SIO203C: PDE Notes A

W.R. Young ¹

February 2011

¹Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093–0230, USA.
## Contents

1 First order PDE's  
   1.1 The simplest partial differential equation  ............... 4  
   1.2 Arbitrary functions  ........................................ 8  
   1.3 The Jacobian  ............................................. 11  
   1.4 The method of characteristics  ............................ 13  
   1.5 Some books  .............................................. 16  
   1.6 Problems  .................................................. 17  

2 Linear evolution equations  
   2.1 Conservation laws  ........................................... 21  
   2.2 A recipe for semilinear evolution equations  ............... 23  
   2.3 Entry on the half-line  ...................................... 26  
   2.4 Age-Stratified Populations  ................................. 29  
   2.5 Probability generating functions  .......................... 33  
   2.6 References specialized to pgfs  ............................ 38  
   2.7 Problems  .................................................. 38  

3 Quasilinear equations  
   3.1 Characteristics and their properties  ....................... 49  
   3.2 More theory  .............................................. 54  
   3.3 The nonlinear advection equation  .......................... 56  
   3.4 Shocks, caustics and multivalued solutions  ............... 60  
   3.5 Problems  .................................................. 63
4 δ-functions and Green's functions 69
  4.1 Review of “patching” ............................... 69
  4.2 δ-sequences ...................................... 73
  4.3 Green’s function of the harmonic oscillator ..... 75
  4.4 The black box .................................... 80
  4.5 Problems ........................................... 83

5 Green's functions for boundary value problems 88
  5.1 A suspended string ................................ 88
  5.2 Solvability conditions ............................... 90
  5.3 Problems ........................................... 93

6 The diffusion equation $u_t = \kappa \nabla^2 u$ 94
  6.1 Origin of the diffusion equation .................. 94
  6.2 Smoothing a discontinuity or “front” .......... 97
  6.3 Discussion of the similarity method ............. 98
  6.4 The Green's function for diffusion on the line .. 99
  6.5 Solution of the diffusive initial value problem 102
  6.6 Diffusion with a source .......................... 106
  6.7 Specialized references to diffusion ............ 108
  6.8 Problems ........................................... 109

7 Diffusion on a half-line and on a finite domain 114
  7.1 Half-line problems and images ........................ 114
  7.2 Heating the Earth .................................. 117
  7.3 Inhomogeneous boundary conditions ............. 120
  7.4 Diffusion of a heat around a wire loop .......... 122
  7.5 Problems ........................................... 126
Lecture 1

First order PDE’s

1.1 The simplest partial differential equation

Here is the simplest example of a first-order partial differential equation (PDE)

\[ h_x = 0, \tag{1.1} \]

where \( h(x, y) \) is a function of two variables. For example, \( h \) might be topographic height of a landscape above the \((x, y)\)-plane. It is fundamental that the solution of \((1.1)\) is a two-dimensional solution surface living in a three-dimensional space.

The solution of the PDE \((1.1)\) is that

\[ h(x, y) = a(y), \tag{1.2} \]

where \( a \) is an arbitrary (even discontinuous) function of \( y \). For example, \( a(y) \) might be

\[ a(y) = \begin{cases} 1, & \text{if } y \text{ is rational;} \\ 0, & \text{if } y \text{ is irrational.} \end{cases} \tag{1.3} \]

This solution of \((1.1)\) is discontinuous everywhere, but differentiable along the lines of constant \( y \), which is all that \((1.1)\) requires.

This reasoning might remind you that the general solution of an ODE involves arbitrary constants, which are determined by initial or boundary conditions. Apparently the solution of a PDE involves determining an arbitrary function from boundary or initial conditions.
Lecture 1. First order PDE’s

Figure 1.1: Left panel: the solid curve is the data curve \( y = \ln x \) (with \( x > 0 \)). The characteristics are the dashed lines \( y = \) constant. Right panel: the solid curve is the data curve \( y = \exp(x) \) (with \(-\infty < x < \infty\)). The characteristics are the dashed lines \( y = \) constant.

The lines of constant \( y \) are called characteristic curves of the PDE (1.1), and a solution like (1.2) that contains an arbitrary function of the characteristic variable is called a general solution of the PDE. It is remarkable that the solution on a particular characteristic curve is independent of the solution on other characteristics — even infinitesimally close characteristics.

Although the PDE (1.1) is trivial, it can be used to make several general points. So let’s wallow further in (1.1) and illustrate how \( a(y) \) might be determined by boundary information.

**Example:** Solve the PDE \( h_x = 0 \) subject to the boundary condition that on the curve \( y = \ln x \) (with \( x > 0 \)), \( h = \sin x \).

The left panel of figure 1.1 shows the data curve \( y = \ln x \) on which \( h(x, y) \) is specified as being equal to \( \sin x \). The dashed straight lines are the characteristics, on which \( h(x, y) \) is constant. Each characteristic intersects the data curve only once, and at this intersection the arbitrary function \( a(y) \) is determined. To determine the arbitrary function \( a(y) \), on the data curve we have \( \sin x = a(\ln x) \), or \( h(x, y) = \sin y \). Notice that the solution is defined everywhere in the \((x, y)\)-plane by the PDE and specification of \( h \) on the data curve.

**Example:** Solve the PDE \( h_x = 0 \) subject to the boundary condition that on the curve \( y = e^x \) (with \(-\infty < x < \infty\)), \( h = \sin x \).
In the right panel of Figure 1.1 we show the data curve $y = \exp(x)$ and the dashed straight lines which are the characteristics of the PDE $h_x = 0$. In this case, only the characteristics in the upper half plane intersect the data curve. Consequently the data determines the arbitrary function $a(y)$ only in the upper half plane. In the lower half plane the function $a(y)$ is still arbitrary: the characteristics in the lower half plane don’t find the data curve. Thus in this example we have incomplete determination of $h(x, y)$.

Focussing then on the upper half plane, at the intersection of the characteristics with the data curve we have $\sin x = a(\exp x)$, so in the region $y > 0$ we have $h(x, y) = \sin \ln y$. We say that the domain of definition of the solution is the upper half plane $y > 0$.

**Example:** Solve PDE $h_x = 0$ subject to the boundary on $y = 0$, $h = \sin x$.

In this case there is no solution: the PDE says that $h$ is constant on every line of constant $y$, and the boundary condition contradicts this on the particular line $y = 0$. If the boundary condition is changed to $h = \pi$ on $y = 0$ then there is no longer an inconsistency — in this case $h(x, 0) = \pi$. But $h(y)$ remains undetermined for $y \neq 0$.

**Example:** Solve the PDE $h_x = 0$, subject to the boundary condition that $h = x$ on the circle $x^2 + y^2 = 1$.

You should draw the circle and the lines of constant $y$. In the regions where $|y| > 1$ the lines of constant $y$ don’t intersect the circle. Thus the boundary condition supplies no information in these regions. All we can say is that $h(x, y) = a(y)$, with $a(y)$ still arbitrary if $|y| > 1$.

Inside the circle $x^2 + y^2 = 1$ the PDE has no solutions, because $h(x, y) = a(y)$ can’t satisfy the boundary condition at both

$$x = +\sqrt{1 - y^2} \quad \text{and} \quad x = -\sqrt{1 - y^2}. \quad (1.4)$$

We might say that in the region $x^2 + y^2 < 1$, the boundary information is inconsistent with the PDE and there is no solution. An alternative point of view is that inside the circle the solution is double valued. That is at each point $(x, y)$ one has a solution

$$h_+(x, y) \overset{\text{def}}{=} +\sqrt{1 - y^2}, \quad (1.5)$$

and a second solution

$$h_-(x, y) \overset{\text{def}}{=} -\sqrt{1 - y^2}. \quad (1.6)$$

---

1When I first started writing these notes I felt obliged to draw all figures. This is a lot of work. However I’ve been told by Roger Samelson that the best educational strategy is to leave out most of the figures, and insist that students provide these themselves. Thus omission of some essential figures is a feature, rather than a bug.
Lecture 1. First order PDE’s

We need more information (e.g., some physical context) to make sense of this situation.

In the remainder of the plane, where $|y| < 1$ and $x^2 + y^2 > 1$, we can say that the solution is

$$h(x, y) = \begin{cases} +\sqrt{1 - y^2}, & \text{if } x \geq +\sqrt{1 - y^2}; \\
-\sqrt{1 - y^2}, & \text{if } x \leq -\sqrt{1 - y^2}. \end{cases} \quad (1.7)$$

Even this construction requires a little bit of reformulation: we’re saying that the solution in the region $x > +\sqrt{1 - y^2}$ ignores that part of the data curve in the region with $x < 0$, and vice versa. This sounds reasonable, but it is not strictly obvious from the bare formulation of the PDE plus data.

The transport equation

The transport equation for the unknown function $h(x, y)$ is

$$ah_x + bh_y = 0, \quad \text{where } a \text{ and } b \text{ are constants.} \quad (1.8)$$

The PDE in (1.1) is the special case $a = 1$ and $b = 0$.

The transport equation is a simple PDE that occurs frequently in applications. By inspection, the general solution of the transport equation is

$$h(x, y) = f(bx - ay), \quad (1.9)$$

where $f(p)$ is an arbitrary function.

We can write the transport equation as

$$a \cdot \nabla h = 0, \quad \text{where } a = (a, b). \quad (1.10)$$

In this form we interpret (1.8) geometrically as saying that the directional derivative of $u(x, y)$ along the vector $a$ is zero. We say that $h$ is transported along the vector $a$. Then the variation of $h$ must be in the direction orthogonal to $a$ i.e.,

$$h = f(p \cdot x), \quad (1.11)$$

where $p$ is any vector orthogonal to $a$. In writing the general solution (1.9) we happened to use $p = a \times \hat{z} = (b, -a)$. 

$$p \cdot a = 0$$
This observation is the key to solving the transport equation in higher dimensions e.g., in three dimensions we can interpret

\[ ah_x + bh_y + ch_z = 0 \]  

as saying that the directional derivative of \( h(x, y, z) \) along the vector \( a \equiv (a, b, c) \) is zero. Therefore the general solution of the three-dimensional transport equation is

\[ h = f(p \cdot x, q \cdot x), \]  

where \( p \) and \( q \) are any two linearly independent vectors orthogonal to \( a \) and \( f \) is an arbitrary function with two arguments.

Notice that the solution of the PDE (1.12) is a three-dimensional surface embedded in a four-dimensional space.

**Example:** Find a general solution of

\[ xh_x - yh_y = 0. \]  

If we change variables to \( \xi = \ln x \) and \( \eta = \ln y \) then the equation is transformed to the transport equation

\[ h_\xi - h_\eta = 0. \]  

The general solution is that

\[ h = f(\xi + \eta) = f(\ln(xy)) = g(xy), \]  

where \( f \) and \( g \) are arbitrary functions.

### 1.2 Arbitrary functions

Now let’s consider a slightly more complicated example of arbitrary functions. What do all of the following functions have in common? Each \( f_n \) is a function of the simpler function

\[ p(x, y) \equiv x^2 + 2y. \]  

\[ f_1(x, y) = x^4 + 4(x^2y + y^2 + 1), \]

\[ f_2(x, y) = \sin x^2 \cos 2y + \cos x^2 \sin 2y, \]

\[ f_3(x, y) = \frac{x^2 + 2y + 2}{3x^2 + 6y + 5}, \]  

(1.17)
Specifically

\[ f_1 = p^2 + 4, \quad f_2 = \sin p, \quad f_3 = \frac{p + 2}{3p + 5}. \]  

(1.19)

This also means that the \textit{contours} of the functions \( f_n(x, y) \) in the \((x, y)\) plane all coincide with the contours of the simple function \( p(x, y) \). Yet another consequence is that a ‘scatterplot’ of say \( f_1 \) against \( f_2 \) produces a curve (rather than a cloud) in the \((f_1, f_2)\)-plane.

Since \( p(x, y) \) is a solution of the PDE

\[ h_x - x h_y = 0, \]  

(1.20)

all of the \( f_n \)'s are also solutions of the PDE above. We can check this by substitution. Suppose that \( h \) is an arbitrary function of \( p \). Then

\[ h(p)_x = p_x \frac{dh}{dp}, \quad \text{and} \quad h(p)_y = p_y \frac{dh}{dp} \]  

(1.21)

so that

\[ h_x - x h_y = \frac{dh}{dp} (p_x - x p_y) = 0. \]  

(1.22)

In fact the \textit{general solution} of the PDE (1.20) is that \( h(x, y) \) an arbitrary function of \( x^2 + 2y \).

Since the function \( p(x, y) \) is so central to the solution of (1.20), let us “change coordinates” from \( x \) and \( y \) to \( p(x, y) \) and \( q(x, y) \). At the moment \( q(x, y) \) is just some other unspecified function of \( x \) and \( y \) (your choice).

We recall the rule for converting derivatives

\[ \partial_x = p_x \partial_p + q_x \partial_q = 2x \partial_p + q_x \partial_q, \]
\[ \partial_y = p_y \partial_p + q_y \partial_q = 2 \partial_p + q_y \partial_q. \]  

(1.23)

These relations imply

\[ \partial_x - x \partial_y = (q_x - qx) \partial_q. \]  

(1.24)

Thus, provided that we can find a \( q(x, y) \) satisfying

\[ x q_y - q_x \neq 0, \]  

(1.25)

we can rewrite (1.20) in terms of \((p, q)\) as simply

\[ h_q = 0. \]  

(1.26)

This should remind you of (1.1) — the solution is that \( h(x, y) \) is a possibly discontinuous arbitrary function of \( p(x, y) \).
**Example:** Solve
\[ u_x - xu_y = e^y. \] (1.27)
We change coordinates to \( p = x^2 + 2y \) and \( q = x \). In these new coordinates
\[ y = \frac{1}{2}p - \frac{1}{2}q^2, \quad \text{and} \quad \partial_x - x\partial_y = \partial_q. \] (1.28)
Thus the PDE is
\[ u_q = e^{\frac{1}{2}p} \cdot \frac{1}{2}q^2, \quad \text{or} \quad u = e^{\frac{1}{2}p} \int_0^q e^{-\frac{1}{2}s^2} \, ds + f(p), \] (1.29)
where \( f \) is an arbitrary function. We can now rewrite this in terms of the original variables \( x \) and \( y \):
\[ u(x, y) = e^y \int_0^x e^{-\frac{1}{2}s^2} \, ds + f(x^2 + 2y). \] (1.30)
This is the general solution of the PDE. The final step is to check by substitution that this solution works (do it).

**Example:** Solve \( u_x - xu_y = e^y \) with the boundary condition \( u(0, y) = 0 \).
The first step is to draw the characteristics curves on which \( p = x^2 + 2y \) is constant. Then you have to figure out where (or if) these characteristics intersect the curve on which the boundary condition is specified (that’s the \( y \)-axis in this problem). This drawing should convince you that every characteristic intersects the \( y \)-axis exactly once.
Applying the boundary condition to the general solution in (1.30) we have
\[ 0 = e^y \int_0^0 e^{-\frac{1}{2}s^2} \, ds + f(y). \] (1.31)
Thus \( f = 0 \) in (1.30) does the trick in this case.

**Example:** Solve \( u_x - xu_y = e^y \) with the boundary condition \( u(x, 0) = 0 \).
Now the boundary condition is applied on the \( x \)-axis. But characteristics with \( p(x, y) < 0 \) don’t intersect the \( x \)-axis. So in the region where \( p < 0 \) the function \( f \) in (1.30) is still arbitrary.
Characteristics with \( p(x, y) > 0 \) intersect the \( x \)-axis twice, at \( x = \pm \sqrt{p} \). This means that the boundary data is inconsistent with the PDE: there is no solution to this problem in the region where \( p > 0 \).
Suppose we ignore the red flag raised by the double intersection and try to determine \( f \) by insisting that
\[ 0 = e^{\frac{1}{2}x^2} \int_0^x e^{-\frac{1}{2}s^2} \, ds + f(x^2). \] (1.32)
Lecture 1. First order PDE’s

It seems that
\[ f(p) = -e^{\frac{1}{2}p} \int_0^{\sqrt{p}} e^{-\frac{1}{2}s^2} ds. \] (1.33)

But if we take the \( +\sqrt{p} \) in the upper limit of the integral then we satisfy the boundary condition on \( x > 0 \), but unfortunately not on \( x < 0 \) (and vice versa).

If we modify the boundary condition to \( u(x > 0, 0) = 0 \) then we’re in business. The solution to the modified problem is
\[ u = e^{y + \frac{1}{2}x^2} \left( \int_0^x e^{-\frac{1}{2}s^2} ds - \int_0^{+\sqrt{y+x^2}} e^{-\frac{1}{2}s^2} ds \right). \] (1.34)

Notice that \( u(x < 0, y) \neq 0. \)

1.3 The Jacobian

Suppose \( u(x, y) \) and \( v(x, y) \) are differentiable functions of \( x \) and \( y \). Then a necessary and sufficient condition for the existence of a functional relation,
\[ \phi(u, v) = 0, \quad \text{or perhaps} \quad u = A(v), \] (1.35)
is that the Jacobian
\[ \frac{\partial(u, v)}{\partial(x, y)} \equiv u_xv_y - u_yv_x \] (1.36)
is zero. For instance, in the example surrounding (1.1), one can show that
\[ \frac{\partial(p, u_n)}{\partial(x, y)} = 0. \] (1.37)

Let’s prove this algebraically by differentiating \( \phi(u, v) = 0 \) with respect to \( x \) and \( y \):
\[ u_x\phi_u + v_x\phi_v = 0, \quad \text{and} \quad u_y\phi_u + v_y\phi_v = 0. \] (1.38)

Eliminating the derivatives \( \phi_u \) and \( \phi_v \) we obtain
\[ \frac{\partial(u, v)}{\partial(x, y)} = 0. \] (1.39)

Therefore vanishing of the Jacobian is necessary for the existence of a function relation between \( u \) and \( v \). The condition is also sufficient.
Lecture 1. First order PDE’s

The geometric interpretation of the condition (1.39) is that \( u(x, y) \) and \( v(x, y) \) are functionally related if their contours coincide, which is equivalent to saying that \( \nabla u \) and \( \nabla v \) are parallel at every point. Now recall that
\[
\hat{z} \cdot (\nabla u \times \nabla v) = |\nabla u||\nabla v| \sin \theta
\]
where \( \theta \) is the angle between \( \nabla u \) and \( \nabla v \). So \( \nabla u \) and \( \nabla v \) are parallel at a point if and only if \( |\nabla u| \neq 0, |\nabla v| \neq 0 \) and \( \theta = 0 \). But it is easy to verify that
\[
\hat{z} \cdot (\nabla u \times \nabla v) = \frac{\partial (u, v)}{\partial (x, y)}.
\]

Solving (really) simple PDE’s with Jacobians

One way of solving a PDE such as
\[
(\cos y)u_y + (x \sin y)u_x = 0
\]
is by recognizing a thinly disguised Jacobian:
\[
(x \cos y)_x u_y - (x \cos y)_y u_x = 0.
\]
Thus the general solution is that \( u \) is an arbitrary function of \( x \cos y \).

But unfortunately not all PDEs of the form
\[
a(x, y)u_x + b(x, y)u_y = 0
\]
are disguised Jacobians: for the LHS to be a Jacobian one must have that \( a_x + b_y = 0 \).

Example Find the general solution of
\[
(x + 2y)u_x - (6x + y)u_y = 0
\]
by showing that the LHS is a Jacobian.

If the LHS is a Jacobian then there is a function \( v(x, y) \) such that
\[
v_y = x + 2y, \quad \Rightarrow \quad v_{yx} = 1.
\]
But we also require
\[
v_x = 6x + y, \quad \Rightarrow \quad v_{xy} = 1.
\]
Lecture 1. First order PDE’s

The equation passes the test so we know that the object of our desire, \( v(x, y) \), exists. Now we proceed to uncover \( v \) by integration

\[
v_x = 6x + y, \quad \Rightarrow \quad v = 3x^2 + xy + A(y),
\]
\[
v_y = x + 2y, \quad \Rightarrow \quad v = xy + y^2 + B(x). \quad (1.48)
\]

Subtracting our different expressions for \( v \) we soon see that

\[
v = 3x^2 + xy + y^2 + \text{constant}. \quad (1.49)
\]

Thus the solution is that \( u \) is an arbitrary function of \( 3x^2 + xy + y^2 \) ■

Example Is the PDE

\[
(x + 2y)w_x + (6x + y)w_y = 0. \quad (1.50)
\]

a Jacobian?

Since \((x + 2y)_x + (6x + y)_y \neq 0\), the LHS above is not a Jacobian ■

In next section we'll develop the method of characteristics to solve non-Jacobian PDE\( \)s such as (1.50).

1.4 The method of characteristics

Consider some function \( h(x, y) \) which you can visualize as the height of a surface above the \((x, y)\)-plane. If you move around on the \((x, y)\)-plane following a curve \( y = y(x) \) then you will observe changes in \( h \) both because \( x \) changes and also \( y \) changes. In fact, the total derivative of \( h(x, y) \) following this moving point is

\[
\frac{dh(x, y)}{dx} = h_x + \frac{dy}{dx}h_y. \quad (1.51)
\]

We will use this result again and again and again.

The PDE

\[
h_x + (x + y)^2h_y = 0 \quad (1.52)
\]

is not a disguised Jacobian and one can’t guess the solution. However comparing \((1.52)\) and \((1.51)\) we see that if we move in the \((x, y)\)-plane along a path \( P \) satisfying

\[
\frac{dy}{dx} = (x + y)^2 \quad (1.53)
\]
then on \( P \) the PDE implies
\[
\frac{dh}{dx} = 0. \tag{1.54}
\]
This means that \( h(x, y) \) is a constant on the path i.e., \( P \) is a contour or level-set of the function \( h(x, y) \).

Now it is easy to integrate (1.53) and we find
\[
\xi(x, y) = \tan^{-1}(x + y) - x \tag{1.55}
\]
where \( \xi \) is a “constant of integration”; curves of constant \( \xi \) are shown in Figure [1.2]. Regarding \( \xi \) is function of \( x \) and \( y \) defined by (1.55), the solution of (1.52) is
\[
h(x, y) = \text{an arbitrary function of } \xi(x, y). \tag{1.56}
\]
Thus the curves in Figure [1.2] are the “level sets” or the “contours” of the function \( h(x, y) \). The collection of special paths \( P \) along which a PDE collapses to an ODE such as (1.54) are characteristics. We’ll give a more careful definition in the next few lectures.

**Example** Find the solution of
\[
v_x + (x + y)^2 v_y = 1, \tag{1.57}
\]
subject to \( v(x, -x) = 0 \).

On the path defined by (1.53) the PDE is
\[
\frac{dv}{dx} = 1, \quad \text{or} \quad v = x + f \left( \tan^{-1}(x + y) - x \right). \tag{1.58}
\]
This is the general solution. The arbitrary function \( f \) is determined by the boundary condition that \( v = 0 \) on the curve \( y + x = 0 \):
\[
0 = x + f(-x). \tag{1.59}
\]
This implies that \( f(s) = s \) and \( v = \tan^{-1}(x + y) \). Figure [1.2] shows that there is a unique intersection, and the characteristics fill up the whole plane. Therefore in this problem the domain of definition is the entire \((x, y)\)-plane.

Notice that in this example the solution \( v \) is not constant on characteristics: the defining quality of characteristics is that the variation of \( v \) is determined by solving an ODE.
Example  Find the general solution of the non-Jacobian PDE

\[(x + 2y)w_x + (6x + y)w_y = 0.\]  \hspace{1cm} (1.60)

On the paths defined by

\[\frac{dy}{dx} = \frac{y + 6x}{2y + x},\]  \hspace{1cm} (1.61)

the PDE collapses to an ODE:

\[\frac{dw}{dx} = 0.\]  \hspace{1cm} (1.62)

That is, \(w\) is constant on the characteristics determined by solving \(1.61\). To integrate \(1.61\) we turn to section 1.7 of BO and observe that the equation is scale invariant i.e., the ODE is unchanged if \(y \rightarrow ay\) and \(x \rightarrow ax\). Following the BO recipe, we first substitute \(y = xY(x)\) so that

\[x \frac{dY}{dx} + Y = \frac{Y + 6}{2Y + 1},\]  \hspace{1cm} (1.63)

and if \(\lambda = \ln x\) then

\[\frac{dY}{d\lambda} = \frac{6 - 2Y^2}{1 + 2Y}, \quad \text{or} \quad \frac{1 + 2Y}{6 - 2Y^2}dY = d\lambda.\]  \hspace{1cm} (1.64)

To do the \(Y\)-integral I used \textsc{Mathematica}, and obtained

\[\frac{\sqrt{3} - 6}{12} \ln \left(\sqrt{3} + Y\right) - \frac{\sqrt{3} + 6}{12} \ln \left(\sqrt{3} - Y\right) = \lambda + c.\]  \hspace{1cm} (1.65)
Lecture 1. First order PDE’s

The constant of integration is \( c \). But \( c \) is also the label of the characteristic, or equivalently \( w \) is constant along curves of constant \( c \). Thus \( w \) is an arbitrary function of

\[
e^c = x^{-1} \left( \sqrt{3} + \frac{Y}{x} \right)^{\frac{\sqrt{3}+6}{12}} \left( \sqrt{3} - \frac{Y}{x} \right)^{-\frac{\sqrt{3}+6}{12}} \] (1.66)

1.5 Some books

Four good general texts on PDEs written by applied mathematicians are:


**Z** Partial Differential Equations of Applied Mathematics by E. Zauderer.

**K** Partial Differential Equations: Analytic Solution Techniques by J. Kevorkian.

**Sn** Elements of Partial differential Equations by I.N. Sneddon.

**CP** and **Sn** have the advantage of brevity. **Z** is a leisurely introduction, with lots of problems. The first two chapters of **Sn** are a good introduction to the geometric theory of first-order PDE’s.

There are several ‘math methods’ books written by physicists for physicists which contain a lot of material on PDE’s. Four good ones are:

**So** Partial Differential Equation by A. Sommerfeld.

**MW** Mathematical Methods of Physics by J. Matthews & R.L. Walker.

**W** Mathematical Analysis of Physical Problems by P.R. Wallace.

**MF** Methods of Theoretical Physics; Part I and II by P.M. Morse & H. Feshbach.

**MF** is useful as a compendium of formulas and methods. **So** is a very readable classic.

Two other classic textbooks on wave dynamics, characteristics and much more are:

**W** Linear and Nonlinear Waves by G.B. Whitham.
Lecture 1. First order PDE’s


W is a tough book, but well worth the effort. The first chapter of ZR is a good introduction to gas dynamics and the wave equation.

General texts on mathematical methods are:

BO Advanced Mathematical Methods for Scientists and Engineers by C. Bender & S.A. Orszag.

JJ Methods of Mathematical Physics by H. Jeffreys & B.S. Jeffreys.

Useful handbooks are:

AS Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables by M. Abramowitz & I. Stegun.

GR Tables of Integrals, Series and Products by I.S. Gradshteyn & I.M. Rhyzik. (The recent editions are edited by A. Jeffrey & D. Zwillinger).

1.6 Problems

Problem 1.1. Evaluate the derivative

\[(\partial_x - x \partial_y) \ln \left( \frac{f_1 f_3}{1 - \sin^{1/3}(f_2)} \right) \cos(x - y),\]

where the functions \(f_n(x,y)\) are defined in (1.17)

Problem 1.2. Find the general solution of \(u_x - xu_y = 1\).

Problem 1.3. Find the general solution of \(u_x - xu_y = f'(x)g(x^2 + 2y)\). Here \(f\) and \(g\) are arbitrary functions and \(f'\) is the derivative of \(f\).

Problem 1.4. (i) Find the solution, and where in the \((x, y)\)-plane there is a solution (i.e., the domain of definition), of the PDE \(h_x - xh_y = 0\), subject to the boundary condition that \(h(x, x^2/2) = x\). (ii) Same PDE, but with the boundary condition \(h(x, x^2/2) = x^2\).
**Problem 1.5.** Consider the transport equation $h_x + h_y = 0$ with the data $h(x, x^2) = f(x)$ i.e., $h$ is specified on the parabola $y = x^2$ as equal to a specified function $f(x)$. Discuss the solution of this problem — where in the $(x, y)$- plane is the solution defined?

**Problem 1.6.** Find an example of a function $u(x, y')$ which is discontinuous at $(x, y') = (0, 0)$, even though all directional derivatives exist at $(x, y') = (0, 0)$.

**Problem 1.7.** (i) Spot the Jacobian and find the general solution of the following PDE

$$(x + y)f_x - (2x + y)f_y = 0.$$  

(ii) Consider the PDE

$$(ax + by)g_x + (cx + dy)g_y = 0,$$

where $a$, $b$, $c$ and $d$ are all constants. Find the condition ensuring that this PDE is a Jacobian, and then find the general solution.

**Problem 1.8.** Is the LHS of

$$e^x \cos y \ u_x - e^x \sin y \ u_y = 0,$$

a Jacobian? How about

$$(\cos x \cosh y - \sin x \sinh y)u_x - (\sin x \sinh y + \cos x \cosh y)u_y = 0?$$

**Problem 1.9.** Let

$$f(x, y) = \int_0^\infty \frac{\exp[-y\sqrt{1 + t^2}] \sin xt}{\sqrt{1 + t^2}} \ dt.$$  

Use integration by parts to show that

$$f(x, y) = \frac{x}{y} \int_0^\infty \exp[-y\sqrt{1 + t^2}] \cos xt \ dt.$$  

Next, differentiate with respect to $x$ and $y$ and show that

$$f_y = \left[ (y/x)f \right]_x.$$  

Show that the solution of this PDE is $f = xg(x^2 + y^2)$ where $g$ is an arbitrary function. Find a simple form for $g$. (From CP.)
Problem 1.10. Find a solution of the PDE
\[ u_{xx}u_{yy} - u^2_{xy} = 0, \quad u(x,0) = \frac{1}{2}x^2, \quad u_y(x,0) = x. \]
(Hint: look for a Jacobian.)

Problem 1.11. Consider the three functions
\[ \phi \equiv x + y + z, \quad \theta = x^2 + y^2 + z^2, \quad \chi \equiv xy + yz + zx. \]
(i) Show that one of these can be written as a function of the other two.
(ii) Find a PDE of the form
\[ au_x + bu_y + cu_z = 0, \]
with \(a, b\) and \(c\) functions of \((x,y,z)\),
whose general solution is that \(u\) is an arbitrary function of \(\phi\) and \(\theta\).

Problem 1.12. Suppose you are given three functions of \((x,y,z)\) (e.g., as in problem \ref{problem:1.11}). Find a three-dimensional generalization of the Jacobian criterion which tests if such a functional relation exists between the three. Give both an algebraic proof and a geometric interpretation.

