(cC_{p+1} + C_{p+2})\theta^{p+2} + \mathcal{O}(\theta^{p+3}),\end{aligned}$$

where we have used $\theta^{p+2}\cos c\theta = \theta^{p+2} + \mathcal{O}(\theta^{p+4})$ and $\theta^{p+1}\sin c\theta = c\theta^{p+2} + \mathcal{O}(\theta^{p+4})$.

The bottom line is that for a method whose order of consistency p (odd), the order of dissipation is $(p+1)$ but, if the order of consistency p (even), the order of dissipation is $(p+2)$.

12.23

When $c > 0$, characteristics travel from left-to-right (see Fig. 14, left) and the domain of dependence of the anchor point (shown as $\circ$) is shaded. The method must be used as

$$U_m^{n+1} = (U_m^n + c\,U_{m-1}^{n+1})/(1+c), \quad m = 1, 2, \ldots, M$$

so the BC must be placed at $x = 0$.

The corresponding situation when $c \le -1$ is shown in Fig. 14 (right) and the BC needs to be at $x = 1$.

12.25

The Lax-Wendroff method is applied at the points (x_m, nk), $0 < m < M$, $n \ge 0$. At these points

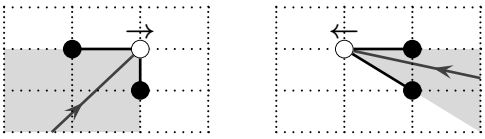


Figure 14: The two modes of operation of the BTBS scheme in Exercise 12.23 with target points (∘). Both modes are stable provided that the characteristics lie in the shaded regions.

the LTE $\mathcal{R}_m^n = 0$ because the exact solution is a quadratic polynomial and the expression for the LTE involves the 3rd derivatives of u (see (12.30)). So

$$\begin{aligned}\mathcal{R}_m^n &= \frac{1}{k}\big(u_m^{n+1} - u_m^n + c\triangle_x u_m^n - \tfrac{1}{2}c^2\delta_x^2 u_m^n\big) = 0,\\ U_m^{n+1} &= \big(1 - c\triangle_x + \tfrac{1}{2}c^2\delta_x^2\big)\, U_m^n,\end{aligned}$$

and so the global error $E = u - U$ itself is a solution of the Lax-Wendroff scheme:

$$\begin{aligned}E_m^{n+1} &= \big(1 - c\triangle_x + \tfrac{1}{2}c^2\delta_x^2\big)\, E_m^n\\ &= \tfrac{1}{2}c(1+c)E_{m-1}^n + (1-c^2)E_m^n + \tfrac{1}{2}c(c-1)E_{m+1}^n\end{aligned}$$

with starting condition $E_m^0 = 0$ and a homogeneous BC $E_0^n = 0$ at $x = 0$. Thus, by a domain of dependence argument, $E_m^n = 0$ for $x_m + t_n \le 1$.
The LTE of the FTBS scheme (see (12.21)) at $x = 1$ is, for $u(x,t) = (x-t)^2$ and $a = 1$,

$$\mathcal{R}_M^n = k^{-1}\big(u_M^{n+1} - cu_{M-1}^n - (1-c)u_M^n\big) = -h(1-c)$$

and so

$$E_M^{n+1} = cE_{M-1}^n - (1-c)E_M^n - hk(1-c)$$

for $n = 0, 1, 2, \ldots$.
Assuming now that $E_m^n \to A_m$ as $n \to \infty$, the Lax-Wendroff equations reduce to

$$(1+c)A_{m-1} - 2cA_m + (c-1)A_{m+1} = 0,$$

Setting

$$C = \tfrac{1}{2}(1-c)^2h^2\left(-\frac{1+c}{1-c}\right)^{1-M}, \qquad \rho = \frac{c+1}{c-1}$$

then the proposed solution is $A_m = C\rho^m$. Substituting, we find,

$$(1+c)A_{m-1} - 2cA_m + (c-1)A_{m+1} = C\rho^m\big(\frac{1+c}{\rho} - 2c + (c-1)\rho\big) = C\rho^m((c-1) - 2c + (c+1)) = 0$$

so these equations are satisfied.
When $E_m^n \to A_m$ as $n \to \infty$, the FTBS equations reduce to $A_M - A_{M-1} = -h^2(1-c)$. Substituting the putative solution into the left hand side of this, we find

$$A_M - A_{M-1} = C\rho^{M-1}(\rho - 1) = \frac{2}{c-1}C\rho^{M-1} = -h^2(1-c)$$

as required.

12.27
The FTCS method is stable only for positive eigenvalues and the FTBS method is stable only for negative eigenvalues but the matix A has eigenvalues $\pm a$ of both signs.

(a) Applying the FTBS method to (12.66a) and the FTBS method to (12.66b) leads to

$$(U_m^{n+1} - V_m^{n+1}) = (U_m^n - V_m^n) - c\triangle_x^-(U_m^n - V_m^n) = c(U_{m-1}^n - V_{m-1}^n) + (1-c)(U_m^n - V_m^n)$$
$$(U_m^{n+1} + V_m^{n+1}) = (U_m^n + V_m^n) + c\triangle_x^+(U_m^n + V_m^n) = (1-c)(U_m^n + V_m^n) + c(U_{m+1}^n + V_{m+1}^n)$$