Hint: Denote the three given functions by \(u(x,y,z)\), \(v(x,y,z)\) and \(w(x,y,z)\). Suppose that there is some unknown functional relation \(\Phi(u,v,w) = 0\), connecting \(u\), \(v\) and \(w\). Differentiating with respect to \(x\), \(y\) and \(z\) we find
\[
\begin{align*}
u_x \Phi_u + v_x \Phi_v + w_x \Phi_w &= 0, \\
v_y \Phi_u + v_y \Phi_v + w_y \Phi_w &= 0, \\
v_z \Phi_u + v_z \Phi_v + w_z \Phi_w &= 0.
\end{align*}
\]
This is a \(3 \times 3\) set of linear equations for the three unknowns \((\Phi_u, \Phi_v, \Phi_w)\). If the only solution is \((\Phi_u, \Phi_v, \Phi_w) = (0,0,0)\) then there is no nontrivial functional relation.

Problem 1.13. Reduce the following PDE’s to the transport equation and construct the general solution:
\[
\begin{align*}
y u_x + xu_y &= 0, \\
x u_x + y u_y &= 0, \\
u_y + e^x u_x &= u, \\
u_x + (x u)_x &= 0.
\end{align*}
\]

Solution: For the first example, divide by \(xy\) make the change of variables \(\xi = x^2\) and \(\eta = y^2\), so that the equation becomes
\[
u_{\xi} + u_{\eta} = 0 \implies u = f(\xi - \eta) = f(x^2 - y^2).
\]
This problem illustrates how one can sometimes quickly beat simple PDE into a soluble form by making a change of variables.
Problem 1.14. Find the general solution of the PDEs

\[ a(x)u_x + b(y)u_y = 0, \]
\[ a(x)v_x + b(y)v_y + c(z)v_z = 0. \]

Problem 1.15. Find the general solution of the PDE \( u_x + (x + y)u_y = 0 \).

Check your answer by substitution.

Problem 1.16. Find the general solution of the PDE \( u_x + (x + y)^2u_y = u \).

Check your answer by substitution.

Problem 1.17. (i) Find the general solution of the PDE \( xyu_x + u_y = 0 \).

(ii) Find the solution satisfying the boundary data \( u(x, 0) = x \). What is the domain of definition of this solution?

Problem 1.18. Consider the PDE \( u_x + \beta u_y = 0 \) with the boundary condition \( u(x, x - 1) = x \). Find the solution and state any condition which must be imposed on the constant \( \beta \) in order for the solution to exist.

Problem 1.19. Find the general solution of \( au_x + bu_y + cu_z = 0 \), and \( pu_x + qu_y + ru_z = 0 \).

Above \( a, b \) etc. are all constants.
Lecture 2

Linear evolution equations

In this lecture we discuss some relatively simple applications of the method of characteristics to linear evolution problems. “Evolution” means changing in time — so one of our independent variables in this lecture is time $t$.

2.1 Conservation laws

First-order PDE’s arise in applications frequently as a result of conservation laws. The fundamental idea here is simple: count the amount of stuff in some control volume and then:

$$\frac{d}{dt} \left[ \text{Amount of Stuff} \right] = \left[ \text{Stuff entering} \right] - \left[ \text{Stuff leaving} \right].$$  \hspace{1cm} (2.1)

Consider traffic on the northbound lanes of a highway as an example. We use the coordinate $x$ to denote distance along the road. The density, $\rho(x, t)$, is the number of cars per length. In the interval $a < x < b$ the total number of cars is

$$\text{Number of cars in the interval } a < x < b = \int_a^b \rho(x, t) \, dx.$$  \hspace{1cm} (2.2)

We have made the continuum approximation by assuming that there is a well defined density which is a smooth function of position. If a car is 5 meters long it makes no sense in (2.2) to pick an interval of length 1 meter. The introduction of the density $\rho(x, t)$ requires a separation
in length scales between the distance over which \( \rho(x, t) \) changes appreciably (e.g. one kilometer) and the distance between cars (e.g. a few meters).

We apply the principle in (2.1) by picking an interval of highway \( a < x < b \), and counting the number of cars which pass \( x = a \) in a time \( dt \). Thus we define the flux (cars per second passing \( x = a \)) as \( f(a, t) \). We can do the same at the other end of the control length and so determine \( f(b, t) \). We are ignoring off-ramps and on-ramps which introduce cars into the middle of \( (a, b) \). Then (2.1) tells us that

\[
\frac{d}{dt} \left[ \text{Number of cars in the interval } a < x < b \right] = f(a, t) - f(b, t). \tag{2.3}
\]

In other words

\[
\frac{d}{dt} \int_{a}^{b} \rho(x, t) \, dx + f(b, t) - f(a, t) = 0. \tag{2.4}
\]

Letting \( a \to b \) in (2.4), with \( x \) sandwiched in the middle, we obtain the differential statement of the conservation law:

\[
\rho_t + f_x = 0. \tag{2.5}
\]

If we can find a relation between \( f(x, t) \) and \( \rho(x, t) \) then the conservation law in (2.5) is a PDE for the density of traffic.

### The linear advection equation

Suppose, for example, that all of the drivers are law abiding so that everyone is moving at exactly \( c = 65 \text{mph} \). In this unrealistic case we get a simple and important example of a first order PDE:

\[
f = c \rho \quad \text{in (2.5)} \quad \Rightarrow \quad \rho_t + c \rho_x = 0. \tag{2.6}
\]

The boxed equation is the linear advection equation — you’ll notice it’s a special case of the transport equation from lecture 1.

There are two ingredients in this argument. There is the conservation law in (2.3) and the connection between flux and density, namely \( f = c \rho \). Once we accept the continuum approximation the integral form of the conservation law in (2.3) is unassailable. But the flux-density relation between might be a matter of debate. In realistic cases, such as traffic...
flow, the connection between flux and density has to be determined empirically and the resulting PDE is nonlinear. Moreover, as we’ll see later, the transition from the integral law (2.3) to the differential law (2.5) can also be difficult: \( \rho \) can spontaneously develop discontinuities (shocks).

By now you should be able to write down the general solution of (2.6) on auto-pilot:

\[
\rho(x, t) = q(x - ct), \tag{2.7}
\]

where \( q \) is arbitrary function. To determine \( q \) we need a combination of initial and boundary conditions. The simplest case is the initial value problem in which \( q \) is determined by specifying an initial condition which applies for all \( x \). For example:

\[
\rho(x, 0) = \exp(-x^2), \quad \Rightarrow \quad \rho(x, t) = \exp[-(x - ct)^2]. \tag{2.8}
\]

The pulse preserves its shape and moves with constant velocity \( c \) (see the top panel of figure 2.1). This is a travelling wave solution of the PDE in (2.6).

## 2.2 A recipe for semilinear evolution equations

We can solve a first-order PDE in the form\(^1\)

\[
\rho_t + c(x, t)\rho_x = s(x, t, \rho), \quad \rho(x, 0) = \rho_0(x), \tag{2.9}
\]

\(^1\)We call this a “semilinear” equation because the right hand side is a possibly nonlinear function of \( \rho \).
Lecture 2. Linear evolution equations

using the following recipe:

- define a characteristic coordinate $\xi(x,t)$ by
  \[
  \frac{dx}{dt} = c(x,t), \quad x(\xi,0) = \xi; \quad (2.10)
  \]
- sketch the characteristic diagram;
- find $\rho(\xi,t)$ by integrating the ODE
  \[
  \frac{d\rho}{dt} = s[x(\xi,t),t,\rho], \quad \rho(\xi,0) = \rho_0(\xi); \quad (2.11)
  \]
- eliminate $\xi$ in favor of $x$ and $t$;
- check by substitution.

Conservation versus non-conservation

Let us apply the recipe and solve the PDE

\[
\rho_t - (x\rho)_x = 0, \quad \rho(x,0) = e^{-x^2}. \quad (2.12)
\]

Notice that this PDE has the form of a conservation law i.e., the flux is $-x\rho$. You can imagine a bunch of particles moving along the $x$-axis with speed $c = -x$, and density $\rho(x,t)$. Since particles don’t disappear, the total number of particles — obtained by integrating $\rho(x,t)$ over $x$ — must be conserved.

Let’s do this explicitly. We integrate\[\text{(2.12)}\] over $a < x < b$ and obtain

\[
\frac{d}{dt} \int_a^b \rho \, dx = \left[ x\rho \right]^b_a. \quad (2.13)
\]

$\rho$ flows in and out of $(a,b)$ at the end points and that accounts for all changes in the total amount of $\rho$ in the interval.

Equation\[\text{(2.12)}\] above can be written in the standard form\[\text{(2.9)}\]:

\[
\rho_t - x \rho_x = \rho, \quad (2.14)
\]
Lecture 2. Linear evolution equations

and we are off to the races. First, we obtain the characteristic coordinate

$$\frac{dx}{dt} = -x, \quad x(0, \xi) = \xi \Rightarrow x = \xi e^{-t}. \quad (2.15)$$

Next we solve the ODE

$$\frac{d\rho}{dt} = \rho, \quad \rho(\xi, 0) = e^{-\xi^2}, \quad \Rightarrow \quad \rho(x, t) = e^{t-\xi^2}. \quad (2.16)$$

Notice that $\xi$ is a 'constant on characteristics' so when we integrate the ODE in (2.16) the constant of integration is an arbitrary function of $\xi$. To determine this arbitrary function of $\xi$ we look at at $t = 0$. At the initial instant $\xi = x$ so that the initial condition for the ODE (2.16) follows from (2.12): $\rho_0(\xi) = \exp(-\xi^2)$.

In the next step of the recipe we eliminate $\xi = x \exp(t)$:

$$\rho(x, t) = \exp\left(t - e^{2t}x^2\right). \quad (2.17)$$

Finally, we check by substitution that this really is the solution of the PDE (2.12).

Notice if the initial condition in (2.12) is changed to $\rho(x, 0) = \rho_0(x)$ then the solution is just $\rho(x, t) = e^t \rho_0(e^t x)$. From this solution we can verify that conserves particles are conserved i.e., you should show that

$$m \equiv \int_{-\infty}^{\infty} \rho(x, t) \, dx \quad (2.18)$$

is constant (independent of time). We know this even before we go to the trouble of solving the equation, but it is interesting to see how it all works out.

Let’s dwell further on the difference between conservation and non-conservation. Unlike (2.12), the PDE

$$\theta_t - x \theta_x = 0, \quad (2.19)$$

is not a conservation equation. Notice what happens if we integrate-by-parts over $a < x < b$:

$$\frac{d}{dt} \int_a^b \theta \, dx = \left[ x \theta \right]^b_a - \int_a^b \theta \, dx \quad (2.20)$$

$\theta$-stuff appears out of thin air in the middle of $(a, b)$. 

SIO203C, W.R. Young, March 21, 2011
2.3 Entry on the half-line

Now we turn to problems defined on the half-line \( x \geq 0 \). Think of \( x = 0 \) as the beginning of a highway and suppose we count the number of cars entering at \( x = 0 \). If this entry rate increases then there is a bulge in the traffic density, \( \rho \), which propagates down the roadway. This bulge is a signal indicating changes in the upstream conditions. In our simplest model, with constant speed \( c \), the signal from \( x = 0 \) reaches \( x \) at time \( x/c \). Another example might be one-dimensional radiative transfer: \( \rho(x,t) \) is the density of photons moving along the positive half of the \( x \)-axis with the speed of light \( c \). The source at \( x = 0 \) is a laser emitting new photons.

Note carefully we’re discussing entry problems. That is

\[
c > 0
\] (2.21)

so that cars or photons are entering at \( x = 0 \). We’ll come back at the end of this section and discuss the more subtle exit problem in which \( c < 0 \).

To formulate the problem mathematically, denote the rate at which cars enter at \( x = 0 \) by \( R(t) \) (cars per second). The conservation equation is our old friend the linear advection equation

\[
\rho_t + c\rho_x = 0.
\] (2.22)

For a PDE model we specify both an initial condition

\[
\rho(x,0) = F(x), \quad \text{for } x > 0,
\] (2.23)

and a boundary condition

\[
c\rho(0,t) = R(t), \quad \text{for } t > 0.
\] (2.24)

Notice the dimensions — \( R \) is cars per second, \( \rho \) is cars per meter and \( c \) is meters per second.

To solve this problem it is essential to draw the \((x,t)\)-plane and the family of curves on which \( \rho(x,t) \) is constant. These characteristics are the lines

\[
x - ct = \xi,
\] (2.25)

where \( \xi \) is the intercept of the characteristic with the \( x \)-axis (see figure 2.1). You can think of \( \xi \) as the name of each member of the family of

The general solution of

\[
\rho = q(x - ct) = q(\xi).
\]
Lecture 2. Linear evolution equations

Figure 2.2: The \((x, t)\)-diagram, or *characteristic diagram*, for the signaling problem. The shaded region is \(x - ct = \xi > 0\), in which the solution is determined by information from \(t = 0\). The unshaded region is \(x - ct = \xi < 0\), in which the solution is determined by information from \(x = 0\).

The general solution of (2.22) can also be written
\[
\rho = p(t - x/c).
\]

Thinking carefully about this argument, you’ll see that characteristics are spacetime curves along which signals propagate. To determine the arbitrary function, \(q\), we follow each characteristic till we hit the boundary of the spacetime domain.

Summarizing, the solution of this “signaling” problem is
\[
\rho(x, t) = \begin{cases} 
F(x - ct), & \text{if } x - ct > 0, \\
c^{-1}R(t - x/c), & \text{if } x - ct < 0. 
\end{cases}
\]

In general \(F(0) \neq R(0)\) so the solution might be discontinuous on the line \(x = ct\). For instance, suppose \(\rho(x, 0) = 0\) and \(R\) is a constant. Then the solution is \(\rho(x, t) = c^{-1}RH(ct - x)\) where \(H\) is the Heaviside step function.
Lecture 2. Linear evolution equations

The exit problem, \( c < 0 \)

If you redraw Figure 2.2 with \( c < 0 \) then you immediately see that the characteristics intersect the boundary twice: the lines of constant \( x - ct \) intersect both the positive \( x \)-axis and the positive \( t \)-axis. Consequently one if arbitrarily specifies both \( F(x) \) and \( R(t) \) then usually the problem has no solution.

In discussing this non-existence one must be sensitively aware of the distinction between the mathematical formulation and the physical situation. The former is an approximation or idealization of the latter. If one is not aware of the physical motivation then one declares that the PDE plus boundary/initial condition has no solution, and that’s the end of the matter. But thinking of the original motivation for this model in terms of traffic flow or photons it seems clear, to me at least, that one can prescribe the initial condition on the positive \( x \)-axis, and one cannot prescribe the exit condition on the positive \( t \)-axis. The exit at \( x = 0 \) takes what it gets and doesn’t complain. If the exit does complain (e.g., if some fraction of the incident photons are reflected back into the domain) then one must reformulate the physical problem to account for this complication.

Example: Formulate a half-line \((x > 0)\) model in which photons move in both directions with speed \( \pm c \). Assume that all photons with speed \( -c \) incident on \( x = 0 \) are reflected back with speed \( +c \). Photons with speed \( +c \) are also emitted at \( x = 0 \) at a rate \( R(t) \).

We need two densities, \( \rho^+(x,t) \) for photons with \( +c \) and \( \rho^-(x,t) \) for photons with speed \( -c < 0 \). The PDEs are

\[
\rho^+_t + c \rho^+_x = 0, \quad \rho^-_t - c \rho^-_x = 0. \tag{2.28}
\]

The boundary condition at \( x = 0 \) is

\[
c \rho^+(0,t) = R(t) + c \rho^-(0,t). \tag{2.29}
\]

We must also supply initial conditions for \( \rho^+ \) and \( \rho^- \) on the half-line. As a sanity check, you should show that

\[
\frac{d}{dt} \int_0^\infty \rho^+(x,t) + \rho^-(x,t) \, dx = R(t) \quad \blacksquare \tag{2.30}
\]
2.4 Age-Stratified Populations

Let us give another example of a context in which the half-line linear advection equation, (2.6) arises. To characterize the age structure of the population of San Diego at \( t = 0 \) we use a histogram:

\[
h_0(a) \, da = \text{the number of people with age } a \in (a, a + da) \text{ at } t = 0.
\]

(2.31)

The age-coordinate, \( a \), is strictly positive so that the histogram lives on the half-line \( a > 0 \).

Suppose for a moment that this population is closed — we ignore births, deaths and immigration. Then it is obvious that at \( t > 0 \) the histogram of ages is \( h(a, t) = h_0(a - t) \). Therefore the evolution of the age structure of this population is described by the PDE

\[
h_t + h_a = 0.
\]

(2.32)

In this example everyone strictly observes the speed limit by moving along the age axis at a rate one year per year — there is no doubt that the connection between flux and density is correct.

Now we face the facts of life by admitting that some people die, just as new people (babies) replace them. We can make the model more realistic by incorporating birth and death. The probability of death depends on age and also on time. This fact is incorporated by adding a loss term to the right hand side of (2.32):

\[
h_t + h_a = -\mu(a, t)h,
\]

(2.33)

where \( \mu(a, t) \) is the probability that in the interval \( (t, t + dt) \) an individual of age \( a \) will die. Births correspond to the \( a = 0 \) boundary condition. If the birth rate is \( b(t) \) (babies per second) then

\[
h(0, t) = b(t).
\]

(2.34)

An increase in \( b(t) \) produces a disturbance which will move through the histogram (a baby boom).

We can convince ourselves that this formulation is sensible by noticing that the population size is

\[
N(t) = \int_0^\infty h(a, t) \, da.
\]

(2.35)
If we integrate (2.33) from \( a = 0 \) to \( a = \infty \), and use the boundary condition in (2.34), we get

\[
\frac{dN}{dt} = b(t) - \int_0^\infty \mu(a,t)h(a,t) \, da.
\] (2.36)

This is the obvious conclusion that the total size of the population changes if there is an imbalance between the birth-rate and the death-rate.

**Populations in equilibrium**

Consider an age-stratified population whose death-rate is depends only on age, \( \mu = \mu(a) \). Suppose further that the birth rate \( b \) is constant. Then we expect that eventually the age-structure of the population will stop changing — the population will be in equilibrium. We can determine this equilibrium by setting \( h_t = 0 \) and looking for a steady solution of the PDE (2.33) with the boundary condition in (2.34). This turns out to be very easy because now we deal only ODE’s:

\[
h_{\text{eq}}(a) = b \exp \left( - \int_0^a \mu(a') \, da' \right).
\] (2.37)

We'll refer to \( S(a) \) as the “survival function”. Using this notation, the total population is

\[
N = b \int_0^\infty S(a) \, da.
\] (2.38)

It seems intuitive that for a population in equilibrium

\[
N = b \times \text{average life span}.
\] (2.39)

In fact, this intuitive expectation is correct because it turns out that

The average life span is the 'renewal' or 'flushing' time.

\[
\text{average life span}, \tau = \int_0^\infty S(a) \, da,
\] (2.40)

(see problem 2.21).

The average age of the population is

\[
\bar{a} = N^{-1} \int_0^\infty a h_{\text{eq}}(a) \, da,
\] (2.41)
and using the results above
\[
\bar{a} = \frac{\int_{0}^{\infty} a S(a) \, da}{\int_{0}^{\infty} S(a) \, da}.
\]
(2.42)

In the special and unrealistic case of constant $\mu$ the survivor function is a decaying exponential, the total population is $N = b/\mu$, and the average age of the population is equal to the average life span, $\tau = \bar{a} = \mu^{-1}$.

**A baby boom**

Let us go back to our earlier model of an age-stratified population. Recall that the PDE formulation is
\[
h_t + h_a = -\mu(a) h, \quad h(0,t) = b(t),
\]
(2.43)
where $b(t)$ is flux of people into the histogram (babies per second). This is naturally a signalling problem on the axis $0 < a < \infty$.

Let us adopt a simple model of the death-rate: assume that $\mu$ is constant. Suppose that the birth rate is also a constant, $b_0$. If we wait for a very long time then the age structure of the population comes into equilibrium. In other words, the solution of (2.33) is steady and we can very quickly find:
\[
h_{eq}(a) = b_0 e^{-\mu a}.
\]
(2.44)
The steady-state population is $N_0 = b_0/\mu$.

We suppose now that the population is in steady state when $t < 0$ but then at $t = 0$ a baby boom starts. We can model this situation by solving the PDE
\[
h_t + h_a = -\mu h, \quad h(a,0) = b_0 e^{-\mu a},
\]
(2.45)
with the birth rate
\[
h(0,t) = b_0 (1 + e^{-\alpha t}).
\]
(2.46)
At $t = 0$ the birth rate $b(t)$ suddenly jumps to $2b_0$. But then $b(t)$ falls back again to $b_0$.

We solve this problem by noticing that the PDE in (2.45) can be written as
\[
(e^{\mu a} h)_t + (e^{\mu a} h)_a = 0.
\]
(2.47)
Thus $\tilde{h} \equiv \exp(\mu a) h$ satisfies the linear advection equation and the general solution of the PDE in (2.45) is therefore

$$\tilde{h}(a, t) = q(a - t), \quad \Rightarrow \quad h(a, t) = e^{-\mu a}q(a - t). \quad (2.48)$$

Now we have to determine the function $q$ using the boundary condition at $a = 0$ and the initial condition at $t = 0$. This is exactly the signaling problem discussed earlier in this lecture. Here is the answer

$$h(a, t) = \begin{cases} b_0 e^{-\mu a}, & \text{if } a - t > 0, \\ b_0 e^{-\mu a} \left[1 + e^{\alpha(a-t)}\right], & \text{if } a - t < 0. \end{cases} \quad (2.49)$$

To visualize the solution we try a MATLAB waterfall plot (see figure 2.3). Is it intuitively obvious to you that the age structure of the population is not altered by the baby boom till $t > a$?

A final remark is that we might want to determine the total size of the population, $N(t)$ in (2.35). Of course we can do this by direct integration of the explicit solution (2.49) over $0 < a < \infty$. But there is an slicker approach which you can use to obtain $N(t)$ even before the PDE is solved. Do you see how to do this?
2.5 Probability generating functions

Probability theory provides interesting and nontrivial applications of the
method of characteristics. An historical example is provided by the en-
gineers of the Ericsson phone company, who are interested in how many
phone lines are busy at time $t$ in Stockholm. This is a random process,
and so they define

$$ p_n(t) = \text{probability that at } t \text{ there are } n \text{ lines busy}. \quad (2.50) $$

We assume that phone conversations are independent and that in an
interval $dt$ people terminate a call with a probability $\mu dt$. Of course, if
there are $n$ conversations the probability of a termination is $n \mu dt$. We
assume that a new call is initiated with a probability $\lambda dt$.

Notice that the initiation process is independent of how many lines
happen to be busy: we are assuming that the population of Stockholm is
much greater than the number of busy lines. In this case the initiation of
a new call in $(t, t + dt)$ happens with probability $\lambda dt$ no matter whether
10 people or 1000 people are chatting on the phone. This might is plau-
sible if the population of Stockholm is $10^6$, but not if the population is
$10^4$. We also neglect the probability that a call finishes and starts on the
same line because that’s proportional to $(dt)^2$.

These assumptions leads to an infinite system of coupled ordinary
differential equations

\[
\begin{align*}
\dot{p}_0 &= -\lambda p_0 + \mu p_1, \\
\dot{p}_1 &= \lambda p_0 - (\lambda + \mu)p_1 + 2\mu p_2, \\
\dot{p}_2 &= \lambda p_1 - (\lambda + 2\mu)p_2 + 3\mu p_3, \\
\dot{p}_3 &= \lambda p_2 - (\lambda + 3\mu)p_3 + 4\mu p_4,
\end{align*}
\]

and so on.

To explain how we arrive at (2.51) consider the gain and loss of prob-
ability of the realizations with no lines busy:

$$ \dot{p}_0(t) = - \{ \text{loss if a call arrives} \} + \{ \text{gain if a realization with one busy line \textbf{"hangs-up"}} \} . \quad (2.52) $$

The loss occurs at a rate $\lambda p_0(t)$ and the gain at a rate $\mu p_1(t)$. If we
Lecture 2. Linear evolution equations

considering realizations with 3 busy lines we would write

\[ \dot{p}_3(t) = + \{ \text{gain from state 2 due to a new call} \} - \{ \text{loss from state 3 to state 4 due to a new call} \} - \{ \text{loss from state 3 to state 2 due to a hang-up} \} + \{ \text{gain from state 4 due to a hang-up} \}. \] (2.53)

\[ \dot{p}_3(t) = + \{ \text{gain from state 2 due to a new call} \} - \{ \text{loss from state 3 to state 4 due to a new call} \} - \{ \text{loss from state 3 to state 2 due to a hang-up} \} + \{ \text{gain from state 4 due to a hang-up} \}. \] (2.54)

The four terms on the right hand side above correspond to the four terms on the right hand side of (2.51). Do you see why probabilists call this a “one-step” process — this a Markov chain in which state \( n \) gains and loses probability only from its immediate neighbours at \( n \pm 1 \).

An essential check on our argument is that if we sum the system (2.51) we find

\[ \frac{d}{dt} \sum_{n=0}^{\infty} p_n(t) = 0, \quad \Rightarrow \quad \sum_{n=0}^{\infty} p_n(t) = 1. \] (2.55)

Thus probability is conserved.

The mean number of lines in action is

\[ \bar{n}(t) = \sum_{n=0}^{\infty} np_n(t). \] (2.56)

From (2.51) it follows that

\[ \frac{d\bar{n}}{dt} = \lambda - \mu \bar{n}. \] (2.57)

This is the mean field description of the process: as \( t \to \infty \) the expected number of busy lines is \( \lambda/\mu \). However we do not know how large the fluctuations about this mean state are likely to be.

This is interesting, but what does it have to with PDE’s? One way of obtaining the complete solution of (2.51) is to introduce the probability generating function:

\[ G(z, t) \equiv \sum_{n=0}^{\infty} p_n(t)z^n. \] (2.58)

Fiddling around with (2.51) we can show that \( G(z, t) \) satisfies

\[ G_t + \mu(z - 1)G_z = \lambda(z - 1)G. \] (2.59)
Lecture 2. Linear evolution equations

Notice this PDE is not a conservation equation, but it is a semilinear equation. Once we specify an initial condition we can solve this first-order PDE and figure out the Taylor series expansion of $G(z,t)$. The coefficients in this expansion are the probabilities $p_n(t)$.

For instance, suppose that at $t = 0$ there are no lines in action. Then $p_0(0) = 1$ and all the other $p_n$’s are zero. In this case the initial condition for (2.59) is

$$G(z,0) = 1.$$  \hfill (2.60)

We are off to the races with our semilinear recipe. The characteristic coordinate, $\xi(x,t)$ is defined by

$$\frac{dz}{dt} = \mu(z - 1), \quad z(0) = \xi.$$  \hfill (2.61)

Solving this system:

$$z = (\xi - 1)e^{\mu t} + 1, \quad \text{or} \quad \xi = e^{-\mu t}(z - 1) + 1.$$  \hfill (2.62)

On the characteristics the PDE collapses to

$$\frac{dG}{dt} = \lambda(z - 1)G = (\xi - 1)e^{\mu t}G.$$  \hfill (2.63)

Integrating this ODE and applying the initial condition $G(0,0) = 1$ we discover that

$$G(z,t) = \exp \left[ (\xi - 1)\frac{\lambda}{\mu}(e^{\mu t} - 1) \right].$$  \hfill (2.64)

Eliminating $\xi(z,t)$ using (2.62) we finally have the solution of the PDE

$$G(z,t) = \exp \left[ (z - 1)\frac{\lambda}{\mu}(1 - e^{-\mu t}) \right].$$  \hfill (2.65)

As a quick check on our work, we notice from the definition in (2.58) that we must have

$$G(1,t) = \sum_{n=0}^{\infty} p_n(t) = 1.$$  \hfill (2.66)

The expression in (2.65) passes this basic test.

We also see from (2.58) that

$$G_z(1,t) = p_1 + 2p_2(t) + 3p_3(t) + \cdots = \tilde{n}(t).$$  \hfill (2.67)
Lecture 2. Linear evolution equations

Taking the $z$-derivative of (2.65) and setting $z = 1$ we quickly find

$$\bar{n}(t) = \frac{\lambda}{\mu} \left(1 - e^{\mu t}\right).$$

But we can also calculate $\bar{n}(t)$ by solving (2.57) with the initial condition $\bar{n}(0) = 0$. This approach also gives (2.68) — we are starting to have some confidence in (2.65)! Of course we should really check (2.65) by substituting it back into the PDE...

After checking (2.65) we are finished with the PDE, but not with generatingfunctionology. We can rewrite the probability generating function in (2.65) as

$$G(z, t) = e^{(z-1)\bar{n}},$$

where $\bar{n}(t)$ is given in (2.68). Recalling the Taylor series expansion

$$e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!},$$

we then find from (2.69) that the probability that $k$ lines are busy is

$$p_k(t) = e^{-\bar{n}} \frac{\bar{n}^k}{k!}.$$

This is a Poisson distribution.

The birth-death processes

Consider a population of bugs reproducing and dying in your refrigerator. $p_k(t)$ is the probability that there are $k$ bugs at time $t$. We suppose that the probability per time that a single bug divides is $\lambda dt$. Thus, if there are $n$ bugs the probability of a birth is $n \lambda dt$. Likewise $n \mu dt$ is probability of a death in $dt$. The mean field model for the expected number of bugs is just

$$\frac{d\bar{n}}{dt} = \gamma \bar{n},$$

where $\gamma \equiv \lambda - \mu$ is the growth rate. This tells us that on average the bug population changes exponentially with time. Of course, even if $\gamma > 0$, bad luck might extinguish a particular bug culture. This is an example
Lecture 2. Linear evolution equations

of a population fluctuation. To quantitatively understand this we turn to the birth-death equations:

\[
\begin{align*}
\dot{p}_0 &= \mu p_1, \\
\dot{p}_1 &= -(\lambda + \mu) p_1 + 2\mu p_2, \\
\dot{p}_2 &= \lambda p_1 - 2(\lambda + \mu) p_2 + 3\mu p_3,
\end{align*}
\] (2.73)

and so on. The k’th equation is

\[
\dot{p}_k = \lambda(k-1)p_{k-1} - (\lambda + \mu)kp_k + \mu(k+1)p_{k+1}.
\] (2.74)

As a sanity check you can verify that probability is conserved. This is like the trunking problem except that the probability of a new bug appearing in \(dt\) is proportional to the number of bugs. Notice also that extinction is an absorbing state — bugs don’t appear out of thin air. (But telephone calls do.)

Once again, the generating function is defined by

\[
G(z,t) \equiv \sum_{k=0}^{\infty} z^k p_k(t).
\] (2.75)

From (2.73), it follows that the generating function satisfies

\[
G_t - \nu(z)G_z = 0, \quad \nu \equiv (z - 1)(\lambda z - \mu).
\] (2.76)

If each realization has \(n_0\) organisms at \(t = 0\) then \(p_k(0) = \delta_{k-n_0}\) and \(G(z,0) = z^{n_0}\).

Equation (2.76) can be solved with method of characteristics. The characteristic coordinate, \(\xi(z,t)\), is the solution of

\[
\frac{dz}{dt} = \nu(z), \quad z(0) = \xi,
\] (2.77)

or

\[
\xi(z,t) = \frac{\mu(1 - e^{\nu t}) + (\mu e^{\nu t} - \lambda)z}{(\mu - \lambda e^{\nu t}) + \lambda(e^{\nu t} - 1)z}.
\] (2.78)

The general solution of (2.76) is therefore \(G(z,t) = \hat{G}(\xi)\), where \(\hat{G}\) is an arbitrary function. If there are \(n_0\) organisms at \(t = 0\) then \(G(z,0) = z^{n_0}\), and consequently \(\hat{G}(\xi) = \xi^{n_0}\). Thus, one finds that

\[
G(z,t) = \left[ \frac{\mu(1 - e^{\nu t}) + (\mu e^{\nu t} - \lambda)z}{(\mu - \lambda e^{\nu t}) + \lambda(e^{\nu t} - 1)z} \right]^{n_0}.
\] (2.79)
The probability of extinction is $p_0(t) = G(0, t)$ and so from (2.79):

$$p_0(t) = \left[ \frac{\mu - \mu e^{(\lambda - \mu)t}}{\mu - \lambda e^{(\lambda - \mu)t}} \right]^{n_0}.$$  

(2.80)

If $\lambda - \mu > 0$ then, as $t \to \infty$, the probability of extinction is $p_0 \to (\mu/\lambda)^{n_0}$. What happens if $\lambda - \mu < 0$?

In general it is difficult to find the coefficient of $z^k$ in the Taylor series expansion of (2.79). But in the special case $n_0 = 1$ one can easily expand $G(z, t)$ in (2.79) and obtain

$$p_0(t) = \mu e^{(\lambda - \mu)t - 1}, \quad p_k(t) = (1 - p_0) \left( 1 - \frac{\lambda}{\mu} p_0 \right) \left( \frac{\lambda}{\mu} p_0 \right)^{k-1}$$

if $k \geq 1$.

(2.81)

Notice that even if the growth rate, $\gamma = \lambda - \mu$, is positive the probability of ultimate extinction is nonzero i.e., $p_0(\infty) = \mu/\lambda$.

### 2.6 References specialized to pgfs

Chapter XVII of the classic:

Feller An Introduction to Probability Theory and its Applications by W. Feller

is my source. (This chapter is self-contained and can be understood without reading the preceding 16 chapters.) For lots more on generating functions see

Wi Generatingfunctionology by Herbert S. Wilf

### 2.7 Problems

**Problem 2.1.** Solve the PDE (2.19) with the initial condition $\theta(x, 0) = \exp(-x^2)$. Calculate

$$q(t) \equiv \int_{-\infty}^{\infty} \theta(x, t) \, dx .$$

**Problem 2.2.** Solve the advection-reaction equations

$$\theta_t + \theta_x = -\theta \phi, \quad \phi_t + \phi_x = \theta \phi ,$$

with the initial condition $\theta(x, 0) = \phi(x, 0) = \frac{1}{2} f(x)$. 

SIO203C, W.R. Young, March 21, 2011
Problem 2.3. Find the solution of the forced linear advection problem:

\[ \rho_t + c \rho_x = s(x,t), \quad \rho(x,0) = 0. \quad (2.82) \]

Solution This problem can be solved using the semi-linear recipe. But for a change of pace we use a slightly different approach. Inspired by (2.7), we change variables to

\[ \xi(x,t) \equiv x - ct, \quad \text{and} \quad t' \equiv t. \quad (2.83) \]

The decoration on \( t' \) reminds us that \( t' \)-derivatives are at constant \( \xi \). Using the usual rule for changing variables

\[ \partial_t = \partial_{t'} - c \partial_\xi, \quad \text{and} \quad \partial_x = \partial_\xi. \quad (2.84) \]

Thus \( \partial_t + c \partial_x = \partial_{t'} \) and the PDE in (2.82) becomes

\[ \rho_{t'} = s(\xi + c t', t'). \quad (2.85) \]

We integrate (2.85) from 0 to \( t' \):

\[ \rho(\xi, t') = \int_0^{t'} s(\xi + c t_1, t_1) \, dt_1. \quad (2.86) \]

Returning to the original variables \( x \) and \( t \), we have the solution of (2.82):

\[ \rho(x, t) = \int_0^t s[x + c (t_1 - t), t_1] \, dt_1. \quad (2.87) \]

It is instructive and essential to check by substitution that (2.87) is the solution of (2.82). To do this expeditiously we notice the general rule that

\[ (\partial_t + c \partial_x) \) (any function of \( x - ct \)) = 0. \quad (2.88) \]

Since the integrand in (2.87) is a function of \( x - ct \) the only surviving contribution is from the upper limit:

\[ (\partial_t + c \partial_x) \int_0^t s[x + c (t_1 - t), t_1] \, dt_1 = \int_0^t \left[ \partial_t \frac{x + c (t_1 - t)}{t_1} \right] \, dt_1 = 0. \quad (2.89) \]

Notice that (2.87) is the solution of (2.82) with the initial condition \( \rho(x, t) = 0 \). Using linear superposition it is now easy to solve (2.82) with the general initial condition \( \rho(x, 0) = \rho_0(x) \) (do it).

Problem 2.4. (i) Solve the PDE

\[ p_t + c p_x = \frac{1}{1 + x^2} - \alpha p, \quad q_t + c q_x = \alpha p, \]

with the initial condition \( p(x,0) = q(x,0) = 0 \). Visualize the solution using MATLAB.

SIO203C, W.R. Young, March 21, 2011
Problem 2.5. Solve the PDE:

\[ \rho_t + (e^{-x}\rho)_x = 0, \quad \rho(x, 0) = e^{-\mu^2 x^2}. \]  

(2.90)

Use MATLAB to visualize the solution (see Figure 2.4 for the answer).

Problem 2.6. Consider traffic on a highway and suppose that the speed is some function of distance, \( c(x) \). (i) Write down a PDE model, analogous to (2.6), describing the evolution of the density \( \rho(x, t) \). (ii) Beat this PDE into a form with constant coefficients by making a change of variables. Use the same tricks that worked in problem 1.11 above. The function

\[ \xi(x) = \int_0^x \frac{dx'}{c(x')}, \]

should figure prominently in your solution. (iii) Find the general solution of the initial value problem with nonuniform speed, analogous to (2.7) and (2.8). (iv) Work out the details supposing that the speed is

\[ c(x) = \frac{e^{2x} + 2e^x}{2e^{2x} + 2e^x}, \]
and that the initial condition is $\rho(x,0) = \exp(-x^2)$. To check your answer against mine show that

$$\rho(0,t) = \frac{2e^{-t}\exp[-\ln^2(\sqrt{1+3e^{-t}}-1)]}{1 + 3e^{-t} - \sqrt{1 + 3e^{-t}}}.$$ 

(v) Use MATLAB to visualise $\rho$ as a function of $x$ at selected times.

**Problem 2.7.** (i) Solve the PDE

$$\rho_t + c\rho_x = \frac{1}{1 + x^2}, \quad \rho(x, 0) = 0,$$

using the formula in [2.87]. (ii) Also solve the PDE by writing the solution as a sum of a particular and homogeneous solution, $\rho(x, t) = \rho_P(x, t) + \rho_H(x, t)$. You should be able to easily spot a particular solution. (iii) Evaluate

$$m(t) = \int_{-\infty}^{\infty} \rho(x, t) \, dx$$

the easy way. (iv) Visualize the solution using MATLAB (see figure 2.5).
Problem 2.8. Solve the PDE

\[ \rho_t - (\tanh x \rho)_x = 0, \quad \rho(x,0) = \rho_0(x). \]

Problem 2.9. (i) Show that solution of the PDE

\[ \rho_t + [c(x)\rho]_x = 0, \quad \rho(x,0) = \rho_0(x), \]

is

\[ \rho(x,t) = \rho_0(\xi) \frac{\partial \xi}{\partial x} \]

where \( \xi(x,t) \) is the characteristic coordinate defined by defined by

\[ \xi_t + c\xi_x = 0, \quad \xi(x,0) = x, \quad (2.91) \]

(ii) Show that \( \xi_x = c(\xi)/c(x) \).

Solution: (i) Let \( J \equiv \xi_x \). Taking an \( x \)-derivative of (2.91) we see that

\[ J_t + (cJ)_x = 0, \quad J(x,0) = 1. \quad (2.92) \]

Now we put \( \rho(x,t) = \rho_0(\xi)J(x,t) \) into the PDE and verify by substitution that we have a solution. Using (2.91) and (2.92) we have

\[ (\partial_t + c\partial_x)\rho_0(\xi)J(x,t) = \rho_0(\partial_t + c\partial_x)J, \]

\[ = -c_x \rho_0 J. \]

Thus \( \rho_0 J \) satisfies the PDE, and \( \rho_0 J \) also satisfies the initial condition. (ii) The solution of (2.91) can be written implicitly as

\[ s(x) - s(\xi) = t \quad \text{where} \quad s(x) \equiv \int_{0}^{x} \frac{dx_1}{c(x_1)}. \]

Differentiating this implicit solution with respect to \( x \) gives \( J = c(\xi)/c(x) \).

Problem 2.10. Consider the trunking problem and suppose that at \( t = 0 \) exactly one line is busy. (i) Find the probability generating function by solving (2.59) with the appropriate intial condition. (ii) Use your solution to obtain \( p_0(t) \) and \( p_1(t) \).

Problem 2.11. Solve the system

\[ \dot{p}_1 = -p_1, \quad \dot{p}_n = (n-2)p_{n-2} - np_n, \]

with the initial condition \( p_n(0) = \delta_{1,n} \).
Hint: show that the generating function satisfies
\[ G_t = (z^3 - z)G_z. \]

Problem 2.12. Consider the birth-death process in (2.73). Define
\[ \alpha_0(t) \equiv \sum_{k=0}^{\infty} p_k, \quad \alpha_1 \equiv \sum_{k=0}^{\infty} kp_k, \quad \alpha_2 \equiv \sum_{k=0}^{\infty} k(k-1)p_k. \]  
(2.93)
Show from (2.73) that
\[ \dot{\alpha}_0 = 0, \quad \dot{\alpha}_1 = \gamma \alpha_1, \quad \dot{\alpha}_2 = 2\lambda \alpha_1 + 2\gamma \alpha_2, \]  
(2.94)
where \( \gamma \equiv \lambda - \mu \) is the growth rate. Interpret these results probabilistically. In particular show that if \( \gamma \leq 0 \) then eventually fluctuations in the size of the population must become comparable to the average population.

Problem 2.13. Fill in all the steps between (2.77) and (2.78).

Problem 2.14. Consider the special case \( \lambda = \mu \) in (2.76). (i) Start from scratch and solve the PDE (2.76) using the initial condition which is appropriate if \( n_0 = 1 \). (ii) Show that
\[ p_0(t) = \frac{\mu t}{1 + \mu t}, \quad p_k(t) = \frac{(\mu t)^{k-1}}{(1 + \mu t)^{k+1}} \text{ if } k \geq 1. \]  
(2.95)

Problem 2.15. A power plant supplies \( m \) machines which use the current intermittently. A machine switches off with a probability \( \mu \) per unit time and switches on with a probability \( \lambda \) per unit time. All the machines operate independently of each other. (i) Construct the coupled ODE’s describing the evolution of
\[ p_k(t) = \text{Probability of } k \text{ machines drawing power at } t, \]  
(2.96)
where \( k = 0, 1, \cdots, m \). (ii) Find the PDE satisfied by
\[ G(z, t) = \sum_{k=0}^{m} p_k(t)z^k. \]  
(2.97)
(iii) Find the solution of this PDE with the initial condition \( p_0(0) = 1 \). (iv) Find the steady asymptotic \( (t \to \infty) \) solution.
Problem 2.16. Consider the trunking problem and suppose that the probability of initiation of a new call is
\[ \lambda(t) = e^{-\mu t} \lambda_0. \] (2.98)
The probability of termination is still the constant \( \mu \). Supposing that \( p_0(0) = 1 \), find \( p_n(t) \).

Problem 2.17. Solve the PDE:
\[ \rho_t + c \rho_x = \alpha(1 - \rho), \quad \rho(x,0) = 0, \quad \rho(0,t) = 0, \]
in the domain \( x > 0 \) and \( t > 0 \). \( \rho(x,t) \) might be the signalling of fluid which is being pumped with speed \( c \) through a semi-infinite pipe \( (x > 0) \). The pipe is heated uniformly, and fluid enters at \( x = 0 \) with \( \rho = 0 \). Discuss both \( c > 0 \) and \( c < 0 \). Sketch the solution at \( \alpha t = 1 \).

Problem 2.18. Give a general formula — analogous to (2.87) — for the solution of the PDE
\[ \rho_t + c \rho_x = s(x,t), \quad \rho(x,0) = 0, \quad \rho(0,t) = 0, \]
in the domain \( x > 0 \) and \( t > 0 \).

Problem 2.19. Consider a highway, \( x > 0 \), on which the speed is \( c = c_0 \exp(-\alpha x) \). Suppose there are no cars on the highway at \( t = 0 \) and then cars enter the highway at \( x = 0 \) at a constant rate, \( R \). (i) Formulate a PDE description of this problem. (ii) Check your formulation by showing that the number of cars on the highway at time \( t > 0 \) is \( Rt \). (iii) Find the position of the car which leaves \( x = 0 \) at \( t = \tau \). (iv) Solve the PDE and determine \( \rho(x,t) \).

Problem 2.20. Consider a population in which the death rate is a decreasing function of time. As an illustrative model consider:
\[ h_t + h_a = -\frac{\mu h}{1 + \alpha t}. \]
Suppose the age structure is in equilibrium at \( t = 0 \): \( h(a,0) = b_0 \exp(-\mu a) \). Notice that both \( \mu \) and \( \alpha \) have dimensions time\(^{-1}\), so \( \nu \equiv \mu / \alpha \) is dimensionless. The constant \( \nu \) appears prominently in the solution.

(i) How must the birth-rate, \( b(t) \), decrease so that the total size of the population remains constant? (ii) Calculate the average age, \( \bar{a}(t) \), of the population in this situation. Hint: you should answer (i) and (ii) without solving a PDE. (iii) Solve the PDE and obtain the age structure of the population.
Problem 2.21. Consider a population whose death rate depends only on age, \( a \). A cohort of babies who leave the maternity ward together with \( a = 0 \). Or a box of light bulbs leaves the factory. Or a bunch of \(^{14}\text{C}\) atoms are created in the upper atmosphere by a cosmic ray shower. The fate of this cohort is determined by solving

\[
h_t + h_a = -\mu(a)h, \quad h(a, 0) = \delta(a).
\]

(i) Solve the PDE above. The survival function

\[
S(a) \equiv \exp\left(-\int_0^a \mu(a') \, da'\right)
\]

should appear prominently in your solution. (ii) Show that the survival function has the interpretation

\[
S(a) = \text{fraction of the cohort still alive at age } a.
\]

(iii) Let \( P(a) \) be the probability density function of life spans e.g., if \( a \) is the age of death

\[
\text{probability that } a_1 < a < a_2 = \int_{a_1}^{a_2} P(a') \, da'.
\]

Find the (simple) relation between \( S(a) \) and \( P(a) \) i.e., given \( S(a) \) how do you calculate \( P(a) \)? (iv) Check your answer by showing that your expression for \( P \) in terms of \( S \) is normalized:

\[
1 = \int_0^\infty P(a) \, da.
\]

(v) As another check on your answer to (iii), show that the average life-span, defined by

\[
\tau = \int_0^\infty a P(a) \, da,
\]

is given in terms of \( S(a) \) by

\[
\tau = \int_0^\infty S(a) \, da.
\]

(vi) Show that if \( \mu \) is independent of \( a \) then \( \bar{a} = \tau \), where \( \bar{a} \) is the average age of the equilibrium population in (2.42). (vii) Find an example which shows that in general the average life-span, \( \tau \), is not equal to the average age of the equilibrium population, \( \bar{a} \). Discuss this “renewal paradox”
e.g., for a population with high infant mortality do you expect $\tau > \bar{a}$ or $\bar{a} > \tau$? Alternatively, suppose because of effective medical care and social planning everyone lives till exactly $a = 100$ and then drops dead. What is the relation between $\tau$ and $\bar{a}$ in this case?

**Problem 2.22.** Construct a model for the growth of hair. Assume that a head has $N \gg 1$ follicles and each follicle extrudes hair at a constant rate, say $c$ centimeters per year. Further, each hair has a constant probability $\alpha$ per unit time of dropping out at the root; when this happens, the follicle keeps extruding so the hair is replaced. Let $h(\ell, t)$ be the density function i.e., $h(\ell, t) d\ell$ is the number of hairs of length $(\ell, \ell + d\ell)$ at time $t$ and therefore

$$N = \int_0^\infty h(\ell, t) \, d\ell.$$  

(i) Given $h(\ell, t)$, how does one calculate the total length of hair? (ii) Write down the PDE satisfied by $h(\ell, t)$ and find the large time solution.

**Problem 2.23.** Change the assumptions of the hair-growth model so that the probability per unit time of a hair dropping out at the root is $\beta \ell$ i.e., longer hairs are more fragile. Find the steady-state density $h(\ell, \infty)$ with this new assumption.

**Problem 2.24.** Yet more hair growth — and this one may be difficult. This time suppose that each hair has a constant probability per unit length $\gamma$ of breaking somewhere along its length. The break-point is picked with uniform probability along the length, and after the break you're left with a shorter hair that keeps growing at speed $c$. Find the equation satisfied by $h(\ell, t)$, and solve it to obtain the steady-state density of hair length.

**Problem 2.25.** Consider a solution of photosensitive molecules occupying the region $x > 0$, and denote the density of molecules (i.e., molecules per cubic meter) by $n(x, t)$. The solution is irradiated by a beam of photons, with density $\rho(x, t)$ photons per cubic meter. A photon is absorbed by a molecule, which then decomposes so that the solution becomes more transparent to light. The model is

$$\rho_t + c\rho_x = -\alpha n \rho, \quad (2.99)$$

$$n_t = -\alpha n \rho, \quad (2.100)$$

where $c$ is the speed of light. The initial conditions are

$$n(x, 0) = n_0 \quad \rho(x, 0) = 0, \quad (2.101)$$
and boundary condition is $\rho(0, t) = \rho_0$ i.e., starting at $t = 0$, $I_0 = c\rho_0$ photons enter the solution at $x = 0$. Determine the non-dimensional parameters in this problem and identify a limit in which the term $\rho_t$ can be neglected i.e., the radiation is “quasi-static”. Solve the problem in this limit by finding analytic expressions for $\rho(x, t)$ and $n(x, t)$. 

---

*SIO203C, W.R. Young, March 21, 2011*
Lecture 3

Quasilinear equations

We are now going to develop the method of characteristics for quasilinear equations

\[ a(x, y, z)z_x + b(x, y, z)z_y = c(x, y, z). \] (3.1)

in detail. Our earlier “semilinear” examples are special cases of (3.1).

We follow Lagrange and view the problem geometrically. The solution, \( z = z(x, y) \), of (3.1) is a surface in the three-dimensional space \((x, y, z)\). The normal to this solution surface is the vector

\[ n = (-z_x, -z_y, 1). \] (3.2)

The quasilinear equation (3.1) is equivalent to

\[ \mathbf{a} \cdot \mathbf{n} = 0, \] (3.3)

where \( \mathbf{a} \equiv (a, b, c) \) is a vector field associated with (3.1). In other words, \( \mathbf{a} \) is normal to \( \mathbf{n} \) and hence \( \mathbf{a} \) lies in the plane tangent to the solution surface.

In this lecture we denote the independent variable by \( z \) to emphasize that the solution surface of (3.1) is in the three-dimensional space \((x, y, z)\).

The characteristic curves (ccs) are everywhere tangent to vector \( \mathbf{a} \). In fluid mechanics \( \mathbf{a} \) is a velocity field and a cc is a streamline. Notice that ccs live in the space \((x, y, z)\). The projections of these three-dimensional ccs down onto the \((x, y)\)-plane are the characteristic base
Lecture 3. Quasilinear equations

Figure 3.1: The totality of characteristics passing through the data curve is the solution surface. characteristics.eps

curves (CBCs). Sometimes, confusingly, people also refer to the characteristic base curves as “characteristics”—I did this myself in the first lecture. I’ll sincerely try to maintain the distinction between Ccs and CBCs in this lecture.

The other important curve associated with the PDE (3.1) is the data curve (DC). The DC is determined by the initial or boundary conditions and also lives in the three-dimensional space \((x,y,z)\). For example, if one requires the solution of (3.1) with the initial condition that \(z(0,y) = 2y\) then the data curve is a straight line determined by the intersection of the two planes \(x = 0\) and \(z = 2y\). Just as the characteristic base curves lie in the \((x,y)\) plane, there is also a data base curve (DBC). Since it is much easier to sketch in two-dimensions than in three, we frequently visualize PDE problems by sketching the CBCs and the DBC.

I have attempted to illustrate Lagrange’s geometric vision for solving PDE’s in figure 3.1. In the streamline analogy, you can imagine that the DC is source of dye which colours each passing streamline. The painted surface is the solution surface of (3.1).

3.1 Characteristics and their properties

On a CC the PDE reduces to an ODE. Points on the Ccs are \([x(t),y(t),z(t)]\) where \(x\), \(y\) and \(z\) satisfy Lagrange’s equations

\[
\frac{dx}{dt} = a, \quad \frac{dy}{dt} = b, \quad \frac{dz}{dt} = c.
\]

Lagrange’s equations

The equations above ensure that the vector field \(\mathbf{a} = (a,b,c)\) is everywhere tangent to a CC.
Lecture 3. Quasilinear equations

Every surface consisting wholly of cc$s is a solution surface. Conversely, every solution surface, \( z = z(x, y) \), is also made wholly of cc$s. To see this, suppose we that we move around on the solution surface by solving

\[
\frac{dx}{dt} = a[x, y, z(x, y)], \quad \frac{dy}{dt} = b[x, y, z(x, y)]. \tag{3.5}
\]

Computing \( \frac{dz}{dt} \) we have

\[
\frac{dz}{dt} = z_x \frac{dx}{dt} + z_y \frac{dy}{dt}, \quad \text{(chain rule)}
\]

\[
= az_x + bz_y, \quad \text{from (3.5)},
\]

\[
= c, \quad \text{from (3.1)}. \tag{3.6}
\]

Thus we recover Lagrange’s equations, (3.4).

The cc$s are a two parameter family of curves. One parameter is \( t \) in (3.4) and the second parameter, \( s \), labels the initial condition: \( x = x(s, t), \quad y = y(s, t) \) and \( z = z(s, t) \). Elimination of \( s \) and \( t \) produces the solution surface \( z = z(x, y) \). Thus the geometric vision is that the solution surface is generated by a family of cc$s parametrized by \( s \).

A quasilinear recipe

We automate the solution of the quasilinear PDE (3.1) with the recipe:

- parameterize the DC: \( x = x(s, 0), \quad y = y(s, 0) \) and \( z = z(s, 0) \);
- solve Lagrange's equations (3.4) using the initial condition from above;
- eliminate the parameters \( s \) and \( t \) and get \( z = z(x, y) \);
- check by substitution.

This recipe is a formulaic version of the total-derivative procedure used in the first lecture to reduce PDEs to ODEs. The parameter \( t \) we introduce in Lagrange’s equations is just a way of treating \( x \) and \( y \) evenhandedly. Let's use the recipe in an example.
Lecture 3. Quasilinear equations

Example: Solve the quasilinear PDE:

\[ z_x + z_y = z, \quad z(0, y) = y^2. \]  

We parameterize the DC by writing

\[ t = 0: \quad x = 0, \quad y = s, \quad z = s^2. \]  

As we vary \( s \) we move along the DC. Next, we solve Lagrange’s equations with the initial condition in (3.8)

\[ \begin{align*}
\frac{dx}{dt} &= 1, \quad \Rightarrow \quad x = t; \\
\frac{dy}{dt} &= 1, \quad \Rightarrow \quad y = t + s; \\
\frac{dz}{dt} &= z, \quad \Rightarrow \quad z = s^2e^t.
\end{align*} \]  

(3.9)

The third step is to eliminate the parameters \( s \) and \( t \):

\[ \begin{align*}
t &= x, \quad s = y - x.
\end{align*} \]  

(3.10)

Eliminating \( s \) and \( t \) in the third equation in (3.7) gives

\[ u = (y - x)^2 \exp(x). \]

The final step is to check by substitution.

Example: In lecture 1 our approach was somewhat clumsier than the quasilinear recipe. Let’s illustrate the differences by discussing the particular example

\[ (x^2 + y^2) z_x - 2xyz_y = 0. \]  

(3.11)

If we divide by \( x^2 + y^2 \) then the left hand side can be regarded as a total derivative:

\[ \frac{dz}{dx} = z_x + \frac{dy}{dx} z_y = 0, \]  

(3.12)

where

\[ \frac{dy}{dx} = -\frac{2xyz}{x^2 + y^2}. \]  

(3.13)

To find the CCS we have to integrate the ODE above. We can do this using tricks from chapter 1 of BO — the ODE is scale invariant i.e., unchanged by \( x \rightarrow ax \) and \( y \rightarrow ay \).

In (3.13) are treating \( x \) and \( y \) differently i.e., we’re regarding \( y \) as a function of \( x \) on the CCS, rather than the reverse. This might lead to issues if the CCS bend around so that there are several values of \( y \) for one value of \( x \). And the ODE in (3.13) has a singularity at the origin which is not in the original PDE.
The alternative approach is to look for a parameterized solution for the ccs. This is the second step of the quasilinear recipe. This alternative route presents us with Lagrange’s equations

\[
\frac{dx}{dt} = x^2 + y^2, \quad \frac{dy}{dt} = -2xy, \quad \frac{dz}{dt} = 0. \quad (3.14)
\]

Of course, if we divide \( dy/dt \) by \( dx/dt \) then we return to (3.13). But another way to solve (3.14), and thus determine the ccs, is to use the new variables \( a = x + y \) and \( b = x - y \):

\[
\frac{da}{dt} = b^2, \quad \frac{db}{dt} = a^2, \quad \Rightarrow \quad \frac{d}{dt} \left( a^3 - b^3 \right) = 0. \quad (3.15)
\]

Hence \( y^3 + 3x^2y \) is constant on characteristics, and the solution is that

\[
z = \text{an arbitrary function of } y^3 + 3x^2y \quad \blacksquare \quad (3.16)
\]

**Example:** Solve

\[
(x^2 + y^2)z_x - 2xyz_y = 0. \quad (3.17)
\]

with \( z(x, x) = \exp(x) \). The general solution is \( z(x, y) = a(y^3 + 3x^2y) \) and the arbitrary function \( a \) is determined by applying the data:

\[
e^x = a \left( 4x^3 \right), \quad \Rightarrow \quad a(z) = e^{(z/4)^{1/3}},
\]

where the one-third power above is defined as \( z^{1/3} = \text{sgn}(z)|z|^{1/3} \).

**More dimensions**

The generalization to dimension \( d \) is clear. The quasilinear PDE is

\[
\sum_{n=1}^{d-1} a_n(x) \frac{\partial z}{\partial x_n} = f(x), \quad (3.18)
\]

where a point in the \( d \)-dimensional solution space is

\[
x = (x_1, x_2, \cdots, x_{d-1}, z). \quad (3.19)
\]

Lagrange’s equations are

\[
\frac{dx_n}{dt} = a_n, \quad (n = 1, \cdots, d - 1), \quad (3.20)
\]

and

\[
\frac{dz}{dt} = f. \quad (3.21)
\]

As initial conditions, we have to supply data on a \((d - 2)\)-dimensional surface, which can be parameterized with \( s_1, s_2 \cdots s_{d-2} \). The general solution will be an arbitrary function with \( d - 2 \) variables.
Lecture 3. Quasilinear equations

**Example:** Find the general solution of the PDE:

\[ xyu_x + zu_y + uz = 0. \]  

(3.22)

In this example the solution surface lives in the four-dimensional space \((x, y, z, u)\). Lagrange’s equations are

\[
\frac{dx}{dt} = xy, \quad \frac{dy}{dt} = z, \quad \frac{dz}{dt} = 1, \quad \frac{du}{dt} = 0. \]  

(3.23)

In this problem \(d = 4\), and we seek a general solution in the form

\[ z = \text{an arbitrary function of two variables}. \]  

(3.24)

The two variables above correspond to the two-dimensional data surface, and this surface is carried into the four-dimensional space \((x, y, z, u)\) by the CCS. Thus the solution “surface” is three dimensional. Because \(u\) is constant on a CC, the two arbitrary functions above are also constant on a CC.

You might be tempted to begin your assault on Lagrange’s equations by dividing the first two equations above to obtain

\[
\frac{dx}{x} = \frac{ydy}{z}. \]  

(3.25)

But don’t treat \(z\) as a constant! This is a dumb mistake because \(x, y, z\) and \(u\) all vary with \(t\) along a characteristic. Thus one cannot integrate the equation above.

Instead notice that Lagrange’s equations tell us that

\[
\frac{d}{dt} \left( y - \frac{1}{2}z^2 \right) = 0, \]  

(3.26)

so that

\[ y - \frac{1}{2}z^2 = A, \]  

(3.27)

where \(A\) is constant on a CC. Thus

\[
\frac{dx}{dz} = xy = x \left( A + \frac{1}{2}z^2 \right), \]  

(3.28)

or

\[ \ln x = B + Az + \frac{1}{6}z^3. \]  

(3.29)

The constants of integration \(A\) and \(B\) are both constant along CCS. We can eliminate \(A\) between (3.27) and (3.29) and obtain

\[ \ln x = B + yz - \frac{1}{3}z^3. \]  

(3.30)

At this point (3.27) and (3.30) define two function \(A(x, y, z)\) and \(B(x, y, z)\) which are both solutions of the PDE. Thus the general solution of the PDE is

\[ u = a \left( y - \frac{1}{2}z^2, \ln x - yz + \frac{1}{2}z^3 \right), \]  

(3.31)

where \(a\) is arbitrary.

\[ \text{SIO203C, W.R. Young, March 21, 2011} \]
**Example:** Find the solution of the PDE:

$$x y u_x + z u_y + u_z = z,$$

(3.32)

with the data $u(x, y, 0) = 0$.

### 3.2 More theory

Does the quasilinear recipe always work? No, because sometimes the problem has no solution and other times there is an infinitude of solutions. Let’s recall a simple example from lecture 1: find the function $z(x, y)$ satisfying the PDE:

$$z_x = 0, \quad z(x, 0) = x.$$

(3.33)

It is obvious that this equation doesn’t have a solution: differentiating the data with respect to $x$, we discover that $z_x = 1$. Thus the DC contradicts the equation.

In (3.33) the DBC coincides with a CBC. But the DC is not a CC: only the base curves coincide and the problem is inconsistently posed. To lift the solution away from the DC the vector field $a$ in (3.3) must pass through DC.

If you apply the quasilinear recipe to the simple example above then the first place you get into trouble is in trying to solve for $x$ and $y$ in terms of $t$ and $s$. It can happen that the mapping from the $(s, t)$ plane to the $(x, y)$ plane is not invertible. The condition for invertibility is that

$$J \equiv \frac{\partial (x, y)}{\partial (s, t)} = x_t y_s - x_s y_t \neq 0.$$

(3.34)

$J$ is the *Jacobian* of the mapping $(s, t) \rightarrow (x, y)$. But the Lagrange’s equations tell us that

$$J = bx_s - ay_s.$$

(3.35)

Hence, in order to use the recipe, we must have $bx_s - ay_s \neq 0$ on the DC.