so, by adding and subtracting, we find

$$\begin{aligned}
U_m^{n+1} &= (1-c)U_m^n + \tfrac{1}{2}c(U_{m-1}^n - V_{m-1}^n) + \tfrac{1}{2}c(U_{m+1}^n + V_{m+1}^n) \\
&= (1 + \tfrac{1}{2}c\delta_x^2)U_m^n + c\triangle_x V_m^n \\
V_m^{n+1} &= (1-c)V_m^n - c(U_{m-1}^n - V_{m-1}^n) + c(U_{m+1}^n + V_{m+1}^n) \\
&= (1 + \tfrac{1}{2}c\delta_x^2)V_m^n + c\triangle_x U_m^n \\
\boldsymbol{U}_m^{n+1} &= (1 + \tfrac{1}{2}c\delta_x^2)\boldsymbol{U}_m^n - (k/h)A\triangle_x \boldsymbol{U}_m^n = \begin{bmatrix} (1 + \frac{1}{2}c\delta_x^2) & c\triangle_x \\ c\triangle_x & (1 + \frac{1}{2}c\delta_x^2) \end{bmatrix} \boldsymbol{U}_m^n.
\end{aligned}$$

(b) With $A = \begin{bmatrix} 0 & -a \\ -a & 0 \end{bmatrix}$ the matrix of eigenvectors is $V = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $\Lambda = \mathrm{diag}(-a, a)$. Therefore $|\Lambda| = |a|I$ (where I is the 2×2 identity matrix) and so $|A| = V|\Lambda|V^{-1} = |a|VV^{-1} = |a|I$. The method now reads

$$\boldsymbol{U}_m^{n+1} = \boldsymbol{U}_m^n - (k/h)A\triangle_x \boldsymbol{U}_m^1 + \tfrac{1}{2}|c|\,\delta_x^2 \boldsymbol{U}_m^n$$

which is identical to the method in part (a) provided $a > 0$.

(c) Following the steps in Example 12.18 the Lax–Friedrichs scheme is found to be

$$\boldsymbol{U}_m^{n+1} = [I - (k/h)A\triangle_x + \tfrac{1}{2}\delta_x^2]\boldsymbol{U}_m^n,$$

which is identical to the method in part (b) when $|c| = 1$.

12.29

Consider the critical factor in the final term on the right hand side of (12.77)

$$\triangle_x^-\big(\varphi_m^n \triangle_x^+ U_m^n\big) = \varphi_m^n \triangle_x^+ U_m^n - \varphi_{m-1}^n \triangle_x^+ U_{m-1}^n$$

but $\triangle_x^+ U_{m-1}^n = \triangle_x^- U_m^n$ so that this becomes

$$\begin{aligned}
\triangle_x^-\big(\varphi_m^n \triangle_x^+ U_m^n\big) &= \varphi_m^n \triangle_x^+ U_m^n - \varphi_{m-1}^n \triangle_x^- U_m^n \\
&= \big(\frac{\varphi_m^n}{\rho_m^n} - \varphi_{m-1}^n\big)\triangle_x^- U_m^n,
\end{aligned}$$

where $\rho_m^n = \triangle_x^- U_m^n / \triangle_x^+ U_m^n$, and the result follows.

12.31

The minmod limiter is a special case of the Chakravarthy–Osher limiter.
Chakravarthy–Osher:

$$\begin{aligned}
\rho \le 0: &\qquad \max\{0, \min(\rho, \psi)\} = \max\{0, \rho\} = 0 \\
0 < \rho \le \psi: &\qquad \max\{0, \min(\rho, \psi)\} = \max\{0, \rho\} = \rho \\
\rho > \psi: &\qquad \max\{0, \min(\rho, \psi)\} = \max\{0, \psi\} = \psi.
\end{aligned}$$

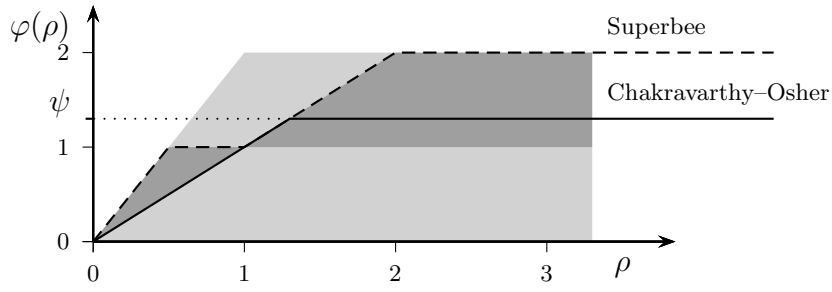


Figure 15: The flux limiters for Exercise 12.31.

Superbee:

$$
\begin{aligned}
\rho \le 0: &\qquad \max\{0, \min(2\rho, 1), \min(\rho, 2)\} = \max\{0, \rho)\} = 0 \\
0 < \rho \le \tfrac{1}{2}: &\qquad \max\{0, \min(2\rho, 1), \min(\rho, 2)\} = \max\{0, 2\rho, \rho\} = 2\rho \\
\tfrac{1}{2} < \rho \le 1: &\qquad \max\{0, \min(2\rho, 1), \min(\rho, 2)\} = \max\{0, 1, \rho\} = 1 \\
1 < \rho \le 2: &\qquad \max\{0, \min(2\rho, 1), \min(\rho, 2)\} = \max\{0, 1, \rho\} = \rho \\
\rho > 2: &\qquad \max\{0, \min(2\rho, 1), \min(\rho, 2)\} = \max\{0, 1, 2\} = 2.
\end{aligned}
$$

http://www.springer.com/978-3-319-22568-5

Essential Partial Differential Equations
Analytical and Computational Aspects
Griffiths, D.F.; Dold, J.W.; Silvester, D.J.
2015, XI, 368 p. 106 illus., 1 illus. in color., Softcover
ISBN: 978-3-319-22568-5