The condition $J \neq 0$ is equivalent to requiring that the data base curve is nowhere tangent to a characteristic base curve. In the example surrounding (3.33) this condition is strongly violated because the data base curve is a characteristic base curves: the two curves are tangent
Lecture 3. Quasilinear equations

everywhere! More subtle problems occur if there are isolated points on the data base curve at which \( J = 0 \) i.e., local points of tangency. This pathology is exhibited by the PDE in problem 3.6 \( J = 0 \) at the isolated point \((x, y) = (0, 0)\) on the DBC.

It all starts to seem clear: if \( J \neq 0 \) on DC then we can start on the DC and confidently step into the wild blue yonder. By this I mean that if \( J \neq 0 \) on DC then the PDE has a solution in the neighbourhood of the DC. In other words, the solution exists locally. But if \( J = 0 \) somewhere on the DBC then we anticipate problems. For instance, the PDE in problem 3.6 has no solution at some points which are arbitrarily close to the DBC (see right panel of figure 3.5). And in the example (3.33) there is no solution at all. But, just to show that there are still surprises, we turn to yet another example.

**Example** Consider the PDE:

\[
3(z - y)^2 z_x - z_y = 0, \quad z(0, y) = y.
\]  

(3.36)

We parameterize the DC with

\[
t = 0: \quad x = 0, \quad y = s, \quad z = s.
\]  

(3.37)

Before pressing on with the recipe, we calculate \( J \) on the DC

\[
J = bx_s - ay_s = (-1) \times 0 - \underbrace{3(z - y)^2}_{\text{zero on the DC}} \times 1 = 0.
\]  

(3.38)

We should be in trouble: this is as bad as it gets because \( J \) is zero everywhere on the DC. But blithely ignoring the red light in (3.38) and continuing with the recipe one very quickly finds that

\[
z = y + x^{1/3}
\]  

(3.39)

is a solution of (3.36). In other words the recipe works, even though \( J = 0 \) everywhere on the DC! This example shows that \( J = 0 \) is necessary, but not sufficient, condition for nonexistence of a solution in the neighbourhood of the DC.

The point of the example above is that our conclusion following (3.35) assumes that \( z(x, y) \) has continuous first derivatives on the DC, and the solution in (3.39) does not. In this example the CCBs are tangent to the DBC everywhere. But still the CC’s manage to leave the DC and make a solution surface. The DBC is a caustic or a envelope of the CBC. Thus even if \( J = 0 \) on the DC, then we can still solve the PDE provided the DBC is a caustic of CBCs.
### 3.3 The nonlinear advection equation

We turn now to a detailed study of an important example quasilinear equation. This is the nonlinear advection equation

\[ u_t + uu_x = 0 \quad \text{or} \quad u_t + \left( \frac{1}{2} u^2 \right)_x = 0. \]  

(3.40)

The final expression in (3.40) shows that the nonlinear advection equation can be written as a conservation equation for the density \( u \), and the flux \( u^2/2 \).

Comparing the linear advection equation

\[ u_t + cu_x = 0 \]  

(3.41)

with nonlinear advection equation (3.40) we see that in the nonlinear case the speed of the disturbance is \( c = u \). Thus, loosely speaking, we anticipate that in the nonlinear case bigger disturbances will travel more rapidly.

We use the method of characteristics to solve (3.40) with the initial condition that

\[ u(x, 0) = F(x), \]  

(3.42)

where \(-\infty < x < \infty\). My discussion follows Whitham, though I am solely responsible for the myrmecophilia.

Imagine an ant moving along some curve \( x = x(t) \) in the \((x,t)\)-plane. The solution of (3.40), \( u(x,t) \), can be visualized as a surface lying above that plane. The value of \( u(x,t) \) observed by the moving ant is obtained from the “total derivative” of \( u(x,t) \), namely

\[ \frac{d}{dt} u(x(t),t) = u_t(x(t),t) + \frac{dx}{dt} u_x(x(t),t). \]  

(3.43)

We now consider a mathematically inclined ant, starting at \( x(0) = \xi \), who adjusts her trajectory so that

\[ \frac{dx}{dt} = u(x,t), \quad x(0) = \xi. \]  

(3.44)

Because \( u(x,t) \) satisfies the advection equation (3.40) this ant observes that

\[ \frac{d}{dt} u(x(t),t) = 0. \]  

(3.45)

AKA the “advective derivative”

On the characteristics pde (3.40) reduces to ode (3.45).
In other words, on a curve determined by \( (3.44) \), \( u(x,t) \) is constant. In fact, since \( u \) satisfies the initial condition \( (3.42) \), the constant value of \( u \) is just
\[
    u(x,t) = F(\xi). \tag{3.46}
\]

In other words, the initial value of \( u \) is \( F(\xi) \) and if ant picks her trajectory according to \( (3.44) \) then \( u \) doesn’t change.

Now that we realize \( u \) is constant on the trajectory, it is trivial to determine that trajectory by integrating \( (3.44) \):
\[
    x = \xi + ut. \tag{3.47}
\]

Finally, eliminate \( \xi = x - ut \) between \( (3.46) \) and \( (3.47) \) to get \( u \) as a function of \( x \) and \( t \). This gives
\[
    u = F(x - ut). \tag{3.48}
\]

Given \( F \) we can, in principle, solve the equation above for \( u(x,t) \) (examples follow). After sorting out some notational distractions, you can also obtain this solution using the quasilinear solution earlier in this lecture.

It is instructive to check by substitution that \( (3.48) \) solves the PDE \( (3.40) \). Taking an \( x \)-derivative of \( (3.48) \) we get:
\[
    u_x = (1 - u_x t)F'(x - ut), \quad \Rightarrow \quad u_x = \frac{F'(\xi)}{1 + tF'(\xi)}. \tag{3.49}
\]

And the \( t \)-derivative of \( (3.48) \) is
\[
    u_t = -(tu)_t F'(x - ut), \quad \Rightarrow \quad u_t = -\frac{uF'(\xi)}{1 + tF'(\xi)} = -uu_x. \tag{3.50}
\]

### A kinky initial condition

Suppose the initial condition is
\[
    u = \begin{cases} 
        1 - |x|, & \text{if } |x| < 1, \\
        0, & \text{if } |x| > 1.
    \end{cases} \tag{3.51}
\]

The solution is shown in figure 3.2 Big values of \( u \) overtake the small values of \( u \); this is **nonlinear steepening**.

To obtain the solution we can deal with the algebraic result in \( (3.48) \) or we can use the geometric construction:
Figure 3.2: The top panel shows the characteristics with the piecewise linear initial condition, $F$ in (3.51). The bottom panel shows the steepening wave. The characteristics first cross, and the solution becomes multivalued, at $(x, t) = (1, 1)$. kink.eps

- Draw the graph of the initial condition, $u = F(x)$, in the $(x, u)$-plane;
- Take each point on the initial curve and slide it sideways a distance $F(x)t$;
- the resulting curve is the graph of the solution $u(x, t)$.

As you see in figure 3.2, this construction gives rise to a multivalued solution when $t > 1$. There is a region of the $(x, t)$-plane in which there are three values of $u$.

A witchy initial condition

As an another example, suppose that we want to solve (3.40) with the initial condition

$$F(x) = \frac{1}{1 + x^2}.$$  \hspace{1cm} (3.52)

This initial condition is the famous Witch of Agnesi. In this case, from (3.47), the characteristics comprise the family of lines

$$x = \xi + \frac{t}{1 + \xi^2}.$$  \hspace{1cm} (3.53)
Lecture 3. Quasilinear equations

Figure 3.3: The left panel shows the characteristics obtained from (3.53). The right panel shows the steepening wave, calculated using (3.55). The plots are at \( t = [0 \ 1/2 \ 1 \ 2 \ 3] t_s \), where \( t_s = 8\sqrt{3}/9 \) is the shock time.

Using MATLAB we visualize this family by specifying \( \xi \) in (3.53) and then plotting \( x \) versus \( t \) (see figure 3.3). We see that characteristics intersect, so that there are some points in the \((x, t)\)-plane at which there are three values of \( u \).

To determine \( u(x, t) \) we must now eliminate \( \xi \) from (3.46) and (3.47):

\[
u = \frac{1}{1 + \xi^2}, \quad x = \xi + ut \quad \Rightarrow \quad u = \frac{1}{1 + (x - ut)^2}.
\]

(3.54)

Equation (3.54) is a cubic which defines \( u \) as an implicit function of \( x \) and \( t \). We can visualize the solution without solving this cubic by specifying \( u \) and \( t \) then solving (3.54) for \( x \) in terms of \( u \) and \( t \):

\[
x_{\pm} = ut \pm \sqrt{1 - \frac{u}{u}}.
\]

(3.55)

In figure 3.3 we show \( u \) as a function of \( x \) at fixed times calculated from (3.55) and plotted with MATLAB. The shock first rears its ugly head at \( t_s = 8\sqrt{3}/9 \approx 1.54 \) when there is a point on the forward face of the pulse at which the slope \( u_x \) is infinite (and negative).

Here is the MATLAB script which produces figure 3.3:

```matlab
close all
c1c
```

SIO203C, W.R. Young, March 21, 2011
3.4 Shocks, caustics and multivalued solutions

How did we determine the shock time \( t_s = 8\sqrt{3}/9 \) in the example of figure 3.3? When \( t < t_s \), \( u(x,t) \) is a single-valued function of \( x \) and the derivative \( u_x \) is finite everywhere. But because of nonlinear steepening, \( u_x \) becomes infinitely negative at a finite time, known as the shock-time and denoted \( t_s \). When \( t > t_s \) the solution is multivalued and there are two locations at which \( u_x = \infty \) — see the right hand panel of figure 3.3.

To determine \( t_s \) we could go back to our earlier expression for \( u_x \) in (3.49) and figure out when and where the infinity first appears. However there is an alternative route which lets us admire some different scenery. Let \( v(x,t) \equiv u_x(x,t) \). Differentiating the advection equation (3.40) we have

\[
v_t + v^2 + uv_x = 0. \tag{3.56}
\]
Lecture 3. Quasilinear equations

Now we notice that \((10.3)\) evaluated on a characteristic curve is
\[
\frac{dv}{dt} = -v^2, \quad \Rightarrow \quad v(\xi, t) = \frac{F'(\xi)}{1 + F'(\xi)t}.
\] (3.57)

We see that if \(F'(\xi) < 0\) then \(v = u_x\) will become infinite in a finite time. The singularity first forms on the characteristic \(\xi_s\) which originates at the point of most-negative slope and so
\[
t_s = \min_{\forall \xi} \left[ -1/F' (\xi) \right] = -1/F'(\xi_s).
\] (3.58)

I like this alternative route because it is obvious from the ODE (3.57) that the slope \(v = u_x\) is monotonically decreasing as we move along each characteristic. Also one can calculate \(t_s\) before solving the PDE.

After the shock forms \(u(x, t)\) is a multivalued function: when \(t > t_s\) there are three values of \(u\) at a single location \(x\). The function \(u(x, t)\), viewed as a surface above the \((x, t)\) plane, is folded and the shock location \((x_s, t_s)\) is the "point" of the fold. There are two creases (or caustics) which originate at \((x_s, t_s)\). The caustic curves, \(x = x_c(t)\), are located by the condition that \(u_x = \infty\), or from (3.49)
\[
1 + tF' (\xi_c) = 0.
\] (3.59)

We have to solve the equation above, together with \(x_c = \xi_c + F(\xi_c)t\), to determine the caustic location (example below).

In many cases the appearence of a multivalued solution at \((x_s, t_s)\) indicates that the PDE model is failing. For example, if \(u(x, t)\) is traffic density it doesn’t make sense that there are three different values of \(u\) at the same location. In later lectures we discuss how the “solution” of a PDE can be extended using physical arguments to describe evolution once \(t > t_s\).

The witch again

In the witchy example:
\[
F = \frac{1}{1 + \xi^2}, \quad F' (\xi) = -2\xi F^2, \quad F'' (\xi) = 2F^2 \left(4\xi^2 F - 1\right).
\] (3.60)

The most negative value of \(F'\) is at \(\xi_s\), where \(F'' (\xi_s) = 0\); a short calculation using \(F''\) in (3.60) then gives \(\xi_s = 1/\sqrt{3}\). Thus
\[
F'(\xi_s) = -3\sqrt{3}/8, \quad \text{and from (3.58):} \quad t_s = 8\sqrt{3}/9.
\] (3.61)
Lecture 3. Quasilinear equations

Can you find the location, \( x_s \), at which the shock forms (see problem 3.18)?

Let’s finally complete this witch example by finding the location of the caustics. We have to find the curves in the \((x,t)\) plane along which \( u_x = \infty \). In the case of the witch, this leads to the system

\[
x_c = \xi_c + \frac{t}{1 + \xi_c^2}, \quad \frac{2\xi_c}{(1 + \xi_c^2)^2} = \frac{1}{t}.
\]

The second equation is the condition that \( u_x = \infty \), or equivalently \( F'(\xi_c) = -1/t \). We must eliminate \( \xi_c \) and find \( x_c \) as a function of \( t \). This can’t be done algebraically. But, once again, it is easy to specify \( \xi_c \), then find \( t \) and finally obtain \( x_c \). Here’s the MATLAB script which produces figure 3.4:

```matlab
close all
clecl
xis=1/(sqrt(3)); %% The xi of the shock.
xi=linspace(0.2*xis,4*xis); %% Specify xi
t=(1+xi.^2).^2./(2*xi); %% find t
xc=xi+t./(1+xi.^2); %% find x_caustic
```

Can you find the location, \( x_s \), at which the shock forms (see problem 3.18)?

Let’s finally complete this witch example by finding the location of the caustics. We have to find the curves in the \((x,t)\) plane along which \( u_x = \infty \). In the case of the witch, this leads to the system

\[
x_c = \xi_c + \frac{t}{1 + \xi_c^2}, \quad \frac{2\xi_c}{(1 + \xi_c^2)^2} = \frac{1}{t}.
\]

The second equation is the condition that \( u_x = \infty \), or equivalently \( F'(\xi_c) = -1/t \). We must eliminate \( \xi_c \) and find \( x_c \) as a function of \( t \). This can’t be done algebraically. But, once again, it is easy to specify \( \xi_c \), then find \( t \) and finally obtain \( x_c \). Here’s the MATLAB script which produces figure 3.4:

```matlab
close all
clecl
xis=1/(sqrt(3)); %% The xi of the shock.
xi=linspace(0.2*xis,4*xis); %% Specify xi
t=(1+xi.^2).^2./(2*xi); %% find t
xc=xi+t./(1+xi.^2); %% find x_caustic
```
Lecture 3. **Quasilinear equations**

```matlab
plot(xc,t,'linewidth',2)
axis([0 4 0 5])
hold on
%%% Now overlay some characteristics
xo=linspace(-2*xis,10*xis,40);
t=linspace(0,5);
for xx=xo
    plot(xx+t./(1+xx.^2),t);
    hold on
end
xlabel('x')
ylabel('t')
```

### 3.5 Problems

**Problem 3.1.** Solve PDE:

\[ u_x + 2x u_y = xy, \quad u(0, y) = 0. \]

**Solution.** Applying the recipe, we find \( u(x, y) = \frac{1}{2}yx^2 - \frac{1}{3}x^4. \)

**Problem 3.2.** Is the curve \( \Gamma: x = s, \quad y = s, \quad u = -\frac{1}{1+s}, \) a CC of the PDE \( u_x + u_y = u^2? \) If it is not, solve the initial value problem; if \( \Gamma \) is a CC, find *all* solutions of the PDE that pass through \( \Gamma. \)

**Problem 3.3.** *(i)* Show that the solution of

\[ yu_x - 2xyu_y = 2xu, \quad u(0, 1 \leq y \leq 2) = y^3, \]

is

\[ u(x, y) = (y + x^2)^4/y. \]

*(ii)* What is domain of definition of the solution in \( y > 0? \)

**Problem 3.4.** Consider the PDE

\[ u_x + u_y = 2, \quad u(x, 2x) = x^2. \]

*(i)* Verify that the general solution is \( u(x, y) = f(y - x) + x + y, \) where \( f \) is an arbitrary function. Solve the problem by: *(ii)* finding \( f \) so that \( u(x, y) \) satisfies the initial condition; *(iii)* using the method of characteristics.
Lecture 3. Quasilinear equations

Problem 3.5. Solve the PDE

\[ u_x + 2xu_y = xy, \quad \text{with } u(0, y) = 1 \text{ on the segment } -\frac{1}{2} < y < \frac{1}{2}. \]

Carefully sketch the region in the \((x, y)\) plane in which this PDE determines \(u(x, y)\). That is, find the domain of definition of the solution.

Problem 3.6. Solve the PDE:

\[ yz_x + xz_y = z - 1, \quad z(x, 2x) = x^2 + 2x + 1. \]

Solution: We parameterize the initial condition with

\[ t = 0: \quad x = s, \quad y = 2s, \quad z = s^2 + 2s + 1. \]  \hspace{1cm} (3.63)

Notice that the DBC is an infinite line in this example (see the right panel of figure 3.5). Lagrange’s equation are

\[ \frac{dx}{dt} = y, \quad \frac{dy}{dt} = x, \quad \frac{dz}{dt} = z - 1. \]  \hspace{1cm} (3.64)

We combine the first two equations and get

\[ \frac{d^2x}{dt^2} - x = 0, \quad \Rightarrow \quad x = A(s)e^t + B(s)e^{-t}. \]  \hspace{1cm} (3.65)
Lecture 3. Quasilinear equations

Also
\[ y = \frac{dx}{dt} = A(s)e^t - B(s)e^{-t}. \]  (3.66)

Now, we let \( t = 0 \) and apply the initial condition in (3.63):
\[ s = A + B, \quad 2s = A - B, \quad \Rightarrow \quad A = \frac{3}{2}s, \quad B = -\frac{1}{2}s. \]  (3.67)

Thus, we now have
\[ x(s,t) = \frac{3}{2}s e^t - \frac{1}{2}s e^{-t}, \quad y(s,t) = \frac{3}{2}s e^t + \frac{1}{2}s e^{-t}. \]  (3.68)

The CBC’s obtained from (3.68) are plotted in the right panel of figure 3.5. Notice that none of the CBCs enter the region in which \( |y| < |x| \). The domain of definition of this problem is the other region, \( |y| > |x| \), which is ventilated by the CBCs which pass through \( y = 2x \).

We solve the third equation in (3.64) for \( z \)
\[ \frac{dz}{dt} = 1 + C(s)e^t. \]  (3.69)

The initial condition in (3.63) now determines \( C(s) \) in (3.69):
\[ z = 1 + s(s + 2)e^t. \]  (3.70)

So much for the first two steps of the recipe.

We are now on the home stretch — we eliminate \( s \) and \( t \) in favor of \( x \) and \( y \). If we add and subtract the expressions in (3.68) we get
\[ y + x = 3se^t, \quad y - x = se^{-t}. \]  (3.71)

Multiplying the equations above together and taking a square root we get
\[ s = \pm \sqrt{\frac{1}{3}(y^2 - x^2)}. \]  (3.72)

(Why do we take the positive branch of the square root?) Dividing the equations (3.71) and taking another square root we get
\[ e^t = \sqrt{\frac{y + x}{3(y - x)}}. \]  (3.73)

Armed with (3.72) and (3.73) we can express \( z \) in (3.70) in terms of \( x \) and \( y \):
\[ z(x, y) = 1 + \frac{1}{3}(x + y) \left[ 2 + \sqrt{\frac{1}{3}(y^2 - x^2)} \right], \quad BYU. \]  (3.74)

The final step is to check that the expression above solves the PDE.
Problem 3.7. Show that the solution of

\[(x + u)u_x + yu_y = u + y^2, \quad u(x, 1) = x, \quad -\infty < x < \infty,\]

is

\[u(x, y) = \frac{x - y^2}{1 + \ln y} + y^2.\]

What is the domain of definition of \(u(x, y)\)?

Problem 3.8. Solve

\[xu_x + yu_y = -xy, \quad u(x, y) = 5 \text{ on } xy = 1.\]

Problem 3.9. Solve

\[yu_x + x^3u_y = x^3y, \quad u(x, x^2) = x^4.\]

Problem 3.10. Show that the solution of

\[x^3u_x = u_y, \quad u(x, 0) = \frac{1}{1 + x^2},\]

is

\[u(x, y) = \frac{1 - 2x^2y}{1 + x^2 - 2x^2y}.\]

Show that the solution is not defined in \(y > 1/2x^2\) even though the data is prescribed for \(-\infty < x < \infty\).

Problem 3.11. Show that the initial value problem

\[(y - x)u_x - (y + x)u_y = 0, \quad u(x, 0) = f(x), \quad x > 0,\]

has no solution if \(f(x)\) is an arbitrary function.

Problem 3.12. Consider the initial-value problems:

\[u_x - u_y = u, \quad u(x, -x + x^2) = e^x,\]
\[u_x - u_y = u, \quad u(x, -x + x^2) = 1.\]

If the solution exists, find it and the domain of existence. If there is no solution, explain why.

Problem 3.13. Show that the initial value problem

\[yu_x + xu_y = pu, \quad u(x, x) = f(x),\]

has no solution unless \(f(x) = ax^p\). If \(f\) has this form, find the solution.
Problem 3.14. (i) Find a solution analogous to (3.48) of the PDE

\[ u_t + c(u)u_x = 0, \quad u(x, 0) = F(x). \]

(ii) Check your answer by substitution. (iii) Considering the special case \( c = u^2 \) and \( F = x \), show that once \( t > 0 \) there are either two real values of \( u(x, t) \) or no real values of \( u(x, t) \) at each point in spacetime. Locate the curve in the \((x, t)\) plane which separates these two behaviours.

Problem 3.15. (i) Find an implicit solution, analogous to (3.48), of

\[ u_t + uu_x = -\alpha u, \quad u(x, 0) = \frac{1}{1 + x^2}. \]

Make sure your answer reduces to (3.48) if \( \alpha \to 0 \). Draw a figure, like 3.3, showing both characteristics and snapshots of \( u(x, t) \). (Use several values of \( \alpha \) so that you understand how this parameter changes the solution). (ii) Calculate \( t_s \) as a function of \( \alpha \). (iii) Find the smallest value of the damping \( \alpha \) which is sufficient to prevent shock formation.

Problem 3.16. (i) Does the PDE

\[ u_t + uu_x = 0, \quad u(x, 0) = \frac{e^x}{1 + e^x}. \]

form a shock? If so, calculate \( t_s \) without solving the PDE. (ii) Find an implicit solution for \( u(x, t) \) and use MATLAB to graph this solution.

Solution. The geometric construction described following (3.51) makes it clear that this initial condition does not form a shock. (Also notice that the initial slope, \( u_x(x, 0) \), is positive everywhere and so according to (3.58) the shock time, \( t_s \), is in the past.) In part (iii), we quickly find

\[ \xi = x - \frac{e^\xi}{1 + e^\xi} t \quad u = \frac{e^\xi}{1 + e^\xi}. \]

Fiddling about with the second equation above, we can express \( \xi \) as an explicit function of \( u \): \( \xi = \ln[u/(u - 1)] \). Next, we can eliminate \( \xi \) and give an equation which defines \( u \) implicitly as a function of \( x \) and \( t \):

\[ \ln \left( \frac{u}{u - 1} \right) = x - ut. \]

We can’t solve this monster for \( u(x, t) \) explicitly. But to graph \( u(x, t) \), just as in the right panel of figure 3.3, we don’t need to...
Problem 3.17. (i) Does the PDE

\[ u_t + uu_x = 0, \quad u(x,0) = \frac{e^{-x}}{1 + e^{-x}}. \]

form a shock? If so, calculate \( t_s \) without solving the PDE. (ii) Find an implicit solution for \( u(x,t) \) and use MATLAB to graph this solution.

Problem 3.18. In the witchy example we showed in (3.60) and (3.61) that the time at which the shock forms is \( t_s = 8\sqrt{3}/9 \). Find the location, \( x_s \), at which the shock forms.

Problem 3.19. Find an expression for \( t_s \), analogous to (3.58), for the PDE

\[ u_t + u^2 u_x = 0, \quad u(x,0) = F(x). \]

Also give a simple formula for \( x_s \).

Problem 3.20. Consider

\[ u_t + u^3 u_x = 0, \quad u(x,0) = \sin x. \]

At what time, \( t_s \), and location, \( x_s \), does the solution \( u(x,t) \) first become singular?
Lecture 4

δ-functions and Green’s functions

Can I assume that you understand the concept of a Green’s function, and δ-functions, at least as far as ODEs are concerned? The material in this lecture is in section 1.5 of BO, but here is a little review in case you need it. In this lecture we’ll only discuss ordinary differential equations, and we’ll begin with evolution problems in which the independent variable is \( t = \) time. In this context the function \( \delta(t) \) embodies the concept of impulsive action. For example, when a hammer pounds a nail into a plank it must exert a force on the head of the nail. If you can draw a graph of this force as function of time then you have a good idea of what \( \delta(t) \) means.

4.1 Review of “patching”

Before we discuss δ-functions we review the solution of ordinary differential equations with discontinuous coefficients and forcing functions. This is the method of “patching”. We illustrate patching by solving a few simple examples from mechanics.

The poster child of discontinuous functions is Oliver Heaviside’s step function

\[
H(t) = \begin{cases} 
  0, & \text{if } t < 0; \\
  1, & \text{if } t > 0.
\end{cases}
\] (4.1)

Notation:

\[ \dot{\theta} = \frac{d\theta}{dt} \]
Lecture 4. δ-functions and Green’s functions

Figure 4.1: Left panel: The step function $H(t)$, and three successive integrals using the definition in (4.5). The singular behavior at $t = 0$ is smoothed out by integration. Right panel: the smoothed step function $H_\varepsilon(t)$ defined in (4.21). In this illustration $\varepsilon = [0.4, 0.2, 0.1, 0.05, 0.025]$. jumpy.eps

Notice we’ve left $H(0)$ undefined. As we solve the following differential equations it will become clear that the value of $H(0)$ is irrelevant. If it makes you feel better you can define it as $H(0) = 1/2$, but it makes no difference.

Consider a differential equation involving $H(t)$, such as the third order example

$$\dddot{y} - 7\dot{y} - 6y = H(t). \tag{4.2}$$

Something on the left hand side must balance the discontinuous forcing function on the right. Suspicion falls heavily on the term with the most derivatives i.e., $\dddot{y}$ in the example above. This means that at $t = 0$ the third derivative $\dddot{y}$ is discontinuous:

$$\lim_{\varepsilon \to 0} \left[ \dddot{y}(-\varepsilon) - \dddot{y}(\varepsilon) \right] = 1. \tag{4.3}$$

The terms with fewer derivatives are better behaved i.e., $\dot{y}(t)$ is continuous at $t = 0$. This is a general principle: Taking derivatives makes a singularity worse. The converse is: Integrating makes a singularity better.

Let’s beat this point to death by considering the Heaviside step $H(t)$,
and the functions obtained by successive integration

\[ I_1(t) = \int_{-\infty}^{t} H(t') \, dt', \tag{4.4} \]

\[ I_2(t) = \int_{-\infty}^{t} I_1(t') \, dt', \tag{4.5} \]

and so on. It is easy to see that \( I_n(t) = H(t)t^n/n! \). The left panel of figure 4.1 shows that the sudden behaviour at \( t = 0 \) becomes less prominent as we successively integrate.

**Example:** Solve

\[ \ddot{\theta} + \sigma^2 \theta = 1, \quad \theta(0) = \dot{\theta}(0) = 0. \]

The oscillator is at rest at \( t = 0 \) and then suddenly a constant force is switched on. When \( t > 0 \) the general solution is

\[ \theta(t) = \underbrace{\sigma^{-2}}_{= \text{particular solution}} + \underbrace{A \cos \sigma t + B \sin \sigma t}_{= \text{homogeneous solution}}. \]

The constants \( A \) and \( B \) are determined so that the initial conditions are satisfied. Thus at \( t = 0 \) we require that both \( \theta \) and \( \dot{\theta} \) are zero. Applying these initial conditions we quickly find

\[ \theta(t) = \sigma^{-2} (1 - \cos \sigma t). \tag{4.6} \]
The differential equation has a steady equilibrium solution, namely \( \theta(t) = 1 \). But the solution we've found above in (4.6) "rings" forever round this equilibrium — see Figure [4.2].

**Example:** Solve

\[
\ddot{\theta} + \sigma^2 \theta = H(t), \quad \theta(t < 0) = 0.
\]

This problem is almost equivalent to the example we just did, and the solution is

\[
\theta(t) = \frac{1}{\sigma^2} (1 - \cos \sigma t) H(t).
\]

This is the solid curve in Figure [4.2]. The \( H(t) \) is inserted above so that we can interpret the solution on the whole time axis \(-\infty < t < \infty\). When \( t < 0 \) the oscillator is at rest and minding its own business: the \( H(t) \) on the right of (4.8) takes care of that. Then suddenly at \( t = 0 \) the oscillator is disturbed by the application of the constant force. We insist that \( \dot{\theta}(t) \) and \( \ddot{\theta}(t) \) are continuous at \( t \). Therefore both \( \theta \) and \( \dot{\theta} \) are zero at \( t = 0^+ \), and we immediately recover the solution in the previous example. Notice that \( \ddot{\theta} \) (the term with the most derivatives) jumps discontinuously:

\[
\ddot{\theta}(0^-) = 0, \quad \text{and} \quad \ddot{\theta}(0^+) = 1.
\]

This is perfectly OK: on the right of the differential equation (4.7) we have the discontinuous function \( H(t) \), which is balanced by the discontinuous function \( \ddot{\theta}(t) \) on the left.

**Example:** Solve

\[
\ddot{\phi} + \sigma^2 \phi = H(t) - H(t - \tau), \quad \phi(t < 0) = 0.
\]

Using linearity and temporal translation invariance, the solution is

\[
\phi(t) = \theta(t) - \theta(t - \tau),
\]

where \( \theta(t) \) is defined in (4.8). One absolutely must include the \( H(t) \) in (4.8) if this beautiful trick is to work. Back in the previous example the \( H(t) \) seems like a fussy little device. But in this example the \( H(t) \) ensures that the solution \( \phi(0 < t < \tau) \) does not anticipate that the forcing on the right of (4.9) is going to switch off at the future time \( t = \tau \).

When \( t > \tau \) the solution is

\[
\phi(t > \tau) = \sigma^{-2} [\cos \sigma (t - \tau) - \cos \sigma t].
\]

We can pick \( \tau \) so that the \( \phi(t) \) above is equal to zero e.g., take \( \sigma \tau = 2\pi \). Alternatively if \( \sigma \tau = \pi \) then the oscillation has maximum amplitude viz., \( \phi(t > \tau) = -2 \cos \sigma t \). So the ultimate amplitude of the oscillation can be nicely controlled by timing when the force is switched off.
4.2 \( \delta \)-sequences

The \( \delta \)-function is not a real function; instead, the \( \delta \)-function is interpreted as the limit (\( \epsilon \to 0 \)) of a sequence of functions with two properties:

\[
\text{(a): } \int_{-\infty}^{\infty} \delta_\epsilon(t) \, dt = 1; \\
\text{(b): } \delta_\epsilon(t) \to 0 \text{ as } \epsilon \to 0 \text{ with } t \neq 0 \text{ fixed.} \tag{4.11}
\]

Because of property (a), the integral of each member of the sequence is independent of the parameter \( \epsilon \). Property (b) is satisfied because the functions become very sharply peaked as \( \epsilon \to 0 \). We will denote the \( \delta \)-sequence by \( \delta_\epsilon(t) \) but for brevity we denote the limit (when \( \epsilon = 0 \)) simply by \( \delta(t) \).

There is a simple recipe for producing \( \delta \)-sequences. Start with some ‘humplike’ function, \( h(t) \), which is normalized so that the integral of \( h \) is equal to one. Here are some examples:

\[
h_1(t) = \frac{e^{-t^2}}{\sqrt{\pi}}, \quad h_2(t) = \frac{1}{\pi} \frac{1}{1 + t^2}, \quad h_3(t) = \frac{\sin t}{\pi t}. \tag{4.12}
\]

To obtain a \( \delta \)-sequence, take

\[
\delta_\epsilon(t) = \frac{1}{\epsilon} h \left( \frac{t}{\epsilon} \right). \tag{4.13}
\]

Make sure you understand why the construction above satisfies the properties in (4.11). (There are other recipes for obtaining a \( \delta \)-sequence; the one above is the most elementary.)

The limit of a \( \delta \)-sequence is only defined when the sequence is inside an integral. For example, if \( f(t) \) is a smooth function then

\[
\int_{-\infty}^{\infty} f(t_1) \delta(t - t_1) \, dt_1 \overset{\text{def}}{=} \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(t_1) \delta_\epsilon(t - t_1) \, dt_1. \tag{4.14}
\]

Because of the integrand on the right of (4.14) is nonzero only in the immediate neighbourhood of \( t_1 = t \), we can pull the slowly varying function \( f(t_1) \) outside the integral so that:

\[
\int_{-\infty}^{\infty} f(t_1) \delta(t - t_1) \, dt_1 = f(t). \tag{4.15}
\]
Lecture 4. \( \delta \)-functions and Green's functions

This is the ‘sifting-property’ of a \( \delta \)-function.

Here is a list of some of the more useful properties of the \( \delta \)-function:

\[
f(t) \delta(t - t_1) = f(t_1) \delta(t - t_1)
\]
(4.16)

\[
t \delta(t) = 0
\]
(4.17)

\[
\delta(-t) = \delta(t)
\]
(4.18)

\[
\delta(at) = |a|^{-1} \delta(t)
\]
(4.19)

In addition, there is the crucial sifting property in (4.15).

Integration by parts shows that the derivative of the \( \delta \)-function has the property

\[
\int_{-\infty}^{\infty} f(t) \delta'(t) \, dt = -f'(0).
\]
(4.20)

We can view \( H(t) \) as the limit of a sequence of very smooth functions. This limiting process is illustrated in the right panel of Figure 4.1 using

\[
H(t) = \lim_{\varepsilon \to 0} \frac{1}{1 + \exp(-t/\varepsilon)}.
\]
(4.21)

It is easy to see that the \( t \)-derivative of the sequence \( H_{\varepsilon}(t) \) defines a perfectly good \( \delta \)-sequence:

\[
\delta_{\varepsilon}(t) = \frac{dH_{\varepsilon}}{dt}.
\]
(4.22)

Thus blithely saying ‘the derivative of the limit is the limit of the derivative’, we get

\[
\frac{d}{dt} H(t) = \delta(t).
\]
(4.23)

This is an extremely useful result, and it reinforces our earlier point that differentiating a singular function makes the singularity worse — \( \delta(t) \) is more singular than \( H(t) \).

**Example:** Suppose we build a \( \delta \)-function using the recipe in (4.13) and the asymmetric hump function:

\[
h_4(t) = \frac{e^{-t^2}}{\sqrt{\pi}} (1 + \alpha t).
\]

Is (4.18) satisfied, even though \( h_4(t) \neq h_4(-t) \)?
Yes: (4.18) means that if \( f(t) \) is any function which is continuous at \( t = 0 \) then
\[
\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(t) \frac{1}{\epsilon} h_4 \left( \frac{t}{\epsilon} \right) \, dt = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(t) \frac{1}{\epsilon} h_4 \left( \frac{-t}{\epsilon} \right) \, dt.
\]

The result above is correct because the symmetry-breaking term proportional to \( \alpha \) in \( h_4(t) \) gives a vanishing contribution in the limit. ■

Example

Although \( t\delta(t) = 0 \), \( t\delta'(t) \neq 0 \). Show that \( t\delta'(t) = -\delta(t) \). ■

If a function is discontinuous at a point, its derivative at that point is a \( \delta \)-function with an amplitude that is proportional to the size of the discontinuity. Equation (4.23) is one example, and another is the function \( \text{sgn}(t) \) which is equal to \( \pm 1 \) according to whether \( x \) is positive or negative. Thus
\[
\frac{d}{dt} \text{sgn}(t) = 2\delta(t). \tag{4.24}
\]

This is not mysterious. If \( a < 0 < b \) then, on one hand, a pedestrian application of the definition of \( \text{sgn}(t) \) gives:
\[
\int_{a}^{b} \text{sgn}(t) \frac{df}{dt} \, dt = \int_{0}^{b} \frac{df}{dt} \, dt - \int_{a}^{0} \frac{df}{dt} \, dt = f(b) - 2f(0) + f(a). \tag{4.25}
\]

On the other hand, we can use (4.23) to integrate by parts:
\[
\int_{a}^{b} \text{sgn}(t) \frac{df}{dt} \, dt = [\text{sgn}(t)f(t)]_{a}^{b} - \int_{a}^{b} 2\delta(t)f(t) \, dt,
= f(b) - 2f(0) + f(a). \tag{4.26}
\]

Although we’ve differentiated a discontinuous function, everything is hunky dory.

### 4.3 Green’s function of the harmonic oscillator

The idea of a Green’s function is illustrated by the following example. Suppose we need to solve the forced oscillator equation
\[
\dot{\theta} + \sigma^2 \theta = f(t), \quad \theta(0) = \dot{\theta}(0) = 0. \tag{4.28}
\]
Although this problem is posed as an initial value problem beginning at \( t = 0 \), we can also think of it as posed on \(-\infty < t < \infty\) provided we agree that
\[
 f(t < 0) = 0 \quad \text{and} \quad \theta(t < 0) = 0 .
\] (4.29)
That is, nothing is happening when \( t < 0 \), and then the force \( f(t) \) switches on at \( t = 0 \) and sets the oscillator into motion.

Suppose we somehow first solve the equation
\[
 \ddot{G} + \sigma^2 G = \delta(t) , \quad G(t < 0) = 0 .
\] (4.30)
Then the solution of (4.28) is
\[
 \theta(t) = \int_{-\infty}^{\infty} G(t - t') f(t') \, dt' .
\] (4.31)
The function \( G(t) \) is called the Green’s function, or sometimes the fundamental solution, of (4.28). The Greens function is the oscillation induced by a \( \delta \)-function kick. It is a remarkable fact that if we can solve the differential equation with this kick-forcing, then we can solve the equation with any forcing \( f(t) \) via the formula (4.31). In the next section we discuss this concept for a general class of linear systems.

Noting\(^1\) that \( G(t) \) and \( f(t) \) are both zero when \( t < 0 \), we can re-write (4.31) as
\[
 \theta(t) = \int_{0}^{t} G(t - t') f(t') \, dt' .
\] (4.32)
This seems simpler than (4.31). However the limits \( -\infty \) in (4.31) are convenient when we prove that (4.31) satisfies (4.28) by substitution:
\[
 \left( \frac{d^2}{dt^2} + \sigma^2 \right) \theta(t) = \int_{-\infty}^{\infty} \left( \frac{d^2}{dt^2} + \sigma^2 \right) G(t - t') f(t') \, dt' ,
\]
\[
 = \int_{-\infty}^{\infty} \delta(t - t') f(t') \, dt' ,
\]
\[
 = f(t) .
\] (4.33)
Notice we used time-translation symmetry above: if \( G(t) \) is the response to a kick at \( t = 0 \), then \( G(t - t') \) is the response to a kick at \( t = t' \).

This is all very well, but we still have to solve (4.30). We begin by noting that the forcing term on the right hand side is zero except in

\(^{1}\)We’ve used causality: \( G(t) \) is the response of an oscillator to an impulsive kick. The oscillator is at rest before its kicked i.e., \( G(t < 0) = 0 \).
the neighborhood of \( t = 0 \). Thus the solution when \( t < 0 \) is simply \( G(t) = 0 \). Then, at \( t = 0 \), the oscillator gets an impulsive kick from the \( \delta \)-function. The effect of the kick is calculated by integrating \((4.30)\) over a small time interval that surrounds \( t = 0 \), say \(-\tau < t < \tau\). Because \( \dot{G}(-\tau) = G(-\tau) = 0 \) the result of this integration is

\[
\dot{G}(\tau) + \sigma^2 \int_{-\tau}^{\tau} G(t) \, dt = 1. \tag{4.34}
\]

As \( \tau \to 0 \) one of the terms on the left hand side must balance the 1 on the right hand side. Some thought shows that it must be \( \dot{G}(\tau) \) since the other term vanishes as \( \tau \to 0 \). This means there is a jump in \( \dot{G} \) at \( t = 0 \): \( G(t) \) is continuous, \( \dot{G}(t) \) is discontinuous at \( t = 0 \) and \( \ddot{G}(t) \) has a \( \delta \)-function component at \( t = 0 \).

To summarize this argument: the effect of \( \delta \)-function function forcing on the right hand side of \((4.30)\) is to produce a jump in \( \dot{G}(t) \) at \( t = 0 \):

\[
G(0^-) = G(0^+) = 0, \quad \dot{G}(0^-) = 0, \quad \dot{G}(0^+) = 1. \tag{4.35}
\]

The conditions at \( t = 0^+ \) can then be used as “effective initial conditions” for the homogeneous equation \( \ddot{G} + \sigma^2 G = 0 \), and the solution is

\[
G(t) = \sigma^{-1} \sin(\sigma t)H(t). \tag{4.36}
\]

The factor \( H(t) \) is inserted to remind us that the oscillator is at rest when \( t < 0 \).

The expression in \((4.31)\) gives the unfortunate impression that the solution at time \( t \) depends on what happens in the future: the integral involves knowing \( f(t') \) with \( t' > t \). Now that we have the Green’s function in \((4.36)\) we see that

\[
\dot{\theta}(t) = \int_{-\infty}^t \sigma^{-1} \sin[\sigma(t-t')] f(t') \, dt'. \tag{4.37}
\]

If the forcing is zero when \( t < 0 \) then we can also replace the lower limit above by zero, so the integral runs from 0 to \( t \). However we might consider a slightly more general problem in which

\[
\lim_{t \to -\infty} f(t) = 0, \quad \text{and} \quad \lim_{t \to -\infty} \dot{\theta}(t) = 0. \tag{4.38}
\]

\(^2\)The same solution is found by variation of parameters. Green’s solution provides much more physical insight.
Lecture 4. \( \delta \)-functions and Green’s functions

The expression in (4.37) is the solution of this problem in which the oscillator is at rest in the distant past.

Once again, the general principle is that higher derivatives of a function become increasingly singular: in an equation like (4.30) it is always the most highly differentiated term (in this case \( \ddot{G} \)) that balances the \( \delta \)-function:

\[
\ddot{G} = \delta(t) \quad \text{as} \quad t \to 0.
\]  \hspace{1cm} (4.39)

The \( \delta \)-function cannot be balanced by \( G \), since then \( \dot{G} \) would be a \( \delta \)-function derivative, and \( \ddot{G} \) would be a double derivative of \( \delta \), and there is nothing to balance this horrible singularity.

Notice also that the bare mathematical problem

\[
\ddot{G} + \sigma^2 G = \delta(t),
\]  \hspace{1cm} (4.40)

has the solution in (4.36), but also at least two other solutions

\[
G_2 = -\sigma^{-1} \sin(\sigma t)H(-t) \quad \text{and} \quad G_3 = \frac{1}{2} \sigma^{-1} \sin \sigma |t|.
\]  \hspace{1cm} (4.41)

All the \( G \)'s above a have a derivative-jump at \( t = 0 \): \( \dot{G}(0^+) - \dot{G}(0^-) = 1 \). We have made specific use of causality to identify (4.36) as the relevant Green’s function. We’ll see later that the solutions above might (and in fact do) appear in other physical problems.

**Example:** Find an integral representation of the causal solution of

\[
\ddot{y} - 7 \dot{y} - 6y = f(t).
\]

The forcing \( f(t) \) is zero when \( t < 0 \), and so is \( y(t) \). Find a condition on \( f(t) \) which ensures that \( \lim_{t \to -\infty} y(t) < \infty \).

The causal Green’s function satisfies

\[
\ddot{g} - 7 \dot{g} - 6g = \delta(t), \quad g(t < 0) = 0.
\]

“Causal” means nothing happens before the kick arrives at \( t = 0 \) i.e., \( g(t < 0) = 0 \). When \( t > 0 \) the general solution of the homogeneous equation is obtained with \( g = e^{\lambda t} \), leading to

\[
\lambda^3 - 7\lambda - 6 = 0 \quad \text{or} \quad \lambda = -1, -2 \text{ and } +3.
\]

The cubic polynomial for \( \lambda \) above is solved by inspired guessing, or with the mathematica command

\text{Roots}[x^3 - 7 x + 6 == 0, x]
Thus the general solution when \( t > 0 \) is
\[
g = ae^{-t} + be^{-2t} + ce^{3t}.
\] (4.42)

We have to determine the constants \( a, b \) and \( c \) by applying initial condition at \( t = 0^+ \), just after the impulse. At \( t = 0 \) the \( \delta \)-kick is balanced by \( \ddot{y} \), so integrating from \( t = -\epsilon \) to \( t = +\epsilon \) we have
\[
\ddot{g}(\epsilon) - \ddot{g}(-\epsilon) = \int_{-\epsilon}^{\epsilon} \delta(t) \, dt
\]
In other words, the second derivative \( \ddot{g} \) jumps discontinuously from zero to one as a result of the kick. The function \( g(t) \) and the first derivative \( \dot{g}(t) \) are continuous at \( t = 0 \), implying that \( g(0) \) and \( \dot{g}(0) \) are zero. Thus we have three effective initial conditions
\[
g(0^+) = 0, \quad \dot{g}(0^+) = 0, \quad \ddot{g}(0^+) = 1 \quad (4.43)
\]
to determine the three constant \( a, b \) and \( c \) in (4.42). The conditions above result in the \( 3 \times 3 \) linear system
\[
\begin{align*}
a + b + c &= 0, \\
a + 2b - 3c &= 0, \\
a + 4b + 9c &= 1.
\end{align*}
\]
Solving the equations above we find
\[
g = \left[ -\frac{1}{4}e^{-t} + \frac{1}{2}e^{-2t} + \frac{1}{20}e^{3t} \right] H(t), \quad (4.44)
\]
where the \( H(t) \) is inserted for causality. The general causal solution is therefore
\[
y'(t) = \int_{0}^{t} g(t - t') f(t') \, dt'.
\]
The long-time behaviour of \( y(t) \) will be dominated by the term in \( g(t) \) proportional to \( e^{3t} \):
\[
y'(t) = \frac{e^{3t}}{20} \int_{0}^{t} e^{-3t'} f(t') \, dt' + \text{decaying terms} e^{-t} \text{ and } e^{-2t}.
\]
To completely destroy the growing term as \( t \to \infty \) one requires that the forcing satisfy
\[
\lim_{t \to \infty} e^{3t} \int_{0}^{\infty} e^{-3t'} f(t') \, dt' = 0.
\]
The advantage of the Green’s function is that one can obtain general results determining important qualitative properties of the solution, such as the condition above. If one wants to solve the equation with a particular \( f(t) \) then it may be best to avoid the Green’s function integral representation and use simpler techniques such as undetermined coefficients.
Example: Find an integral representation of the solution of
\[ \dddot{h} - 7\ddot{h} - 6h = \delta(t). \]
satisfying \( \lim_{t \to \pm\infty} h(t) = 0. \)

This problem is different from the previous one because we don’t seek a causal solution i.e., we might have \( h(t < 0) \neq 0 \). Instead we’re after a solution on the whole \( t \)-axis, \(-\infty < t < \infty\), which is zero as \( t \to -\infty \) and as \( t \to +\infty \).

Let’s work out the details. The homogeneous problem is the same as the previous case, and the conditions at \( \pm\infty \) demand
\[ h(t) = \begin{cases} \begin{array}{ll} pe^{3t}, & \text{for } t < 0; \\ qe^{-t} + re^{-2t}, & \text{for } t > 0. \end{array} \end{cases} \] (4.45)

The construction above secures the vanishing of \( h(t) \) as \( t \to \pm\infty \). Now the three constants of integration are determined by patching at \( t = 0 \) i.e., by requiring that \( h \) and \( \dot{h} \) are continuous at \( t = 0 \), and that \( \ddot{h}(0^+) - \ddot{h}(0^-) = 1 \). Thus
\begin{align*}
p &= q + r, \\
3p &= -q - 2r, \\
q + 4r - 9p &= 1.
\end{align*}

Solving the equations above, we find we can write the solution as
\[ h(t) = g(t) - \frac{1}{20}e^{3t}, \]
where \( g(t) \) is the causal solution back in (4.44). With hindsight this is obvious: the difference \( g(t) - h(t) \) satisfies the homogeneous equation, and we construct \( h(t) \) from \( g(t) \) simply by subtracting the \( t \to +\infty \) growing component \( e^{3t} \).

4.4 The black box

I have introduced the concept of the Green’s function as the impulse-response of the oscillator equation in (4.30). However the idea is much more general and important than this example might suggest. Suppose we have a device (the legendary ‘black box’) that accepts an input \( U(t) \) and produces an output \( V(t) \). We don’t enquire about the internal workings of the black box. Obviously this is a very general concept e.g., one can view ones colleagues as black boxes.

Some black boxes are simpler than others. The very simplest examples are:
Lecture 4. \(\delta\)-functions and Green’s functions

Figure 4.3: The black box with input \(U(t)\) and output \(V(t)\).

BlackBox.eps

1 Causal; the black box can’t anticipate future events.

2 Time translationally invariant; if the input \(U(t)\) produces \(V(t)\), then \(U(t - t_1)\) produces \(V(t - t_1)\).

3 Linear; the response to \(\alpha U_1(t) + \alpha_2 U_2(t)\) is \(\alpha V_1(t) + \alpha_2 V_2(t)\).

An example of this type of simple black box is a simple thermometer. The input to a thermometer might be the “true temperature” of the air surrounding the bulb; the output is the “measured temperature”. If the true temperature is changing rapidly (for instance, if the thermometer is mounted on the wing of a plane) then the measured temperature will be some smoothed or averaged or “filtered” version version of the true temperature.

Another example is the measured deflection \(V(t)\) of a structure at some point in response to a varying load \(U(t)\) at some other point. If the load is not too great then we expect that the response varies linearly with the load. Buildings don’t start shaking before an earthquake, so the system is causal. And if the material in the structure isn’t aging then the response to a standard load now and six months from now will be the same. These are the three requirements above.

The three assumptions are sufficient to completely characterize the operation of the black box if we know the response produced by the special input \(\delta(t)\). This is the impulse response, or the Green’s function of the black box, denoted by \(G(t)\). Given \(G(t)\), the response to an arbitrary input \(U(t)\) is

\[
V(t) = \int_{-\infty}^{t} G(t - t') U(t') \, dt'.
\]  

(4.46)

Thus, if we measure or calculate the response of the black box to a single


**Lecture 4. δ-functions and Green’s functions**

Figure 4.4: The input is $H(t)$ and output $V(t)$ ramps up over a time $\tau$ to the final level ($\tau = 2$ in this illustration). 

input, $\delta(t)$, then we know how the system responds to *any* input by evaluating the integral in (4.46).

Because the response $V(t)$ cannot depend on the future behaviour of the input $U(t)$, the upper limit of integration in (4.46) is $t' = t$. (Anyway, because $G(t < 0) = 0$ you can write $+\infty$ as the upper limit if you like — it won’t make any difference.)

To prove (4.46), partition up the $t$-axis into small intervals of length $dt$ and then approximate the input $U(t)$ as a ’staircase’, with a constant value in each $dt$-interval. The response to each ‘pulse’ in the staircase is then given by a single Green’s function (multiplied by $dt$) and (4.46) is the result of linearly superposing these responses.

**Example:** Suppose that the input temperature is $H(t)$ and the observed response to this step function input is

\[
V(t) = \begin{cases} 
0, & \text{if } t < 0; \\
t/\tau, & \text{if } 0 < t < \tau; \\
1, & \text{if } \tau < t.
\end{cases}
\]

This ramp is shown in Figure 4.4. Find the Green’s function. The constant $\tau$ is a response time of the thermometer.

The Green’s function of the thermometer is obtained by using (4.23) i.e., by taking the derivative of the step-response $V(t)$ in Figure 4.4,

\[
\frac{d}{dt}H(t) = \delta(t)
\]

\[
G(t) = \tau^{-1} [H(t) - H(t - \tau)].
\]

We’ve used linearity to say that if the input $U(t)$ elicits the response $V(t)$, then the response to $dU/dt$ is $dV/dt$. 

---

SIO203C, W.R. Young, March 21, 2011
Lecture 4. \(\delta\)-functions and Green’s functions

Now using (4.46) and (4.47) we can write down a simple expression for the response of this hypothetical thermometer to any input:

\[
V(t) = \frac{1}{\tau} \int_{t-\tau}^{t} U(t') \, dt'.
\] (4.48)

We see that the thermometer is averaging the arriving signal \(U(t)\) over the most recent times. The averaging is done ‘on-the-fly’, so that the measurement at \(t\) depends on the past temperature back to \(t - \tau\).”

Example: Find the response of the thermometer to a signal \(u(t) = e^{i\omega t}\). Discuss the two limits \(\omega \tau \gg 1\) and \(\omega \tau \ll 1\).

Example: Suppose two of these thermometers are connected in series so that the input of the second is the output of the first. Find the response of this composite system to the inputs \(\delta(t)\) and \(H(t)\).

4.5 Problems

In example 1 of section 7.4, BO give short shrift to the elementary but important method of patching. Their discussion of \(\delta\)-functions and Green’s function in section 1.5 is a little more complete, but still not very detailed. The first six or seven problems are examples of “patching”.

Problem 4.1. Use patching to obtain non-trivial solutions of the differential equations

\[
\begin{align*}
\gamma' + \text{sgn}(x)\gamma &= 0, & \gamma(\pm \infty) &= 0, \\
\gamma'' + \text{sgn}(x)\gamma &= 0, & \gamma(-\infty) &= 0.
\end{align*}
\]

Problem 4.2. Solve

\[
\gamma''' + \gamma = H(x), \quad \gamma(-\infty) = 0, \quad \gamma(\infty) = 1.
\]

(Above, \(H(x)\) is the Heaviside step function.)

Problem 4.3. Consider the oscillator problem

\[
\ddot{\chi} + \sigma^2 \chi = \frac{t}{\tau} H(t) - \left( \frac{t}{\tau} - 1 \right) H(t - \tau), \quad \chi(t < 0) = 0.
\]

Draw a graph of the right hand side \(r(t)\) as a function of \(t\). Solve the initial value problem and discuss how the amplitude of the ultimate oscillation depends on \(\tau\).
Problem 4.4. (a) Solve the forced-damped oscillator problem
\[ \ddot{\theta} + \mu \dot{\theta} + \sigma^2 \theta = H(t), \quad \theta(t < 0) = 0. \]
Check your answer by comparing it with the dashed curve in Figure 4.2.
(b) Show that as \( t \to \infty \) half the work done by the force \( H(t) \) is used to increase the potential energy of the oscillator and the other half is dissipated by drag.

Problem 4.5. Solve
\[ \ddot{\theta} + \sigma^2 \theta = 0, \quad \theta(t < 0) = a \cos(\sigma_1 t + \alpha), \]
where
\[ \sigma^2(t) = \begin{cases} \sigma_1^2, & \text{if } t < 0; \\ \sigma_2^2, & \text{if } t > 0. \end{cases} \]

Problem 4.6. Solve the forced oscillator equation
\[ \ddot{\theta} + \theta = e^{-|t|}, \quad \lim_{t \to -\infty} \theta(t) = 0. \]
Check your answer by showing that the amplitude of the oscillation at \( t = +\infty \) is maximized by taking \( \alpha = 1 \).

Problem 4.7. Reconsider the forced oscillator problem in the example surrounding (4.9). Obtain the Green’s function in (4.36) by taking the limit \( \tau \to 0 \) of the solution in (4.10).

Problem 4.8. (i) Show that
\[ f(n, z) = \sqrt{n} \int_{-\infty}^{\infty} \frac{dx}{[1 + (x-z)^2]^n} \]
is independent of \( z \) and becomes independent of \( n \) in the limit \( n \to \infty \). (A heuristic argument based on the size of the interval in which the integrand is significantly different from zero is all that is required.) (ii) Use the well known result
\[ \left(1 + \frac{2y}{n}\right)^n \to e^y \quad \text{as} \quad n \to \infty \]
to evaluate \( f(\infty, z) \). (iii) Approximately evaluate the integral
\[ I \equiv \int_0^{\infty} \frac{\cos x^2}{(1 + x^2)^4} \, dx. \]
Problem 4.9. Evaluate

\[ I \equiv \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{\epsilon \sin (e^x)}{\epsilon^2 + x^2} \, dx. \]

and

\[ J = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \sin (e^x) \frac{\epsilon}{dx \epsilon^2 + x^2} \, dx. \]

Problem 4.10. Make a \( \delta \)-sequence with the function \( \text{sech}^2(x) \).

Problem 4.11. Explain why the dimension of \( \delta(t) \), with \( t \)-time, is \( (\text{time})^{-1} \).

In general \( \delta \)-functions are dimensional quantities (unlike bona fide functions such as \( \cos \omega t \)).

Problem 4.12. Use integration by parts to evaluate \( \int_{-1}^{1} |t| f''(t) \, dt \). (Don’t be afraid to differentiate \( |t| \) twice.)

Problem 4.13. (i) Show that if \( f(x) \) is a function which vanishes at \( x_n \) then

\[ \delta[f(x)] = \sum_n \frac{\delta(x - x_n)}{|f'(x_n)|}. \quad (4.49) \]

(ii) Calculate

\[ \int_{-\infty}^{\infty} \delta(\sin x) \exp(-|x|) \, dx. \quad (4.50) \]

(iii) Show that

\[ \delta(t^2 - a^2) = \frac{[\delta(t - a) + \delta(t + a)]}{2|a|}. \quad (4.51) \]

Problem 4.14. If \( (x, y) = r(\cos \theta, \sin \theta) \), show that

\[ \int \delta(a - x \cos \phi - y \sin \phi) \, d\phi = \frac{2H(r-a)}{\sqrt{r^2 - a^2}}. \quad (4.52) \]

Problem 4.15. Verify by substitution that (4.37) satisfies (4.28).

Problem 4.16. Consider the pendulum in the strongly over-damped limit. After suitable scaling the equation of motion is

\[ \mu \ddot{\theta} + \dot{\theta} + \theta = f(t), \]

where \( \mu \ll 1 \). Obtain the Green’s function \( G(t) \) for this problem. Since \( \mu \ll 1 \) one might be tempted to neglect the term \( \mu \dot{\theta} \). Make this approximation and obtain the Green’s function of the approximate equation. Use MATLAB to make a graphical comparison between the two Green’s functions \( \mu = 0.01 \) and 0.1.
Problem 4.17. Suppose that the output of our hypothetical thermometer is observed to be:

\[ V(t) = \begin{cases} 
0, & \text{if } t < 0; \\
1 - \exp(-\alpha t), & \text{if } t > 0.
\end{cases} \]

Reconstruct the input, \( U(t) \). (This one might be tough.)

Problem 4.18. Find the impulse response of a linear system in which the output \( V(t) \) is obtained from the input \( U(t) \) by solving

\[ \ddot{V} + 2\dot{V} + V = \dot{U}. \]

Problem 4.19. Find the Green’s function of the equation

\[ \frac{d^n y}{dt^n} = f(t), \quad \text{with} \quad \lim_{t \to -\infty} y(t) = 0. \]

(Assume that \( f(t) \) vanishes at least exponentially fast as \( t \to -\infty \).) Consider the function \( Y(t) \) defined by

\[ \left( \frac{d}{dt} \right)^{17} Y = e^{-t^2} \cos(mt), \quad \text{with} \quad \lim_{t \to -\infty} Y(t) = 0. \]

Show that when \( t \gg 1 \)

\[ Y(t) \sim A(m)t^{16} + O(t^{15}), \]

and find \( A(m) \).
Problem 4.20. The figure above shows a function $A(x)$ and its derivative $dA/dx$. Here are some differential equations:

\[ B''' + B = H, \quad C'' - C = -H, \quad D''' - D = -H. \]

$H(x)$ on the RHS is the Heaviside step function, which is zero if $x < 0$ and one if $x > 0$. (i) Which of the differential equations does $A(x)$ satisfy? (Guesses don’t count — explain your reasoning.) (ii) Sketch a graph of $A''(x)$, paying particular attention to the behaviour of $A''$ near $x = 0$. 


Lecture 5

Green’s functions for boundary value problems

5.1 A suspended string

In our earlier examples of Green’s functions and δ-functions the independent variable is time. The examples we discussed were evolutionary problems and consequently we frequently made arguments using causality and stationarity. But there is another equally important class of models based on the impulses or point singularities in the spatial domain. As an example of a spatial Green’s function, consider the wave equation

\[ \rho u_{tt} - Tu_{xx} = -\rho g, \quad u(0) = u(L) = 0. \]  

(5.1)

The tension in the string is \( T \) (same dimensions as force), and \( \rho \) is the mass per unit length of the string; the boundary conditions are that the ends of the string are fixed at \( x = 0 \) and \( x = L \). The right hand side of (1) is a force per length which, in a gravitational field, is simply the weight of the string. If the string has nonuniform thickness then \( \rho \) is a function of \( x \).

Example: Find the steady \( (u_t = 0) \) solution of (1) when \( \rho \) is a constant.

Example: Consider a string with uniform density \( \rho_0 \). A small bead of mass \( m \) is then attached to the middle \( (x = L/2) \) of the string. Use a \( \delta \)-function

---

1We’ll discuss the wave equation in much more detail later. The material in this section can be understood without reference to this important physical example.
to represent the total mass density, \( \rho(x) \) of the composite system. Find the steady \((u_t = 0)\) solution of (1) for this composite system and sketch the shape of the string when \( m \gg \rho_0 L \).

Our goal is to find an integral representation of the steady solution, \( u_t = 0 \), of (1) when \( \rho \) is some given function of \( x \). After suitable non-dimensionalization, the problem is

\[
y'' = w(x), \quad y(0) = y(1) = 0,
\]

where \( w \propto \rho \) is a prescribed function of \( x \). The ends of the string at \( x = 0 \) and \( x = 1 \) are fixed so that that there is no displacement. Unlike the oscillator equation, this is a *boundary value problem*; the independent variable, \( x \), is a spatial variable (not time); there is no principle of causality here. A second difference is that the problem is not translationally invariant: the point \( x = 1/4 \) is not the same as \( x = 1/2 \) because \( x = 1/2 \) is maximally distant from the supports at \( x = 0 \) and \( x = 1 \). Thus, for instance, if we load the string with a point mass then the displacement at \( x = 1/4 \) is less than the displacement at \( x = 1/2 \).

To solve (5.2) we first must solve for the displacement produced by an impulse load at an arbitrary point on the string. Thus, the Green’s function, \( G(x, \xi) \), satisfies

\[
G'' = \delta(x - \xi), \quad G(0) = G(1) = 0.
\]

If we integrate (5.3) over a small interval containing the \( \delta \) function then we obtain the jump condition at \( x = \xi \):

\[
G'(\xi^+) - G'(\xi^-) = 1.
\]

\( G \) itself is continuous at \( x = \xi \). Next, we consider the interval \( 0 < x < \xi \) and write down the most general solution of \( G_{xx} = 0 \) which satisfies the boundary condition, namely, \( G(x) = Ax \). We do the same in the interval \( \xi < x < 1 \). Thus, in this second interval, \( G = B(1 - x) \).

Enforcing continuity of \( G \) at \( \xi \) and the jump condition in (4) gives

\[
A\xi = B(1 - \xi), \quad \text{and} \quad -B - A = 1.
\]

Solving for \( A \) and \( B \), we finally have the Green’s function:

\[
G(x, \xi) = \begin{cases} 
-(1 - \xi)x, & \text{if } 0 < x < \xi; \\
-(1 - x)\xi, & \text{if } \xi < x < 1.
\end{cases}
\]
Now that we have the Green's function, the solution of (5.2) is obtained by linear superposition

\[ y(x) = \int_0^1 G(x, \xi) w(\xi) \, d\xi \]  

(5.7)

You can check that (5.7) satisfies (5.2) by substitution; you'll need to use the sifting property of the \( \delta \)-function.

The Green's function above is not a function only of \( x - \xi \) because the problem is not translationally invariant. However there is a certain symmetry between \( x \) and \( \xi \). If you draw the contours of constant \( G \) in the \( (x, \xi) \) plane you will see that the function is unchanged if you reflect in the line \( x = \xi \). In terms of algebra, this symmetry means that

\[ G(x, \xi) = G(\xi, x). \]  

(5.8)

This is the principle of reciprocity.

**Example:** Suppose that the boundary conditions in are inhomogeneous

\[ y'' = w(x), \quad y(0) = 0, \quad y(1) = 1. \]

(For example, one end of the string is clamped at a higher level than the other.) Find an integral representation for the solution of this problem.

If we represent the solution as \( y(x) = x + \tilde{y}(x) \), then \( \tilde{y}'(x) \) satisfies the problem with homogeneous boundary conditions. Thus the general solution is

\[ y(x) = x + \int_0^1 G(x, \xi) w(\xi) \, d\xi, \]

where \( G(x, \xi) \) is the Green's function in (5.6).

### 5.2 Solvability conditions

Now suppose that both ends of the string are ‘free’. This means that the boundary conditions of (2) are changed so that

\[ y''' = w(x), \quad y'(0) = y'(1) = 0. \]  

(5.9)

Proceeding as before, we hope to obtain the Green's function by solving

\[ G''' = \delta(x - \xi), \quad G'(0) = G'(1) = 0. \]  

(5.10)
Lecture 5. Green’s functions for boundary value problems

However, when we try solve (5.10), it all goes horribly wrong: there is no solution.

With hindsight we see that if we integrate (5.9) from $x = 0$ to $x = 1$ we find that $0 = \int_0^1 w(x) \, dx$; this is a contradiction unless the integral of $w$ happens to be zero. Thus, generally, (5.9) does not have a solution. So the problem is not with our Green’s function technique — we are trying to solve a problem without a solution, and the method tells us this very quickly. You cannot ask for more than this from mathematics alone.

Returning to the physical problem back in (5.1) makes the source of our difficulties clear. The string is in a gravitational field, but the ends are unsupported. Thus the string falls freely. The string does remain stretched between $x = 0$ and $x = L$; the ends can push and pull in the horizontal direction, but they cannot exert a vertical force — this is what I mean by ‘free’. (You can imagine some configuration of curtain rings and slippery vertical rods etcetera.)

In fact, if we integrate (5.1) from $x = 0$ to $x = L$, with free boundary conditions $u_x = 0$, then we obtain

$$\int_0^L \rho u_{tt} \, dx = \int_0^L \rho g \, dx. \tag{5.11}$$

Thus the problem does not have a steady solution. This is where all our difficulties spring from: we assumed a steady solution and none exists. Here is the cure: if the string starts from rest then the solution of (5.11) is that

$$\ddot{u} = \frac{1}{2} gt^2, \quad \ddot{u}(t) \equiv \int_0^L \rho u \, dx / \int_0^L \rho g \, dx; \tag{5.12}$$

$\ddot{u}(t)$ is the mass weighted average of the displacement (that is, the vertical position of the center of gravity of the string). So now we look at the displacement relative to this falling level

$$u'(x, t) \equiv u(x, t) - \ddot{u}(t). \tag{5.13}$$

Substituting (5.13) into (5.1) we find that $u'$ satisfies (1) with the right hand side replaced by zero. This result is expected because $u'$ is the displacement in a freely falling frame of reference (in which there is no apparent gravity).

Now return to the mathematical problem in (5.9) and suppose that it does happen that

$$\int_0^1 w(x) \, dx = 0. \tag{5.14}$$
If \( w \) does satisfy this 'solvability condition' then we do not obtain a contradiction when we integrate (8) from \( x = 0 \) to \( x = 1 \) and we optimistically guess that the differential equation does have a solution. Let us now find an integral representation of this solution. The correct modification of (5.10) is that

\[ G'' = \delta(x - \xi) - 1, \quad G'(0) = G'(1) = 0. \]  

The solution of (14) where \( x \neq \xi \) is that

\[ G(x, \xi) = \begin{cases} 
-\frac{1}{2}x^2 + A, & \text{if } 0 < x < \xi; \\
-\frac{1}{2}x^2 + x + B & \text{if } \xi < x < 1.
\end{cases} \]  

(5.16)

This construction satisfies the boundary conditions at \( x = 0 \) and \( x = 1 \). Now we hope to determine the constants \( A \) and \( B \) from the jump conditions at \( x = \xi \). One of these conditions is that \( G(x, \xi) \) is continuous at \( x = \xi \), so that

\[ A = \xi + B. \]  

(5.17)

The other condition, obtained by integrating across the \( \delta \)-function is that

\[ G'(\xi^+, \xi) - G'(\xi^-, \xi) = 1. \]  

(5.18)

But when we apply this condition we discover that it is automatically satisfied by the solution in (5.16). Thus we have only one condition, namely (5.17), to determine the two constants \( A \) and \( B \). So the most general solution of (14) is that

\[ G(x, \xi) = A + \begin{cases} 
-\frac{1}{2}x^2 - \xi, & \text{if } 0 < x < \xi; \\
-\frac{1}{2}x^2 + x - \xi, & \text{if } \xi < x < 1.
\end{cases} \]  

(5.19)

Above, \( A \) is any constant. This seems strange, but the arbitrary constant \( A \) is expected because now, in retrospect, it is obvious that if \( y(x) \) is a solution of (5.9) then so is \( y(x) + A \). The Green’s function is telling us that the original boundary value problem has an infinite number of solutions. This example gives you a glimpse of the problems which are presented by inhomogeneous boundary value problems: sometimes there is no solution and other times there are an infinite number of solutions!
5.3 Problems

Problem 5.1. Find the Green’s function of the boundary value problem

\[ y'' = w(x), \quad y(0) = 0, \quad y'(1) = 0. \]

Give a simple formula for \( y'(0) \) in terms of \( w(x) \).

Problem 5.2. Find the Green’s function of

\[ y'' - y = f(x); \quad y(\pm\infty) = 0. \]

Problem 5.3. Find a Green’s function representation of the solution of the inhomogeneous boundary value problem

\[ \left[ x - \frac{d}{dx} \right]^2 y - y = f(x), \quad y(1) = y(\infty) = 0. \]

Problem 5.4. Consider the boundary value problem

\[ y'' = w(x), \quad \text{with} \ y(0) = 0 \text{ and} \ y'(1) = \alpha y(1); \]

the boundary condition at \( x = 1 \) is ‘mixed’. Find the Green’s function and show that there is a particular value of \( \alpha \) for which the problem has no solution for an arbitrary \( w(x) \). Considering this exceptional case, find the “solvability condition” — the analog of (5.14) — which ensures that the problem does have a solution and then obtain the appropriate Green’s function.

Problem 5.5. Find solutions of the differential equations

\[ y''' + y = \delta(x), \]

\[ \left[ \frac{d^2}{dx^2} - 1 \right]^2 y = \delta'(x), \]

which vanish at \( x = \pm\infty \). In both cases give an explicit real expression for \( y(0) \). In both cases calculate

\[ \int_{-\infty}^{\infty} y(x) \, dx, \quad \text{and} \quad \int_{-\infty}^{\infty} xy(x) \, dx. \]

(look for a way to evaluate the integrals without having to explicitly solve the equation.)
Lecture 6

The diffusion equation

\[ u_t = \kappa \nabla^2 u \]

6.1 Origin of the diffusion equation

The equation governing heat diffusion in a solid is

\[ \rho c u_t = \nabla \cdot (k \nabla u) \]  \hspace{1cm} (6.1)

where \( u(x, t) \) is the temperature. Also in (6.1), where \( \rho \) is the density of the solid (kilograms per cubic meter), \( c \) the heat capacity (Joules per kilogram per Kelvin) and \( k \) the conductivity (find the units yourself). This is in the form of a conservation law

\[ \partial_t [\text{energy density}] + \nabla \cdot [\text{energy flux}] = 0. \]  \hspace{1cm} (6.2)

The flux is \( f = -k \nabla u \), so if the temperature is uniform then there is no flux of energy. In the 17th century, people thought of heat as “caloric fluid” that flows down the temperature gradient, from hot regions to colder regions. This caloric theory has been superseded by thermodynamics, but I think it still provides some useful intuition.

Notice that if we integrate (6.2) over a closed volume \( V \), with surface \( \partial V \), then with Gauss’s theorem we have

\[ \frac{d}{dt} \int_V \rho c u \, dv = \int_{\partial V} k \nabla u \cdot n \, ds, \]  \hspace{1cm} (6.3)

energy in \( V \)  \hspace{1cm} energy flow through the surface \( \partial V \)
Lecture 6. The diffusion equation $u_t = \kappa \nabla^2 u$

where $\mathbf{n}$ is the unit outward normal to the surface of $\partial V$. This is the multi-dimensional analog of an integral conservation law.

If all three of the three material properties $\rho$, $c$ and $k$ are constant we can bundle them into one combination,

$$\kappa \equiv \frac{k}{\rho c}, \quad (6.4)$$

called the diffusivity. Then the simplest form of the diffusion equation is

$$u_t = \kappa \nabla^2 u. \quad (6.5)$$

**Random walkers**

Another context in which the diffusion equation arises is the random walk. In the one-dimensional case, $u(x,t)$ is the density (walkers per meter) and $\dot{f} = -\kappa u_x$ is the flux (walkers per second). In this example

$$\kappa = \frac{\text{(mean square step length)}^2}{2 \times \text{average time between steps}}. \quad (6.6)$$

Let’s sketch the random walk model and explain how it produces the diffusion equation.

We begin with the classical model in which the random walk is discrete in both space and time. The walker (traditionally a drunk) takes a step at times $t = \tau, 2\tau, \cdots$. The steps are $\pm a$ with probability 1/2. We consider an ensemble of walkers and let $\rho(x,t)$ be the number of walkers at site $x$ at time $t$. It is easy to see that the evolution of the ensemble of walkers is given by

$$\rho(t,x) = \frac{1}{2} \rho(t-\tau, x-a) + \frac{1}{2} \rho(t-\tau, x+a). \quad (6.7)$$

i.e. half of the walkers at site $x-a$ hop to the right and land at $x$ and so on. Equation (6.7) is a partial difference equation i.e., a discrete system with two independent variables. With an initial condition, such as

$$\rho(x,0) = \delta(x), \quad (6.8)$$

it is easy to iterate a few times and find $\rho(x,\tau), \rho(x,2\tau)$ etc. You might like to do this and see if you can guess the general result (hint: Pascal’s triangle).
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

We can obtain the diffusion equation (6.5) as an approximation to (6.7) when \( \rho(x,t) \) is slowly changing i.e., when the spatial variation of \( \rho \) is on a scale much larger than the hopping length \( a \), so that in one iteration from \( t \) to \( t + \tau \) there is only a small change in \( \rho(x,t) \). This “macroscopic evolution equation” follows easily from (6.7) if we rewrite it as

\[
\rho(t,x) - \rho(t-\tau,x) = \frac{1}{2} [\rho(t-\tau,x-a) - 2\rho(t-\tau,x) + \rho(t-\tau,x+a)] .
\]

(6.9)

Now we make the slowly varying assumption: we expand everything in sight in the small parameters \( a \) and \( \tau \):

\[
\rho(t-\tau,x) = \rho(t,x) - \tau \rho_t(t,x) + O(\tau^2) ,
\]

(6.10)

and

\[
\rho(t-\tau,x \pm a) = \rho(t,x) - \tau \rho_t(t,x) \pm a \rho_x(t,x) + \frac{1}{2} a^2 \rho_{xx}(t,x) + O(a^3) .
\]

(6.11)

Putting all this into (6.9) and isolating the surviving term on each side we get

\[
\rho_t = \kappa \rho_{xx}
\]

(6.12)

where

\[
\kappa \equiv \frac{a^2}{2\tau} .
\]

(6.13)

Why do we keep the \( O(a^2) \) terms in (6.11), and neglect the \( O(\tau^2) \) terms in (6.10)? This seems inconsistent. Neglect of the \( O(\tau^2) \) terms can be justified by taking a distinguished limit in which \( a \to 0 \) and \( \tau \to 0 \) with ratio \( \kappa \) in (6.13) fixed; with this procedure \( \tau \) and \( a^2 \) have the same order. In taking this limit we suppose that instead of (6.8) we have a smooth initial condition containing a length scale \( \ell \):

\[
\rho(x,0) = f \left( \frac{x}{\ell} \right) ,
\]

(6.14)

and that \( \ell \) is fixed as \( a \) and \( \tau \) go to zero with \( \kappa \) fixed. In this limit, the term \( \tau \rho_{tt} \) — and all the other higher-order terms such as \( a^4 \rho_{xxxx} \) — disappear.
6.2 Smoothing a discontinuity or “front”

Considering the problem
\[ u_t = \kappa u_{xx}, \quad u(x,0) = A \text{ sgn}(x), \quad (6.15) \]
we quickly see that the only dimensionally consistent form for the solution is
\[ u(x,t) = A \times \text{some function of } x/\sqrt{\kappa t}. \quad (6.16) \]
Thus we are inspired to try the guess
\[ u(x,t) = AU(\xi), \quad \xi \equiv x/\sqrt{\kappa t}. \quad (6.17) \]

We start calculating
\[ u_x = \frac{A}{\sqrt{\kappa t}} U', \quad u_{xx} = \frac{A}{\kappa t} U'', \quad u_t = -\frac{A}{2t} \xi U' \quad (6.18) \]
and substitute into the PDE:
\[ -\frac{A}{2t} \xi U' = \kappa \frac{A}{\kappa t} U'' \quad \Rightarrow \quad -\xi U' = 2U''. \quad (6.19) \]
Our efforts are crowned with success when all the t’s cancel and we are left with the boxed ODE for \( U(\xi) \).

Now we solve that ODE using separation of variables. Integrating twice
\[ U' = C_1 e^{-\xi^2/4}, \quad U(\xi) = C_1 \int_0^\xi e^{-v^2/4} \, dv + C_2. \quad (6.20) \]
The constants \( C_1 \) and \( C_2 \) are determined by insisting that
\[ \lim_{\xi \to \pm \infty} U(\xi) = \pm 1, \quad (6.21) \]
or:
\[ -1 = -\sqrt{\pi} C_1 + C_2, \quad 1 = \sqrt{\pi} C_1 + C_2, \quad \Rightarrow \quad C_1 = \frac{1}{\sqrt{\pi}}, \quad C_2 = 0. \quad (6.22) \]
Thus we now have
\[ U(\xi) = \frac{1}{\sqrt{\pi}} \int_0^\xi e^{-v^2/4} \, dv. \quad (6.23) \]
Adjusting our notation, we rewrite the solution in (6.23) as
\[ u(x,t) = \text{erf}(\eta), \quad (6.24) \]
Lecture 6. The diffusion equation $u_t = \kappa \nabla^2 u$

where the similarity variable is now

$$\eta \equiv \frac{x}{2\sqrt{\kappa t}}.$$  \hfill (6.25)

The erf function is shown in figure 6.1.

6.3 Discussion of the similarity method

This similarity method is a powerful technique for solving or simplifying both linear and nonlinear PDE’s. Above I have also demonstrated the powerful educational technique of glossing quickly over the profound part of the method — the guess in (6.17) — and dwelling at length on algebraic technicalities.

To explain the similarity method in more detail suppose that Tom and Jerry are both struggling with (6.15). Tom is measuring length in meters and seconds, but Jerry is an engineer and prefers to use furlongs and fortnights. Tom uses $x$ and $t$ to denote space and time and Jerry uses $x'$ and $t'$. The numerical values of these coordinates can be converted one to the other by

$$x' = \lambda x, \quad t' = \mu t.$$  \hfill (6.26)

It follows that Tom and Jerry’s numerical values for the diffusivity are related by $\kappa' = \lambda^2 \kappa / \mu$.

Even before Tom & Jerry solve the PDE they realize that the solution must satisfy

$$u = AF(x, t, \kappa) = AF(\lambda x, \mu t, \lambda^2 \kappa / \mu).$$  \hfill (6.27)
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

All that we have done is assume the existence of a formula which gives \( u \) once either set of variables is specified. The essential point is that this is the same formula no matter which set of units is used. Then as we freely vary both \( \lambda \) and \( \mu \) the value of \( u \) at the same point in space-time is unchanged i.e., (6.27) is always true for all values of \( \mu \) and \( \lambda \). This implies that \( x, t \) and \( \kappa \) appear in \( F \) only in the combination \( x^2 / (\kappa t) = x'^2 / (\kappa' t') \) — that’s where we started back in (6.17).

One difficulty with the argument above is that the initial condition can never be totally discontinuous. Suppose the initial condition is really something like

\[
\begin{align*}
  u(x, 0) &= A \tanh(x / \ell) = A \tanh(x' / \ell') .
\end{align*}
\]

(6.28)
i.e., some smooth function which changes from \(-A\) to \(+A\) over a distance \( \ell' = \mu \ell \). As soon as Tom and Jerry admit the existence of this length they must also amend (6.27) to

\[
\begin{align*}
  u &= AF(x, t, \kappa, \ell) = AF(\lambda x, \mu t, \mu \kappa / \lambda^2, \lambda \ell) .
\end{align*}
\]

(6.29)
Now as one freely varies \( \lambda \) and \( \mu \) the most that one can deduce is that

\[
\begin{align*}
  u &= AF(x / \ell, \kappa t / \ell^2) = AF(x' / \ell', \kappa' t' / \ell'^2) .
\end{align*}
\]

(6.30)
This is a useful result because it tells us that we can economically present the solution using a function of two (rather than four) variables. But there is no longer a similarity solution.

On the other hand, it seems reasonable that if we consider a sequence of problems with smaller and smaller \( \ell' \)’s then in the limit we should approach the \( \text{sgn}(x) \) initial condition and the erf similarity solution. Alternatively, with fixed \( \ell \), once the diffusively spreading front becomes much thicker than \( \ell \) then \( \ell \) is an “irrelevant” parameter. In this case we get the erf solution as the asymptotic description of a diffusing front. This possibility is consistent with (6.29) if

\[
\begin{align*}
  (p, q) \to \infty \quad \Rightarrow \quad F(p, q) \to \tilde{F}(p / \sqrt{q}) .
\end{align*}
\]

(6.31)

### 6.4 The Green’s function for diffusion on the line

To solve the initial value problem

\[
\begin{align*}
  u_t &= \kappa u_{xx} , \quad u(x, 0) = f(x) ,
\end{align*}
\]

(6.32)
Lecture 6. The diffusion equation $u_t = \kappa \nabla^2 u$

we need the Green’s function, which is defined by

$$g_t = \kappa g_{xx}, \quad g(x, 0) = \delta(x). \quad (6.33)$$

Once we possess $g(x, t)$ it is easy to see by substitution that

$$u(x, t) = \int_{-\infty}^{\infty} f(x')g(x - x', t) \, dx' \quad (6.34)$$

is the solution of (6.32). To find $g(x, t)$ we use the erf-solution as the thin end of the wedge: since

$$\delta(x) = \frac{1}{2} \frac{d}{dx} \text{sgn}(x), \quad (6.35)$$

the Green’s function of the diffusion equation is

$$g(x, t) = \frac{1}{2} \frac{\partial}{\partial x} \text{erf}(\eta),$$

$$= \frac{e^{-x^2/4\kappa t}}{\sqrt{4\pi \kappa t}}. \quad \text{BYU} \quad (6.36)$$

Another derivation of the diffusion Green’s function

An alternative route to (6.36) is to argue on dimensional grounds that

$$g(x, t) = \frac{1}{\sqrt{\kappa t}} G \left( \frac{x}{\sqrt{\kappa t}} \right). \quad (6.37)$$

As an exercise you should substitute this guess into the diffusion equation and recover (6.36) by solving the resulting ODE.

Where does the factor $1/\sqrt{\kappa t}$ on the RHS of (6.37) come from? Apparently

$$\text{dim}(g) = \text{length}^{-1}. \quad (6.38)$$

Because of the initial condition in (6.33) the dimensions of $g$ are the same as the dimensions of $\delta(x)$. Now $\delta$-functions have the peculiar property that

$$\text{dim} [\delta(\theta)] = \frac{1}{\text{dim}(\theta)}. \quad (6.39)$$

That is, $\delta$-functions have the inverse dimension of their argument. (Other functions, such as $\cos x$ or $\ln t$, are dimensionless.) In the problem
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

(6.33), the argument of the \( \delta \)-function is \( x \), with dimensions of “length”, and therefore \( g \) has dimensions \((\text{length})^{-1}\).

Another way to understand the factor \( 1/\sqrt{\kappa t} \) is to observe that if we integrate the diffusion equation from \( x = -\infty \) to \( x = \infty \) we have:

\[
\frac{d}{dt} \int_{-\infty}^{\infty} g(x, t) \, dx = [\kappa g_x(x, t)]_{x=-\infty}^{x=\infty} = 0.
\]  

(6.40)

(We assume the solution decays quickly as \( x \to \pm \infty \) so that there is no flux of \( g \)-stuff from \( \pm \infty \).) Using the initial condition we deduce that

\[
\int_{-\infty}^{\infty} g(x, t) \, dx = 1.
\]  

(6.41)

Now the factor \( 1/\sqrt{\kappa t} \) is just what the similarity guess in (6.37) needs in order to the satisfy the conservation law (6.41):

\[
1 = \int_{-\infty}^{\infty} G \left( \frac{x}{\sqrt{\kappa t}} \right) \frac{dx}{\sqrt{\kappa t}} = \int_{-\infty}^{\infty} G(\xi) \, d\xi.
\]  

(6.42)

This little argument based on conservation laws also tells us how to determine a constant of integration which arises as we solve the ODE which results from stuffing (6.37) into the diffusion equation.

**Convolution**

In (6.34) we have the solution of the diffusion equation as the convolution of the initial condition \( f(x) \) with the Green’s function \( g(x, t) \). Let us discuss the convolution of two functions in more detail. Alternative terms used to describe

\[
c(x) = \int_{-\infty}^{\infty} a(x') b(x - x') \, dx' = \int_{-\infty}^{\infty} a(x - x') b(x') \, dx',
\]  

(6.43)

are faltung, superposition integral, running mean, smoothing, scanning and blurring.

To appreciate the origin of these terms, consider the case when \( b(x) \) in (6.43) is a “box-car filter”

\[
b(x) = \begin{cases} 
1, & \text{if } |x| < \frac{1}{2}, \\
0, & \text{otherwise.}
\end{cases}
\]  

(6.44)
Then at every point \( x \), \( c(x) \) is an averaged, or smoothed or low-pass filtered version of \( a(x) \). Explaining this requires drawing some examples, such as visually computing the convolution

\[
c(x) = \int_{-\infty}^{\infty} \text{sgn}(x') b(x - x') \, dx' = \begin{cases} 
-1, & \text{if } x < -\frac{1}{2}, \\
2x, & \text{if } -\frac{1}{2} \leq x \leq \frac{1}{2}, \\
+1, & \text{if } \frac{1}{2} < x.
\end{cases} \tag{6.45}
\]

The box-car filter smooths out the sudden jump at \( x = 0 \) into a ramp.

The convolution in (6.34) shows that the value of \( u(x,t) \) is determined by averaging the initial condition of \( f(x) \) over an interval of width \( \sqrt{\kappa t} \) centered on the point \( x \). The Green’s function (6.36) is a Gaussian filter, and the range of the filter grows as \( \sqrt{\kappa t} \). Thus as \( t \) increases the width of the averaging interval increases, and therefore \( u(x,t) \) becomes increasingly smooth. The effect of convolving \( \text{sgn}(x) \) with the box-car is the ramp in (6.45) and convolution of \( \text{sgn}(x) \) with the Gaussian is the erf in Figure 6.1. The two results are qualitatively similar: the discontinuity is smeared out over a distance comparable to the filter width.

### 6.5 Solution of the diffusive initial value problem

Having the general solution of the infinite-line diffusion problem in (6.34) and (6.36) is very pleasant — particularly because the interpretation as a convolution-average provides intuition. But be aware that (6.34) may not be the easiest way to solve simple problems. For example, consider the initial value problem

\[
\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2}, \quad u(x,0) = \cos kx. \tag{6.46}
\]

Guessing that the solution is separable

\[
u = A(t) \cos kx, \tag{6.47}
\]

we quickly find by substitution that

\[
\frac{dA}{dt} = -\kappa k^2 A, \quad \Rightarrow \quad u(x,t) = e^{-\kappa k^2 t} \cos kx. \tag{6.48}
\]

If we can convince ourselves that the diffusion equation has a unique solution, then we have shown

\[
e^{-\kappa k^2 t} \cos kx = \int_{-\infty}^{\infty} \cos kx' \frac{e^{-\frac{(x-x')^2}{4\kappa t}}}{\sqrt{4\pi\kappa t}} \, dx'. \tag{6.49}
\]

---

SIO203C, W.R. Young, March 21, 2011
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

**Example:** Solve the initial value problem

\[
    u_t = \kappa u_{xx}, \quad u(x, 0) = e^{-m^2x^2}.
\]  

(6.50)

We can guess that if we release a suitably scaled \( \delta \)-pulse at \( t_0 < 0 \) then at \( t = 0 \) this pulse will have evolved into the Gaussian above. So consider

\[
    u_t = \kappa u_{xx}, \quad u(x, t_0) = A \delta(x),
\]

(6.51)

with solution

\[
    u = \frac{A e^{-x^2/4\kappa(t-t_0)}}{\sqrt{4\pi \kappa(t-t_0)}}. \tag{6.52}
\]

Evaluating this at \( t = 0 \) we obtain \( t_0 = -1/(4m^2\kappa) \), and we also determine the amplitude \( A \). The final answer is

\[
    u(x, t) = \frac{e^{-m^2x^2/(1+4m^2\kappa t)}}{\sqrt{1+4m^2\kappa t}} \tag{6.53}
\]

**Example:** Solve the initial value problem

\[
    u_t = \kappa u_{xx}, \quad u(x, 0) = xe^{-m^2x^2}.
\]

(6.54)

If \( u \) is a solution of the diffusion equation, then so is \( u_x \). So, noting that the initial condition in this example is

\[
    xe^{-m^2x^2} = -\frac{1}{2m^2} \frac{d}{dx} e^{-m^2x^2}, \tag{6.55}
\]

we can employ the solution of the previous example to conclude that

\[
    u(x, t) = \frac{1}{2m^2} \frac{d}{dx} \frac{e^{-m^2x^2/(1+4m^2\kappa t)}}{\sqrt{1+4m^2\kappa t}} = \frac{xe^{-m^2x^2/(1+4m^2\kappa t)}}{(1+4m^2\kappa t)^{3/2}} \tag{6.56}
\]

**New solutions from old**

In the examples above we’ve used some quasi-obvious properties of the diffusion equation to manufacture new solutions from old solutions. Just to be clear, let’s summarize all these tricks. The most important observation is that the diffusion equation is linear: if \( u_1 \) and \( u_2 \) are solutions then so is an arbitrary linear combination \( a_1 u_1 + a_2 u_2 \).

Moreover, if \( u(x, t) \) is a solution of the diffusion equation then:

1. so is \( u(-x, t) \), and so is \( u(x + x_0, t + t_0) \) where \( x_0 \) and \( t_0 \) are arbitrary constants;
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

2. so is \( \partial_x^m u(x,t) \), and so is \( \partial_t^n u(x,t) \);

3. so is \( u(\alpha x, \alpha^2 t) \), where \( \alpha \) is an arbitrary constant.

Under the operations above the diffusion equation is transformed into itself.

The transformational properties of differential equations is an important field of mathematics, initiated by Lie over a century ago. This is not the main focus of these lectures. But just to show that Lie group methods lead to non-obvious results, I remark that the transformation

\[
\begin{align*}
    t' &= t + c, \\
    x' &= x - 2ct - c^2, \\
    u' &= u e^{cx - c^2 t - c^3/3} 
\end{align*}
\]  

transforms the diffusion equation to itself:

\[
    u_t = \kappa u_{xx} \quad \Rightarrow \quad u'_t = \kappa u'_{x'x'}. 
\]  

There is nothing quasi-obvious about this transformation.

**Uniqueness**

To prove that the diffusion equation has a unique solution, suppose to the contrary that the initial value problem

\[
    v_t = \kappa \nabla^2 v, \quad v(x,0) = f(x). \tag{6.59}
\]

has two solutions, \( v_1(x,t) \) and \( v_2(x,t) \) say. Then the difference \( u = v_1 - v_2 \) also satisfies the diffusion equation, but with the initial condition

\[
    u(x,0) = 0. \tag{6.60}
\]

Now let’s proceed assuming that \( u \) might be non-zero and show that this leads to a contradiction. Multiply the diffusion equation by \( u \) and shuffle the result into the following form

\[
    \frac{1}{2} \left( u^2 \right)_t = \kappa \nabla^2 \left( u^2 \right) - \kappa |\nabla u|^2. \tag{6.61}
\]

Notice that although \( u \) satisfies a conservation equation, \( u^2 \)-stuff is not conserved: \( u^2 \) is destroyed at a rate \( \kappa |\nabla u|^2 \). Integrating over the entire space\footnote{OK — there is an assumption of rapid decay at infinity so that \( \int u \nabla u \cdot \mathbf{n} \, dA \) can be dismissed. In finite domain with homogeneous BCs there is no assumption necessary.} we have

\[
    \frac{1}{2} \frac{d}{dt} \int u^2 \, dV = -\kappa \int |\nabla u|^2 \, dV. \tag{6.62}
\]
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

Now integrate with respect to time

\[
\frac{1}{2} \int u^2 \, dV = -\kappa \int_0^t \int |\nabla u|^2 \, dV \, dt ; \tag{6.63}
\]

This is a contradiction: the left is positive and the right is negative.

The moment method

Even when guessing doesn’t work, one can still extract very useful information about the solution without actually solving the diffusion equation.

Suppose, for example, that the initial condition in (6.32) is a hump, such as \( f(x) = \text{sech} x \). We can use (6.34) and write down an integral representation of the solution immediately. But if we are content with less than total information we can get a good qualitative feel by considering the moments of the solution:

\[
m_p(t) = \int_{-\infty}^{\infty} x^p u(x, t) \, dx . \tag{6.64}
\]

If we think of \( u(x, t) \) as a density then \( m_0 \) is the total mass and

\[X(t) \equiv \frac{m_1}{m_0}\]

is the center of mass. The variance or moment of inertia is

\[
\sigma^2 \equiv m_0^{-1} \int_{-\infty}^{\infty} (x - X)^2 u(x, t) \, dx = (m_2 / m_0) - X^2 . \tag{6.66}
\]

If we know \( m_0, X \) and \( \sigma \) then we know how much stuff there is, where the stuff is, and how far its spread out.

Multiplying (6.32) by \( x^p \), with \( p = 0, 1 \) and \( 2 \), and integrating from \( x = -\infty \) to \( x = \infty \) we see

\[
m_0 = 0, \quad m_1 = 0, \quad m_2 = 2\kappa m_0 . \tag{6.67}
\]

The proof is integration by parts; the key intermediate step is

\[
\int_{-\infty}^{\infty} x^2 u_{xx} \, dx = [x^2 u_x]_{-\infty}^{\infty} - 2xu \bigg|_{-\infty}^{\infty} + 2 \int_{-\infty}^{\infty} u \, dx . \tag{6.68}
\]
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \)

Substituting (6.67) into (6.65) and (6.66) we find that both \( m_0 \) and \( X \) are constant while
\[
\sigma^2 = \sigma_0^2 + 2\kappa t. \tag{6.69}
\]
Thus diffusion leaves the mass and the center of mass fixed but increases the width of the hump linearly with time. This gives us a good qualitative picture for the evolution of an initial pulse of heat.

The maximum principle

We can also show that if \( u(x, t) \) is a solution of (6.32) then the maximum (and minimum) of \( u(x, t) \) for all \( x \) and \( t \geq 0 \) is at \( t = 0 \). That is, the largest possible value of \( u \) is present in the initial condition \( f(x) \).

We can deduce this maximum principle from the Green’s function solution (6.34). This expresses \( u \) at the point \((x, t)\) as an average of \( f \) over an interval of width \( \sqrt{\kappa t} \) centered on \( x \). If \( u \) had a maximum at \( x \) with \( t > 0 \), then this \( u \) would be larger than every value of \( f \) contributing to the average. This is impossible.

Another proof is by contradiction: suppose that there is an isolated maximum of \( u(x, t) \) at some time \( t > 0 \). Then at this point
\[
u_x = u_t = 0, \quad \text{and} \quad u_{xx} < 0. \tag{6.70}
\]
The conditions above are inconsistent with \( u(x, t) \) being a solution of the diffusion equation.

6.6 Diffusion with a source

Consider the forced heat equation
\[
u_t - \kappa u_{xx} = s(x, t), \quad u(x, 0) = f(x). \tag{6.71}
\]
Without any loss of generality we can take \( f(x) = 0 \) — we can always use linear superposition to add on the part of the solution due to the initial condition.

If we knew the Green’s function
\[
G_t - \kappa G_{xx} = \delta(x)\delta(t) \tag{6.72}
\]
then we could instantly write the solution of (6.72) as
\[ u(x,t) = \int_{-\infty}^{\infty} dx' \int_{0}^{\infty} dt' \, G(x - x', t - t') s(x', t') . \] (6.73)

To determine \( G(x, t) \) we integrate (6.72) from \( t = 0^- \) to \( t = 0^+ \):
\[ G(x, 0^+) - G(x, 0^-) - \int_{0^-}^{0^+} G_{xx}(x, t) \, dt \approx \delta(x) , \] (6.74)
and conclude that
\[ G(x, 0^+) = \delta(x) . \] (6.75)

In other words, the inhomogeneous Green’s function problem is equivalent to the initial value problem solved back in (6.33). Thus we have wasted a perfectly good symbol, \( G \), because:
\[ G(x, t) = g(x, t) = \frac{e^{-x^2/4\kappa t}}{\sqrt{4\kappa \pi t}} H(t) . \] (6.76)

We add the Heaviside factor \( H(t) \) to remind us that the Green’s function is zero if \( t < 0 \). The Green’s function \( g(x, t) \) does double-duty: the solution of (6.71) is
\[ u(x, t) = \int_{-\infty}^{\infty} f(x') g(x - x', t) \, dx' + \int_{-\infty}^{\infty} dx' \int_{0}^{t} dt' \, s(x', t') g(x - x', t - t') . \] (6.77)

**Example:** Solve the forced diffusion equation
\[ v_t = \kappa v_{xx} + \exp(-m^2 x^2) , \quad v(x, 0) = 0 . \] (6.78)

Let’s save the formula (6.77) for exercise [6.9] Instead, notice that in an earlier example we considered the initial value problem
\[ u_t = \kappa u_{xx} , \quad u(x, 0) = \exp(-m^2 x^2) , \] (6.79)
and obtained the solution in (6.53). Now you can easily show by substitution that
\[ v(x, t) = \int_{0}^{t} u(x, t') \, dt' = \int_{0}^{t} \frac{e^{-m^2 x^2/(1+4m^2 \kappa t')}}{\sqrt{1 + 4m^2 \kappa t'}} \, dt' \] (6.80)
Lecture 6. The diffusion equation \( u_t = \kappa \nabla^2 u \) solves (6.78). We beat this into a standard form involving an exponential integral with the substitution

\[
z = \frac{m^2 x^2}{1 + 4m^2 \kappa t'}, \quad dt' = -\frac{x^2}{4\kappa} \frac{dz}{z^2}; \quad (6.81)
\]

thus

\[
v(x, t) = \frac{|x|}{4m\kappa} \int_{m^2 x^2 / 1 + 4m^2 \kappa t}^{m^2 x^2 / (1 + 4m^2 \kappa t')} e^{-z} \frac{dz}{z^{3/2}}. \quad (6.82)
\]

6.7 Specialized references to diffusion

For an interesting perspective on diffusion problems see

**Be** Random Walks in Biology by H.C. Berg.

A classic handbook on diffusion is


A recent book emphasizing probability theory is

**Re** A Guide to First-Passage Processes by S. Redner.

For similarity solutions, with lots of physical examples, see

**Ba** Scaling, Self-similarity, and Intermediate Asymptotics by G.I. Barenblatt.

For Lie group methods see

Lecture 6. The diffusion equation $u_t = \kappa \nabla^2 u$

Figure 6.2: The solution of problem 6.2. diffutz.eps

6.8 Problems

Problem 6.1. Reconsider the hopping model in (6.7), assuming that the drunk hops to the right with probability $p$ and to the left with probability $q$, where $p + q = 1$. If $p > q$ then there is a bias towards the right. Find a distinguished limit in which the macroscopic evolution of the slowly varying density $\rho(x, t)$ is determined by the advection-diffusion equation

$$\rho_t + c \rho_x = \kappa \rho_{xx}. \quad (6.83)$$

Your answer should include expressions for $c$ and $\kappa$ in terms of $\tau$, $a$, $p$ and $q$.

Problem 6.2. Write a MATLAB program to iterate (6.7) with $\tau = a = 1$. The initial condition is

$$\rho(x, 0) = \text{sgn}(x),$$

where $x = [-40 : 1 : 40]$. As boundary conditions take $\rho(\pm40) = \pm1$. Compare your numerical solution of the discrete system (6.7) with the appropriate erf similarity solution at $t = 1, 4, 16, 64$ and 256.
Lecture 6. The diffusion equation $u_t = \kappa \nabla^2 u$

The most boring part of this problem is getting MATLAB to subplot at the right times and in the right place. Here’s some of my code showing how this drudgery is handled. The educational part of the problem is to fill in the ???’s.

```matlab
x = [-40:1:40]; % The discrete grid
xx = [-40:0.25:40]; % For plotting the erf solution
u = sign(x);
n = length(u);

% The discrete grid
xx = [-40:0.25:40]; % For plotting the erf solution
u = sign(x);
n = length(u);

n = length(u);
tplot = [1 4 64 256];
nPlot = 0;
for t=[1:1:256]
    %?????????
    %?????????
    %?????????
    if t==4^(nPlot)
      nPlot = nPlot + 1;
      subplot(5,1,nPlot)
      plot(x,u,'.','markersize',10)
      hold on
      plot(xx,erf(????))
      time = num2str(t)
      text(-15,0.4,['t='time])
    end
end
```

Problem 6.3. Obtain the Green’s function for the damped advection-diffusion equation by solving

$$g_t + c g_x = \kappa g_{xx} - \beta g, \quad g(x,0) = \delta(x).$$

**Hint:** transform the damped advection-diffusion equation to the plain old diffusion equation and use results from the lecture.

Problem 6.4. Apply the method of moments to the initial value problem

$$\rho_t + c \rho_x = \kappa \rho_{xx} - \beta \rho, \quad \rho(x,0) = \rho_0(x).$$

The initial condition is a compact humplike density e.g., something like $\rho_0(x) = \text{sech}(x)$. Obtain the mass, center of mass and the width of the hump as functions of time.

Problem 6.5. Fill in all the steps between the similarity guess (6.37) and the solution in (6.36). Make sure you explain carefully how the similarity ansatz is consistent with the initial condition $g(x,0) = \delta(x)$. 

SIO203C, W.R. Young, March 21, 2011
Lecture 6. The diffusion equation $u_t = \kappa \nabla^2 u$

Figure 6.3: Diffusive smoothing of the initial condition in problem 6.7.

Problem 6.6. Find a similarity solution of the nonlinear diffusion equation

$$u_t = \kappa (u^m)_{xx}, \quad u(x, 0) = \delta(x).$$

Hint: The answer is in section 51 of Landau & Lifshitz Fluid Mechanics.

Problem 6.7. Consider the diffusion equation $u_t = \kappa u_{xx}$ with the initial condition

$$u(x, t) = \begin{cases} 
1, & \text{if } |x| < a; \\
0, & \text{otherwise}.
\end{cases}$$

(i) Define nondimensional variables so that $(a, \kappa) \to 1$. (ii) Use linearity and the erf-solution in (6.24) to solve this initial value problem. (iii) Use MATLAB to illustrate your solution (see figure 6.3).

Problem 6.8. Solve the forced diffusion equation

$$u_t - \kappa u_{xx} = \cos kx, \quad u(x, 0) = 0.$$
Problem 6.9. (i) As an exercise in Gaussian integrals, use the formula (6.77) to show that the solution of
\[ u_t = \kappa u_{xx} + h(t)m e^{-m^2x^2} \frac{1}{\sqrt{\pi}} , \quad u(x,0) = 0, \]
is
\[ u(x,t) = \frac{m}{\sqrt{\pi}} \int_0^t h(t') e^{-m^2x^2/(1+4m^2\kappa(t-t'))} \frac{1}{\sqrt{1+4\kappa m^2(t-t')}} dt'. \]
Above, \( h(t) \) is some function of time. (ii) Consider the limit \( m \to \infty \). Show that the heat source is \( h(t)\delta(x) \), and that the temperature at the source is
\[ u(0,t) = \frac{1}{2\sqrt{\pi \kappa}} \int_0^t h(t') \frac{1}{\sqrt{t-t'}} dt'. \]
(iii) Find a heating function \( h(t) \) that makes \( u(0,t) \) independent of time.

Problem 6.10. (i) Show that the similarity guess
\[ u(x,t) = t^a f(\eta), \quad \eta = x/2\sqrt{\kappa t} \]
a possible solution of the problem
\[ u_t - \kappa u_{xx} = \delta(x)qt^{q-1}, \quad u(x,0) = 0, \quad (q > 0), \]
only if \( q = a + \frac{1}{2} \). (ii) Substitute the guess into the PDE and show that \( f' \) satisfies the ODE
\[ f'' + 2\eta f' - 4af = 0. \]
(iii) Reduce the ODE above to quadratures in the special case \( q = 1 \).

Problem 6.11. A function \( F(x,y) \) is said to be homogeneous of degree \( n \) if
\[ F(\lambda x, \lambda y) = \lambda^n F(x,y), \]
for any constant \( \lambda \). For example, \( F = x/y \) is homogeneous with \( n = 0 \). (i) Give examples of homogeneous functions of degree \( n = 0 \), 1 and 2. (ii) Obtain a PDE satisfied by \( F(x,y) \) and solve this equation in terms of an 'arbitrary function'. [Hint: differentiate with respect to \( \lambda \) and then set \( \lambda = 1 \).]

Problem 6.12. After (6.27) I said “As we freely vary both \( \lambda \) and \( \mu \) the value of \( u \) at the same point in space-time is unchanged i.e., (6.27) is always true. This implies that \( x, t \) and \( \kappa \) appear only in the combination \( x^2/(\kappa t) = x'^2/(\kappa' t') \).” Use the method of problem 6.11 to prove this assertion.
Problem 6.13. (i) Find the exponents $a$ and $b$ which make the similarity guess

$$u(x,t) = t^a v(\eta), \quad \eta \equiv x/(\nu t)^b$$

a possible solution of the dispersive wave equation

$$u_t = \nu u_{xxx}, \quad u(x,0) = \delta(x).$$

(ii) Substitute the guess into the PDE and find the ODE satisfied by $v(\eta)$.
(iii) Solve the ODE in terms of a special function you might find in the appendix of BO.

Problem 6.14. Consider the PDE

$$u_t = \kappa u_{xxx}, \quad u(x,0) = A \text{sgn}(x).$$

What are the dimensions of $\kappa$? Construct a similarity solution analogous to (6.17), and find the ODE satisfied by $U(\xi)$. Brownie points for solving this ODE.

Problem 6.15. Suppose you are given the solution, $u(x,t)$, of the homogeneous initial value problem:

$$u_t = u_{xxx}, \quad u(x,0) = e^{-x^2}.$$  

(i) Express the solution of the forced initial value problem

$$v_t = v_{xxx} + e^{-x^2}, \quad v(x,0) = 0,$$

in terms of $u(x,t)$. (ii) Suppose that $\chi(t) = 1$ when $0 < t < 1,$ and $\chi(t) = 0$ otherwise. Express the solution of the forced initial value problem

$$w_t = w_{xxx} + e^{-x^2} \chi(t), \quad w(x,0) = 0,$$

in terms of $v(x,t)$. (iii) Express the solution of the forced problem

$$h_t = h_{xxx} - 2x e^{-x^2} \delta(t), \quad h(x,t < 0) = 0,$$

in terms of $u$. 

SIO203C, W.R. Young, March 21, 2011
Lecture 7

Diffusion on a half-line and on a finite domain

In the previous lecture we discussed diffusion on the whole line \(-\infty < x < \infty\). Now let’s turn to the simplest diffusion problem with a boundary, namely the half-line \(x > 0\).

7.1 Half-line problems and images

Consider the problem of drunks random walking along the half-line \(x > 0\). There is a manhole at \(x = 0\) and so the drunks who reach \(x = 0\) immediately disappear. This is equivalent to the diffusion problem

\[
\frac{du}{dt} = \kappa u_{xx} \quad u(x,0) = f(x), \quad u(0,t) = 0,
\]

(7.1)

where the dimensions of \(u\) are drunks per meter. There is an absorption boundary condition at \(x = 0\).

The number of surviving drunks at time \(t\) is

\[
S(t) = \int_0^\infty u(x,t) \, dx,
\]

(7.2)

and it follows from (7.1) that

\[
\frac{dS}{dt} = -\kappa u_x(0,t).
\]

(7.3)
Lecture 7. Diffusion on a half-line and on a finite domain

The method of images. In the left hand panel the absorption boundary condition is enforced with a negative heat pulse at \( x = -\xi \).

The interpretation of the relation above is that the population decreases because of flux into the manhole.

One way of describing the demographics of these unfortunate random walkers is to use the histogram

\[
P(t)\,dt = \text{the number of lifetimes in the interval } (t, t + dt). \quad (7.4)
\]

There is a simple relation between \( P(t) \) and \( S(t) \) because

\[
S(t) = \text{the number of lives longer than } t,
= \int_t^\infty P(t')\,dt'. \quad (7.5)
\]

The derivative of this relation is

\[
P(t) = -\frac{dS}{dt}, \quad (7.6)
\]

and so from (7.3)

\[
P(t) = \kappa u_x(0, t). \quad (7.7)
\]

Does this connection between the flux at \( t \) and the density of lifetimes make sense to you?

The Green’s function of the initial value problem in (7.1), that is absorption Green’s function \( g^a(x, \xi, t) \), is the solution of

\[
g^a_t = \kappa g^a_{xx}, \quad g^a(x, 0) = \delta(x - \xi), \quad g^a(0, t) = 0. \quad (7.8)
\]
Once we have \( g^a \), the solution of (7.1) is

\[
    u(x, t) = \int_0^\infty g^a(x, \xi, t) f(\xi) \, d\xi.
\]  

(7.9)

This is a little more complicated than our earlier formulas because \( g^a \) is a function of \( x \) and \( \xi \) separately (not just \( x - \xi \)).

The \( \delta \)-function at \( x = \xi \) in (7.8) might correspond to a bar which closes at \( t = 0 \). All the patrons are suddenly ejected onto the street and start wandering back and forth. The unlucky ones disappear into the manhole at \( x = 0 \).

To solve this problem, take the \( \delta \)-function at \( x = \xi \) and reflect it in the boundary (\( x = 0 \)) with a negative sign. Thus on the infinite line the initial condition is odd and the diffusion equation preserves this symmetry so that \( g^a(0, \xi, t) \) remains zero at all times. The solution of (7.8) can then be written down using linear superposition:

\[
    g^a(x, \xi, t) = g(x - \xi, t) - g(x + \xi, t),
\]

(7.10)

Now we can calculate the number of survivors from the initial \( \delta \)-pulse:

\[
    S(t) = \frac{1}{\sqrt{4\pi\kappa t}} \left[ \int_0^\infty e^{-(x-\xi)^2/4\kappa t} \, dx - \int_0^\infty e^{-(x+\xi)^2/4\kappa t} \, dx \right],
\]

\[
    = \frac{1}{\sqrt{\pi}} \left[ \int_{-\xi/\sqrt{4\kappa t}}^{\xi/\sqrt{4\kappa t}} e^{-u^2} \, du - \int_{\xi/\sqrt{4\kappa t}}^{\infty} e^{-u^2} \, du \right],
\]

(7.11)

When \( \sqrt{4\kappa t} \gg \xi \), the \( \text{erf} \) simplifies to

\[
    S(t) \approx \frac{\xi}{\sqrt{\pi\kappa t}}.
\]  

(7.12)

Thus death is certain, \( S(t) \to 0 \), but the average lifetime,

\[
    \bar{t} = N^{-1} \int_0^\infty t P(t) \, dt = N^{-1} \int_0^\infty S(t) \, dt,  
\]

(7.13)

is infinite. This indicates that a small fraction of the population make large positive excursions before encountering the absorbing boundary.
7.2 Heating the Earth

As physical examples of half-line diffusion, we consider two classic problems solved by Stokes in the nineteenth century.

Stokes’s first problem

The first problem is

\[ u_t = \kappa u_{xx}, \quad u(0, t) = A \cos \omega t. \]  \hspace{1cm} (7.14)

You can think of this as periodic heating the solid Earth (with the coordinate \( x \) running downwards from the surface at \( x = 0 \)).

The diffusion length (skin-depth) is

\[ \ell = \sqrt{\frac{\kappa}{\omega}}. \]  \hspace{1cm} (7.15)

If we’re thinking of the annual cycle, and we use \( \kappa \) typical of dry soil:

\[ \kappa = 2 \times 10^{-3} \text{cm}^2 \text{s}^{-1}, \quad \omega = \frac{2\pi}{1 \text{ year}}. \]  \hspace{1cm} (7.16)

leading to

\[ \ell \approx 1 \text{ meter}. \]  \hspace{1cm} (7.17)

Because of linearity, the diurnal and annual cycles can be superimposed. The diurnal heating has a penetration depth which is smaller by a factor of \( \sqrt{365} \) than the annual cycle.

Since

\[ \cos \omega t = \frac{1}{2} e^{i\omega t} + \frac{1}{2} e^{-i\omega t}, \]  \hspace{1cm} (7.18)

we can obtain the solution with the substitution

\[ u(x, t) = e^{-i\omega t} U(x) + e^{i\omega t} U^*(x). \]  \hspace{1cm} (7.19)

This gives

\[ U'' + i \ell^{-2} U = 0, \quad U(0) = \frac{1}{2} A, \quad \lim_{x \to \infty} U(x) = 0. \]  \hspace{1cm} (7.20)

The standard guess \( U = e^{px} \) results in \( p^2 + i \ell^{-2} = 0 \), or

\[ p = \pm \sqrt{-i} \ell^{-1} = \pm (1 - i)q, \quad \text{where} \quad q = \frac{1}{\sqrt{2}} \ell. \]  \hspace{1cm} (7.21)
Lecture 7. Diffusion on a half-line and on a finite domain

To ensure decay as \( x \to +\infty \) we must take \( p = -(1-i)q \), so that

\[
  u(x, t) = A e^{-qx} \left( e^{i qx - i \omega t} + e^{-i qx + i \omega t} \right),
  = A e^{-qx} \cos(\omega t - qx).
\] (7.22)

Notice there is a damped signal propagating downwards from the surface — the phase speed is \( \omega/q \).

The phase shift relative to the surface is \( \pi \) at a depth \( x_\ast = \sqrt{2 \pi \ell} \approx 4.4 \) meters. (7.23)

When it is Summer at the surface, it is Winter at a depth of 4.4 meters. The attenuation factor at this depth is

\[
  e^{-qx_\ast} = e^{-\pi} \approx \frac{1}{23}.
\] (7.24)

**Stokes’s second problem**

In the second problem the heating is turned on suddenly at \( t = 0 \)

\[
  u_t = \kappa u_{xx}, \quad u(x, 0) = 0, \quad u(0, t) = 1.
\] (7.25)

Dimensional analysis indicates that the solution has the form

\[
  u(x, t) = U(\eta), \quad \text{with} \quad \eta = \frac{x}{2 \sqrt{\kappa t}}.
\] (7.26)

Substitution into the diffusion equation gives

\[
  U'' + 2\eta U' = 0, \quad \text{with} \quad U(0) = 1, \quad \lim_{\eta \to -\infty} U(\eta) = 0.
\] (7.27)

Integrating this ODE and applying the boundary conditions we have

\[
  U = \text{erfc}(\eta). \quad \text{erfc}(z) \equiv 1 - \text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_{z}^{\infty} e^{-v^2} \, dv
\] (7.28)

The thermal signal penetrates into solid a distance \( \sqrt{\kappa t} \) in time \( t \).
Kelvin’s estimate of the age of the Earth

According to CJ, in deep mines the thermal gradient is about $1{}^\circ\text{C}$ every 25 meters. There is some geographic variation and on land the geothermal gradient $\Gamma$ is in the range

$$10^{-2}\text{K m}^{-1} < \Gamma < 5 \times 10^{-2}\text{K m}^{-1}.$$  

These numbers refer to non-Volcanic regions, and the differences result from variations in the thermal diffusivity of rocks.

Let’s follow Kelvin and use the solution of Stokes’s second problem to estimate the age of the Earth. Assume that we start at $t = 0$ with molten rock at a temperature $u_0 = 1200\text{K}$, and that this rock cools because the surface at $x = 0$ is held at $u = 0$. This is a slight modification of Stokes’s second problem, and the solution is

$$u(x,t) = u_0 \text{erf} \left( \frac{x}{2\sqrt{\kappa t}} \right), \quad \Rightarrow \quad u_x(0,t) = \frac{u_0}{\sqrt{\pi \kappa t}}.$$  \hspace{1cm} (7.29)

The elapsed time is expressed in terms of the surface gradient as

$$t = \frac{u_0^2}{u_x \pi \kappa}.$$  \hspace{1cm} (7.30)

Kelvin took a diffusivity for “average rock” of about $1.18 \times 10^{-6}\text{m}^2\text{s}^{-1}$. Using the smallest geothermal gradient from (7.2), this gives for the age of the Earth$^1$

$$t \sim 123 \times 10^6\text{years}.$$  \hspace{1cm} (7.31)

Note that $\sqrt{\kappa t} \sim 68\text{ kilometers}$, so that we can ignore the sphericity of the Earth — after 123 million years the heat is still being extracted from a thin surface layer.

The age of the Earth is closer to 4.5 billion years and the usual explanation for Kelvin’s gross underestimate is that Kelvin was unaware of radioactive heating. However Richter (1985) shows that inclusion of radioactive heating does not substantially increase estimates based on the diffusion equation. Richter argues that it is mantle convection, which

---

$^1$According to CJ, Kelvin’s opinion was that the Earth could be no older than 24 million years. Kelvin used $u_x = 35\text{K per kilometer}$, typical of measurements near Edinburgh. In any event, whether it is 123 million years or 24 million years, this estimate is far too short to be consistent with the geologic record. For an amusing account of the age of the Earth, see *A Short History of Nearly Everything* by Bill Bryson.
efficiently mixes the entire heat of the Earth up to the surface, that explains how the geothermal gradient can still be larger as 10 degrees per kilometer after 4.5 billion years.

### 7.3 Inhomogeneous boundary conditions

How to solve

\[ u_t = \kappa u_{xx}, \quad u(0, t) = u_0(t), \quad u(x, 0) = 0? \]  \hspace{1cm} (7.32)

This mathematical problem corresponds to an isothermal solid medium occupying the half-space \( x > 0 \). At \( t = 0 \) the boundary temperature at \( x = 0 \) is prescribed to be \( u_0(t) \). This thermal forcing diffuses into the interior, and we seek the interior temperature distribution \( u(x, t) \).

We previously considered the special case \( u_0(t) = H(t) \) and obtained the \( \text{erfc} \) similarity solution, which we write here as

\[ u(x, t) = \text{erfc}(\eta)H(t). \]  \hspace{1cm} (7.33)

The Heaviside function ensures that \( u(x, t < 0) = 0 \) i.e., there is no disturbance before the steady surface heating switches on at \( t = 0 \).

Now consider “top-hat” forcing that switches on at \( t = 0 \) with an amplitude \( \tau^{-1} \), and then switches off at \( t = 0 \). We can represent this forcing function as

\[ u_0(t) = H(t) - H(t - \tau) \frac{\tau}{\tau}. \]  \hspace{1cm} (7.34)

We’ve taken the amplitude of the boundary forcing to be \( \tau^{-1} \), so that as we take the limit \( \tau \to 0 \) we obtain a \( \delta(t) \). We can use linear superposition, and the solution in (7.33), to solve the diffusion equation with the boundary temperature in (7.34):

\[ u(x, t) = \frac{H(t)\text{erfc}(\eta) - H(t - \tau)\text{erfc}(\eta')}{\tau}, \]  \hspace{1cm} (7.35)

where \( \eta' = x/\sqrt{4\kappa(t - \tau)} \). The solution is shown in Figure 7.2.

Taking \( \tau \to 0 \) the top-hat in (7.34) becomes a \( \delta(t^+) \). The solution (7.35) in this same limit \( \tau \to 0 \) is:

\[ \eta(x, t - \tau) \approx \eta(x, t) - \tau \eta_t(x), \]
Figure 7.2: The solution in (7.35) with $\tau = \kappa = 1$. The left panel shows the subtraction of the two erfc’s. The right panel shows the solution surface above the $(x, t)$-plane (again with $\tau = 1$). halfDiff.eps
Lecture 7. Diffusion on a half-line and on a finite domain

\[ G(x, t) = -\eta t \frac{2}{\sqrt{\pi}} e^{-\eta^2} = \frac{x}{2\sqrt{\pi \kappa t^3}} e^{-x^2/4\kappa t}. \]  

(7.36)

This is the Green’s function for the inhomogeneous half-line problem. The solution of (7.32) is therefore:

\[ u(x, t) = \int_0^t u_0(t - t') G(x, t') dt'. \]  

(7.37)

The philosophy behind this method is that we approximate the forcing \( q(t) \) arbitrarily closely by a sequence of narrow top-hats. We know the response to a single top-hat so we use linear superposition to write the response to the sequence — that’s the guts of (7.37) and every other Green’s function formula in these last few lectures.

### 7.4 Diffusion of a heat around a wire loop

Consider a loop of very thin wire. We suppose that the length \( L \) of the loop is much greater than the thickness of the wire, so that heat conduction in loop is governed by the one-dimensional diffusion equation

\[ u_T = \kappa u_{XX}, \quad u(X, 0) = f(X). \]  

(7.38)

The coordinate \( 0 < X < L \) runs around the loop. The shape of the loop is unimportant: it might be a circle or a square. What is important is that the point \( x = 0 \) is the same as the point \( x = L \), and the solution is periodic with period \( L \)

\[ u(x, t) = u(x + L, t). \]

We can crack problems like this using Fourier series. To use the tools we’ve developed so far, we first change the coordinate so that the problem is posed on the fundamental interval \((-\pi, \pi)\) This is accomplished with the change of variable

\[ x = \pi \frac{2X - L}{L}, \quad t = \kappa \left( \frac{2\pi}{L} \right)^2 T. \]  

(7.39)

[^2]: A more detailed description would require solving the three-dimensional diffusion equation. But since the thickness of the wire is small, the temperature rapidly becomes uniform in the transverse \((y \text{ and } z)\) directions. We’re interested in the slower process of heat diffusion around the loop.
Figure 7.3: Solution of $u_t = u_{xx}$ with the initial condition $u(x,0) = \sqrt{x}$. The initial square wave is represented approximately taking 20 terms in (??). The solution is shown at $t = 0.001, 0.01, 0.1$ and 0.5. Diffusion instantly smooths the discontinuity at $x = 0$ and $x = \pi$ so that the Gibbs oscillations disappear.

As $X$ goes from 0 to $L$, $x$ runs from $-\pi$ to $\pi$. On this fundamental interval, the diffusion problem is

$$u_t = u_{xx}, \quad u(x,0) = f(x),$$

(7.40)

with the periodicity condition

$$u(x,t) = u(x + 2\pi, t).$$

(7.41)

It is easy to see that the Fourier series

$$u(x,t) = a_0 + \sum_{k=1}^{n} e^{-k^2 t} \left( a_k \cos kx + b_k \sin kx \right)$$

(7.42)
solves the diffusion equation, and satisfies the periodicity condition. To satisfy the initial condition we use the previous results to determine the constants $a_k$ and $b_k$ above. For example, if the initial condition is $f(x) = \sqrt{x}$ then from (??) we can immediately write down the solution

$$u(x,t) = \frac{4}{\pi} \sum_{k=1}^{n} \frac{\sin[(2k+1)x]}{2k+1} e^{-(2k+1)^2 t}.$$  

(7.43)

This solution, using 20 terms in the series (7.43), is shown in figure 7.3. We remarked previously that the Fourier series for $\sqrt{x}$ converges slowly because smooth sinusoids struggle to approximate the discontinuities. However diffusion immediately blurs the initial discontinuities into nice smooth erf’s. Once this happens the Fourier series converges very quickly indeed: with $t > 0$ the coefficients in (7.43) decrease exponentially with $k$. Consequently, as you can see in Figure 7.3, the Gibbs oscillation disappears very quickly.

\section*{Solution of other PDEs on a loop}

Let’s consider other loopy PDEs defined on $-\pi < x < \pi$: we identify $x = -\pi$ with $x = +\pi$ and look for periodic solutions

$$u(x,t) = u(x + \pi, t).$$  

(7.44)

Some examples are

$$u_t + cu_x, \quad u_t = \beta u_{xxx}, \quad u_t = -\nu u_{xxxx}. \quad (7.45)$$

We suppose that these PDEs come equipped with an initial conditions,

$$u(x,0) = f(x) = \sum_{n=-\infty}^{\infty} f_n e^{inx}.$$

(7.46)

Because the coefficients in (7.45) are constant, it is easy to see that the solution for each complex sinusoid in the sum has the form

$$u(x,t) = e^{inx - i\omega t}.$$  

(7.47)

In the three examples in (7.45), the “dispersion relation” is

$$\omega = cn, \quad \omega = \beta n^3, \quad \omega = -i\nu n^4.$$  

(7.48)
The expressions above show that to solve the PDE there has to be a connection between the frequency $\omega$ and the wavenumber $n$ i.e., the frequency is a function of the wavenumber. Once we have the dispersion relation, the solution of the PDE with the initial condition in (7.46) is simply

$$u(x,t) = \sum_{n=-\infty}^{\infty} f_n e^{inx - i\omega t}. \quad (7.49)$$

**Example** Solve $u_t = \beta u_{xxx}$ on the loop $-\pi \leq x \leq \pi$ with the initial condition

$$u(x,0) = e^{\gamma \cos x}. \quad (7.50)$$

Let’s use a complex Fourier series to represent the initial condition. The coefficients in this series are

$$f_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx + \gamma \cos x} \, dx = I_n(\gamma), \quad (7.51)$$

where $I_n$ is the modified Bessel function. Thus the solution of the PDE is

$$u(x,t) = \sum_{n=-\infty}^{\infty} I_n(\gamma) e^{\lambda (nx - \beta n^3 t)},$$

$$= I_0(\gamma) + 2 \sum_{n=1}^{\infty} I_n(\gamma) \cos \left( n x - \beta n^3 t \right). \quad (7.52)$$

We've used $I_n(\gamma) = I_{-n}(\gamma)$ to write the series in terms of cosines. The solution is shown in Figure 7.4.

% Solution of the dispersive wave equation on a loop
beta=1; gamma = 8; X = linspace(-pi,pi,100); T = [0:0.01:1];
N=40; % truncation of the Fourier series
udata = zeros(length(T),length(X));
nloop = 0;
for tt = T
    nloop = nloop + 1;
    u = besseli(0,gamma)*ones(1,length(X));
    for n = 1:N
        u = u + 2*besseli(n,gamma)*cos(n*X - beta*n^3*tt);
    end
    udata(nloop,:)=u;
end
waterfall(X,T,udata), xlabel('x'), ylabel('t')
colormap([0 0 0]), axis([-pi pi 0 max(T)]), grid off
set( gca,'ztick',[-0.1*exp(-gamma) exp(gamma)] )
view(-60,15), pbaspect([ 1 1 0.75])
Figure 7.4: A solution of the dispersive wave equation with $\beta = 1$ and $\gamma = 8$. dispersivewaterfall.eps

To compute the solution in Figure 7.4, the Fourier series in (7.52) has been truncated at $n = 40$ terms. Figure 7.5 indicates that $n = 40$ is overkill: $I_n(\gamma)$ decreases very rapidly with $n$. This fast convergence is because the initial condition in (7.50) is smooth i.e., infinitely differentiable.

7.5 Problems

Problem 7.1. Solve the half-line $(0 < x < \infty)$ diffusion problem

$$u_t = \kappa u_{xx} - e^{-\alpha x} \cos(\omega t).$$

The boundary condition at $x = 0$ is no flux of heat, $\kappa u_x(0, t) = 0$. Discuss the structure of the solution in the two limits $\alpha^{-1} \gg \sqrt{\kappa/\omega}$ and $\alpha^{-1} \ll \sqrt{\kappa/\omega}$.

Problem 7.2. Consider an ensemble of $N \gg 1$ random walkers wandering in the half-plane $y > 0$; the walkers are released from the point $(x, y) = (0, \xi)$ at $t = 0$. A cliff runs along the $x$-axis, so that walkers fall to their death on a first encounter with $y = 0$. (i) Formulate and solve the
Figure 7.5: With $I_n(\gamma)$ as a function of $n$ at fixed values of $\gamma$ between 1 and 10. This shows that coefficients of the Fourier series in (7.52) decrease very quickly with $n$ i.e., the series converges very quickly. To better estimate the rate of convergence you can apply the saddle point method to analyze the integral in (7.51) for $n \gg 1$ (see BO section 6.6).
two-dimensional random walk model. (ii) Calculate the survival function analogous to (7.2). (iii) How does the number of corpses at the foot of the cliff vary with \( x \) as \( t \to \infty \)?

**Problem 7.3.** Solve the PDE

\[
 u_t + c u_x = \kappa u_{xx}, \quad u(0, t) = \cos \omega t, 
\]
on the half-line \( x > 0 \). There is a nondimensional parameter, \( p \equiv c/\sqrt{\omega \kappa} \). Discuss the structure of the solution in the limits \( c \to \infty \) and \( c \to -\infty \). Can you obtain these limiting solutions more simply from the original PDE by neglecting the diffusive term and solving the resulting first-order (in \( x \)) problem?

To assist grading and discussion of this problem let us agree to use the following notation

\[
 \nu \equiv \frac{|c|}{2\kappa}, \quad \lambda \equiv \frac{4}{p^2} = \frac{4\omega \kappa}{c^2}. 
\]

The identity

\[
 \sqrt{1 + i\lambda} = \pm \left[ \sqrt{\frac{r + 1}{2}} + i\sqrt{\frac{r - 1}{2}} \right], \quad r \equiv \sqrt{1 + \lambda^2} 
\]

(which applies only if \( \lambda > 0 \)) might also be useful. If you use the identity, then prove or verify it and explain what happens if \( \lambda < 0 \).

**Problem 7.4.** (i) Find the reflection Green’s function

\[
 g^r_t = \kappa g^r_{xx}, \quad g^r(x, 0) = \delta(x - \xi), \quad g^r_x(0, t) = 0, 
\]

using the method of images. (ii) The center of mass of the pulse is

\[
 X(\xi, t) \equiv \frac{\int_0^\infty x g^r(x, \xi, t) \, dx}{\int_0^\infty g^r(x, \xi, t) \, dx}. 
\]

Show in the large-time limit (i.e., when \( \sqrt{\kappa t} \gg \xi \)) that \( X(\xi, t) \sim a \sqrt{t} \) and calculate the constant \( a \).

**Problem 7.5.** (i) Show that the Green’s function of the advection-diffusion equation

\[
 g_t + c g_x = \kappa g_{xx}, \quad g(x, 0) = \delta(x), 
\]
Lecture 7. Diffusion on a half-line and on a finite domain

Figure 7.6: Solution of problem 7.5 with $c = -1$, $\xi = 1$ and $\kappa = 1/200$.advecDiffAbs.eps

is

$$g(x, t) = \frac{e^{-(x-ct)^2/4\kappa t}}{\sqrt{4\pi\kappa t}}.$$  

(iii) Now consider the advection-diffusion equation on a half-line $x > 0$ with an absorbing boundary condition

$$g_{a}^{a} + cg_{xx}^{a} = \kappa g_{xx}^{a}, \quad g^{a}(x, 0) = \delta(x - \xi), \quad g^{a}(0, t) = 0.$$  

If $c > 0$ then the drunks have a bias to walk away from the manhole. Guess that a solution might be

$$g^{a}(x, t) = \frac{1}{\sqrt{4\pi\kappa t}} \left[ e^{-(x-\xi-ct)^2/4\kappa t} - \alpha e^{-(x+\xi-ct)^2/4\kappa t} \right],$$  

and determine $\alpha$ by substitution. (iii) Show that the probability that a drunk escapes the manhole is $1 - \exp(-c\xi/\kappa)$ if $c > 0$ and zero if $c < 0$ (see figure 7.6).

Problem 7.6. Use a Fourier series to solve the half-line ($x > 0$) heating problem:

$$u_{t} = \kappa u_{xx}, \quad u(0, t) = \text{sqr}(\omega t).$$  

This is heating the Earth with a uniformly hot Summer followed by a corresponding Winter. Suppose that $\kappa = 2 \times 10^{-3}\text{cm}^{2}\text{s}^{-1}$ and $T = 2\pi/\omega = 1$ year = $3.15 \times 10^{7}$ seconds. Plot the temperature over a yearly cycle at $x = 0$, $x = 1$ meter and $x = 4$ meters.

Problem 7.7. Consider the half-line ($x > 0$) heating problem:

$$u_{t} = u_{xx}, \quad u(0, t) = \sin t H[\sin t].$$
Problem 7.7. Solve the hyper-diffusion problem
\[ u_t = -u_{xxxx}, \quad u(x, 0) = \text{sqr}(x), \]
on the loop \(-\pi \leq x \leq \pi\). Use MATLAB to make a figure analogous to Figure 7.8. Does the hyperdiffusion equation have a maximum principle?

Problem 7.9. (i) Solve Laplace’s equation \( \nabla^2 u = 0 \) for \( r < a \) with the boundary condition \( u(a, \theta) = x^4 \). (ii) With the boundary condition \( u(a, \theta) = \sin^3 \theta \).

Problem 7.10. (i) Considering the half-line problem
\[ u_t = \kappa u_{xx}, \quad u(x, 0) = 0, \quad \kappa u_x(0, t) = -1, \]
show that
\[ \int_0^\infty u(x, t)dx = t. \]
Lecture 7. Diffusion on a half-line and on a finite domain

(ii) Solve the problem with a similarity ansatz. (iii) Adapt the argument in section [7.3] to obtain a Green’s function solution of the applied flux problem

\[ u_t = \kappa u_{xx}, \quad u(x, 0) = 0, \quad \kappa u_x(0, t) = -f(t). \]

(iv) Show that the surface temperature is given in terms of the applied flux by

\[ u_0(t) = \int_0^t \frac{f(t')}{\sqrt{\pi \kappa (t - t')}} dt'. \]

(v) Calculate the surface temperature if the flux is \( f(t) = t^p \). To check your answer show that \( u_0 \) is constant if \( p = -1/2 \).

**Hint:** In part (iii) the similarity ansatz results in the ODE \( U'' + 2\eta U' - 2U = 0 \). Solve this by guessing a special solution, followed by reduction of order.

**Problem 7.11.** Consider the \( x > 0 \) forced problem

\[ u_t = s + \kappa u_{xx}, \quad u(0, t) = 0, \quad u(x, 0) = u_0, \]

where \( s \) is a uniform source and \( u_0 \) is the uniform initial temperature. Show that the surface gradient is

\[ u_x(0, t) = 2s \sqrt{\frac{t}{\pi \kappa}} + \frac{u_0}{\sqrt{\pi \kappa t}}. \]

**Problem 7.12.** Give an alternative solution of (7.32) by writing

\[ w(x, t) \equiv u(x, t) - q(t). \]

Thus \( w(x, t) \) satisfies a forced diffusion equation with the homogeneous boundary condition \( w(0, t) = 0 \). Construct a Green’s function solution of this equation by applying Duhamel’s principle to the absorption Green’s function in (7.10). Make sure your answer agrees with (7.37).

**Problem 7.13.** Considering the problem (7.32), find a compact expression for the “inventory” and the surface flux,

\[ h(t) \equiv \int_0^\infty u(x, t') dt', \quad \text{and} \quad f(t) \equiv -\kappa u_x(0, t), \]

in terms of the applied surface temperature \( u_0(t) \). Suppose that \( u_0(t) \) decays faster than any power of \( t \) at large time e.g., \( u_0(t) = \exp(-t) \). Show that at large times

\[ h(t) \sim \frac{\alpha_0}{t^{1/2}} + \frac{\alpha_1}{t^{3/2}} + \frac{\alpha_2}{t^{5/2}} + \cdots \]

and determine the constants \( \alpha_0 \) and \( \alpha_1 \) in terms of \( u_0(t) \).

SIO203C, W.R. Young, March 21, 2011
Problem 7.14. (i) Show that if $\theta_1(x, t)$ and $\theta_2(y, t)$ are solutions of the one-dimensional diffusion equations:

$$
\theta_1 t = \kappa \theta_1_{xx}, \quad \theta_2 t = \kappa \theta_2_{yy}
$$

then $\theta(x, y, t) \equiv \theta_1(x, t)\theta_2(y, t)$ is a solution of the two-dimensional diffusion equation. (ii) Use this trick to find the two-dimensional Green’s function

$$
g_t = \kappa(g_{xx} + g_{yy}), \quad g(x, y, t) = \delta(x)\delta(y)\delta(t) .
$$

(iii) Consider diffusion in a quarter-plane $(x, y) > 0$ and suppose that initially $\theta(x, y, 0) = 0$. At $t > 0$ the edges $x = 0$ and $y = 0$ are held at fixed temperature:

$$
\theta(x, 0, t) = \theta(0, y, t) = 1 .
$$

Find the temperature distribution in the quarter plane.
Lecture 8

**Laplace’s equation** $\nabla^2 u = 0$

In the case of steady heat diffusion with uniform conductivity we obtain Laplace's equation

$$\nabla^2 u = 0. \tag{8.1}$$

If there is a steady heat source on the right hand side of the diffusion equation, as in (6.71), then in steady state we have the inhomogeneous form of Laplace’s equation, known as Poisson’s equation

$$u_t = s + \kappa \nabla^2 u.$$

Poisson’s equation is one of the most important equations in applied mathematics and physics.

For example, the gravitational potential $\varphi(\mathbf{x})$ associated with a mass density $\rho(\mathbf{x})$ is obtained by solving

$$\nabla^2 \varphi = 4\pi G \rho, \tag{8.3}$$

where $G$ is the gravitational constant. Once we possess $\varphi$, the force of gravity is $\mathbf{f} = -\nabla \varphi$. Poisson’s equation also determines the electrical potential associated with a charge density. And in geophysical fluid dynamics, Poisson’s equation determines the streamfunction $\psi$ from the potential vorticity $q$ e.g.,

$$\nabla^2 \psi + \beta y = q. \tag{8.4}$$
8.1 The \( \delta \)-function in \( d = 3 \)

Suppose we have a small, intense source of heat sitting at the origin of a conducting medium e.g., a hot radioactive pellet at \( \mathbf{x} = 0 \) surrounded by concrete. We model this heat source as \( s(\mathbf{x}) = q\delta(\mathbf{x}) \), so that steady-state Poisson’s equation is

\[
\kappa \nabla^2 u = -q\delta(\mathbf{x}).
\]  

(8.5)

The constant \( q \) is total strength of the source and \( \delta(\mathbf{x}) \) is a three-dimensional \( \delta \)-function i.e, a quantity localized at the point \( \mathbf{x} = (x, y, z) = 0 \), but with non-zero integral. Before we solve \( \text{(8.5)} \) let’s discuss how this idealized function is constructed.

To make a three-dimensional \( \delta \)-function we consider a sequence of functions with a parameter \( \epsilon \):

\[
\delta_{\epsilon}(\mathbf{x}) = \frac{1}{\epsilon^3} \Delta \left( \frac{r}{\epsilon} \right).
\]  

(8.6)

Notice that \( \epsilon \) has the dimensions of length. The function \( \Delta(s) \) is normalized by requiring that

\[
4\pi \int_{0}^{\infty} \Delta(s) s^2 \, ds = 1.
\]  

(8.7)

The conditions above ensure that if we integrate over all space then

\[
\int \delta_{\epsilon}(\mathbf{x}) \, dv = 1, \quad \text{independent of } \epsilon.
\]  

(8.8)

Some popular choices for \( \Delta(s) \) are the Gaussian

\[
\Delta(s) = \pi^{-3/2} e^{-s^2},
\]  

(8.9)

or the top-hat

\[
\Delta(s) = \frac{3}{4\pi} \begin{cases} 
1, & \text{if } s < 1; \\
0, & \text{if } s > 1. 
\end{cases}
\]  

(8.10)

Any equation involving \( \delta(\mathbf{x}) \) can be interpreted — or must be interpreted — by backing up to the non-singular function \( \delta_{\epsilon}(\mathbf{x}) \) and taking the limit as \( \epsilon \to 0 \). This is how BO explain the one-dimensional \( \delta \)-function in their section 1.5 — see also lecture 4. It is the same here,
Lecture 8. Laplace’s equation $\nabla^2 u = 0$

except that in $d = 3$ we have a factor $\epsilon^{-3}$ in (8.6). In $d = 3$ the exponent $-3$ is required so that the normalization in (8.8) is maintained as we dial $\epsilon$ all the way down to zero. It doesn’t matter whether we use the Gaussian or the top-hat for $\Delta$: the internal details of $\Delta(s)$ should be irrelevant in the limit $\epsilon \to 0$.

Following the prescription above, the equation

$$\delta(x) = 0 \quad \text{if } x \neq 0,$$

really means

$$\lim_{\epsilon \to 0} \delta_{\epsilon}(x) = 0, \quad \text{provided that } x \neq 0. \quad \text{(8.12)}$$

It is easy to see this is true if we use the Gaussian or top-hat for $\Delta$.

A formula like

$$\delta(x) = \delta(x)\delta(y)\delta(z) \quad \text{(8.13)}$$

is interpreted by backing up to

$$\delta_{\epsilon}(x) = \delta_{\epsilon}(x)\delta_{\epsilon}(y)\delta_{\epsilon}(z), \quad \text{(8.14)}$$

where the three one-dimensional $\delta$-functions on the right are constructed following the recipe in an earlier lecture. It is easy to verify that this product of three one-dimensional $\delta$-functions satisfies our three-dimensional requirements in (8.6) through (8.8).

A most important property of the three-dimensional $\delta$-function is “sifting”. If $f(x)$ is a continuous function then:

$$f(x) = \int f(x') \delta(x - x') \, dv = \int f(x - x') \delta(x') \, dv. \quad \text{(8.15)}$$

The value of $f$ at the location of the $\delta$-singularity is sifted out of the integral. This crucial result should be obvious if you think about replacing $\delta$ by $\delta_{\epsilon}$ in the integrand, and then taking the limit $\epsilon \to 0$.

The hot pellet

Now let’s return to our radioactive pellet and the Poisson equation (8.5). A main property of the three-dimensional $\delta$-function on the right of (8.5) is

$$\int_V \delta(x) \, dv = \begin{cases} 1, & \text{if } x = 0 \text{ is in the volume } V; \\ 0, & \text{if } x = 0 \text{ is not in } V. \end{cases} \quad \text{(8.16)}$$
Lecture 8. Laplace’s equation $\nabla^2 u = 0$

Thus if we integrate (8.5) over any volume $V$ containing the origin then with Gauss’s divergence theorem

$$\int_{\partial V} \kappa \nabla u \cdot \hat{n} \, ds = -q. \quad (8.17)$$

This says that in steady state all the heat released at $x = 0$ has to diffuse through the control surface $\partial V$.

Suppose we pick a special $V$ which is a sphere of radius of $r$. From symmetry, $u$ is a function only of $r$, and thus

$$\nabla u \cdot \hat{n} = \frac{du}{dr}. \quad (8.18)$$

Hence (8.17) is

$$\kappa \frac{du}{dr} = -\frac{q}{4\pi r^2}. \quad (8.19)$$

The result above can be obtained more formally. Using spherical coordinates, (8.5) is

$$\kappa \frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} u = -q \frac{\delta(r^+)}{4\pi r^2} \quad \Rightarrow \quad \phi = -\frac{Gm}{r} \quad (8.21)$$

The $r^{-2}$ on left and right cancel, and $\delta(r^+)$ is concentrated just on the positive side of $r = 0$ so that $\int_0^r \delta(r_1^+) \, dr_1 = 1$. Then integrating with respect to $r$ we quickly get back to (8.19). If this is the first time you’ve seen $\delta(x) = \frac{\delta(r^+)}{4\pi r^2}$, then this might seem mysterious. (See the exercises.)

In any event, using either of the methods above, we arrive at (8.19). Integrating this first-order ODE we find that the steady-state temperature is

$$u = u_\infty + \frac{q}{4\pi \kappa r} \quad (8.22)$$

where the constant of integration $u_\infty$ is the temperature at great distances from the hot pellet.

1I hope you also recognize that this solution of Poisson’s equation is equivalent to the Newtonian potential of a point mass

$$\nabla^2 \varphi = 4\pi Gm \delta(x), \quad \Rightarrow \quad \varphi = -\frac{Gm}{r}.$$

The potential at $r = \infty$ is conventionally taken to be zero.
8.2 The Laplacian Green’s function in $d = 1, 2$ and $3$

Now consider the problem of finding the spherically symmetric Green’s function in $d = 1, 2$ and $3$. This means solving the Poisson equation

$$\nabla^2 g = \delta(\mathbf{x}).$$  

(8.23)

The answer is

$$g(\mathbf{x}) = \begin{cases} \frac{1}{2} |x|, & \text{in } d = 1; \\ \frac{1}{2\pi} \ln \frac{1}{r}, & \text{in } d = 2; \\ \frac{1}{4\pi r}, & \text{in } d = 3. \end{cases}$$  

(8.24)

These solutions are not unique: one can add any constant to the solutions above.

Aside from notational differences, the case $d = 3$ is the hot-pellet solution from the previous section. The $d = 1$ and $d = 2$ cases follow in the examples.

Example: Find a “radially symmetric” solution of the one-dimensional Poisson equation

$$\frac{d^2 g}{dx^2} = \delta(x).$$  

(8.25)

Away from the singularity at $x = 0$, we have the one-dimensional Laplace equation $g_{xx} = 0$, with general solution

$$g(x) = \begin{cases} A_- + B_- x, & \text{if } x < 0, \\ A_+ + B_+ x, & \text{if } x > 0. \end{cases}$$  

(8.26)

We also have jump conditions, obtained by integrating (8.25) across the $\delta$-function at $x = 0$. This gives

$$\frac{dg}{dx}(0^+) - \frac{dg}{dx}(0^-) = 1.$$  

(8.27)

There is no jump in $g(x)$ at $x = 0$ i.e., $g(x)$ is continuous, or $g(0^+) = g(0^-)$. These conditions give

$$A_- = A_+,$$  

$$B_+ - B_- = 1.$$  

(8.28)

But we’re looking for a symmetric solution, so $B_- = -B_+$, so $B_+ = 1/2$ and $B_- = -1/2$. Thus the Green’s function is

$$g(x) = A + \frac{1}{2} |x|.$$  

(8.29)

There is no way to determine $A$, except perhaps by considering the initial value problem. ■
Lecture 8. Laplace’s equation $\nabla^2 u = 0$

**Example:** Find a “radially symmetric” solution of the two-dimensional Poisson equation

$$g_{xx} + g_{yy} = \delta(x). \quad (8.30)$$

Integrate over the area enclosed by a circle of radius $r$, and perimeter $2\pi r$, to obtain

$$\oint g_r \, d\ell = 1. \quad (8.31)$$

Since the solution is axisymmetric

$$g_r = \frac{1}{2\pi r}, \quad \text{or} \quad g = \frac{\ln r}{2\pi} \quad \blacksquare$$

We might as well record the general case: in dimension $d$ the Green’s function is

$$g(x) = \frac{1}{(2-d) \Omega_d r^{d-2}}, \quad (d \geq 3), \quad (8.31)$$

where $\Omega_d$ is the area of a $d$-dimensional unit sphere (see the exercises). Thus if we’re confronted with the infinite-space Poisson equation

$$\nabla^2 u = \rho, \quad (8.32)$$

then the solution is just

$$u(x) = \int \rho(x') g(x - x') \, dx', \quad (8.33)$$

where $g$ depends on the dimension $d$ of the Euclidian space.

### 8.3 The $d$-dimensional diffusion Green’s function

In dimension $d$, with $x = [x_1, x_2, \ldots, x_d]$, the diffusion Green’s function $G(x, t)$ is determined by

$$G_t = \kappa \left( G_{x_1 x_1} + G_{x_2 x_2} + \cdots + G_{x_d x_d} \right), \quad (8.34)$$

with the initial condition

$$G(x, 0) = \delta(x). \quad (8.35)$$
Lecture 8. Laplace's equation $\nabla^2 u = 0$

Since the $d$-dimensional $\delta$-function can be written as product of one-dimensional $\delta$-functions,

$$\delta(\mathbf{x}) = \delta(x_1)\delta(x_2) \cdots \delta(x_d),$$

(8.36)

it is easy to verify that the solution of (8.35) and (8.35) is

$$G(\mathbf{x}, t) = g(x_1, t)g(x_2, t) \cdots g(x_d, t),$$

$$= (4\pi\kappa t)^{-d/2} \exp(-r^2/4\kappa t),$$

(8.37)

where $g(x_1, t)$ is the one-dimensional Green's function in (6.36) and $r^2 = x_1^2 + x_2^2 + \cdots + x_d^2$.

Once we have $G(\mathbf{x}, t)$, the solution of the $d$-dimensional initial value problem

$$u_t = \kappa \nabla^2 u, \quad u(\mathbf{x}, 0) = f(\mathbf{x}),$$

(8.38)

is given by the convolution

$$u(\mathbf{x}, t) = \int G(\mathbf{x} - \mathbf{x}', t)f(\mathbf{x}')\, d\mathbf{x}'.$$  

(8.39)

8.4 The efficiency of diffusion depends on $d$

You may have noticed in (8.24) that only in $d = 3$ does the Laplacian Green's function decay to zero as $r \to \infty$. With $d = 1$ or 2 the Laplacian Green's function diverges as $r \to \infty$. This is an indication that diffusion is more efficient in higher-dimensional spaces — heat gets away from the $\delta(\mathbf{x})$-source faster in $d = 3$. A supporting observation is that according to (8.37), $G(0, t) \propto t^{-d/2}$ i.e., the central value of the pulse decays faster with time in large dimensions.

To illustrate the difference between diffusion in $d = 2$ and $d = 3$, consider the problem of a hot radioactive rod embedded in a large mass of concrete, with uniform temperature, $u = u_\infty$, at large distances from the rod.

If the rod is very long and stretches along the $z$-axis, then we might expect that the problem is effectively two dimensional, and we could try to find the steady state temperature distribution by writing

$$u(r, t) = u_\infty + v(r),$$  

(8.40)
Lecture 8. Laplace’s equation $\nabla^2 u = 0$

where now $r = \sqrt{x^2 + y^2}$ is the distance of $(x, y, z)$ from the $z$-axis. The steady temperature anomaly must satisfy

$$\kappa \left( v_{xx} + v_{yy} \right) + q\delta(x)\delta(y) = 0, \quad \text{with } \lim_{r \to \infty} v(r) = 0. \quad (8.41)$$

But the general axisymmetric solution is $v = A + B \ln r$, and this cannot satisfy the condition of decay at $r = \infty$. The best we can do is to ignore the condition at $\infty$ and determine $B$ by balancing the heat input with diffusion, so that

$$v \overset{?}{=} A + \ln\left(\frac{r^{-1}}{2\pi}\right). \quad (8.42)$$

We just can’t satisfy the condition at $r = \infty$. There are at least two ways of resolving this problem:

1. Recognize that the rod has finite length, $2a$, and solve a steady three-dimensional diffusion problem with the source

   $$s(x) = q\delta(x)\delta(y)\chi(z),$$

   where $\chi(z) = 1$ if $|z| < a$ and zero otherwise.

2. Remain in $d = 2$ and solve an unsteady diffusion equation in which the rod is inserted into an isothermal medium at $t = 0$.

Both problems are instructive, and provide good applications of Green’s method. Let’s explore them in turn.

The steady $d = 3$ problem

In this case we use the $d = 3$ Green’s function to write down the solution

$$v(x, y, z) = \frac{q}{4\pi \kappa} \int_{-a}^{a} \frac{dz'}{\sqrt{x^2 + y^2 + (z - z')^2}}. \quad (8.43)$$

Feeding this command

Assuming\[r > 0 && z > 0 && a > 0,\]
\[\text{Integrate}[1/\text{Sqrt}[r^2 + (z - u)^2], \{u, -a, a\}]\]
Lecture 8. Laplace’s equation $\nabla^2 u = 0$

Figure 8.1: Left panel shows isotherms round the hot rod in the $(x,z)$-plane computed from (8.44). The contour interval is logarithmic. Right panel shows the $E_1(\eta^2)$ (the solid curve) and the small-$r$ approximation $-2\ln\eta - 0.57721$ (the dashed curve). hotRod.eps

into mathematica\footnote{See also classic texts on potential theory such as Foundations of Potential theory by O.D. Kellogg, or Newtonian Attraction by A. S. Ramsey, where this integral, and many others, are done analytically.}, produces

$$\log\left[-\frac{(a + \text{Sqrt}[r^2 + (a - z)^2] - z)/(a + z - \text{Sqrt}[r^2 + (a + z)^2])}{(a + \text{Sqrt}[r^2 + (a - z)^2] - z)/(a + z - \text{Sqrt}[r^2 + (a + z)^2])}\right]$$

(8.44)

(8.45)

$R_\pm$ is the distance between $\mathbf{x}$ and the end of the rod at $z = \pm a$. Figure 8.1 shows $v(x,0,z)$.

This solution has the expected behaviour. For example, on the plane $z = 0$ the temperature simplifies to

$$v(x,y,0) = \frac{q}{4\pi\kappa} \ln \left[\frac{\sqrt{x^2 + y^2 + a^2} + a}{\sqrt{x^2 + y^2 + a^2} - a}\right].$$

(8.46)
Lecture 8. Laplace’s equation $\nabla^2 u = 0$

It is easy to see from this expression that close to the rod $G \propto \ln r$, and far from the rod one recovers the $r^{-1}$ decay of the $d = 3$ Green’s function (see exercises).

The unsteady $d = 2$ problem

In this case we solve the two-dimensional initial value problem

$$v_t = \kappa \left( v_{xx} + v_{yy} \right) + q \delta(x) \delta(y), \quad \text{with } v(x, 0) = 0. \quad (8.47)$$

The method of section 6.6 gives the integral representation

$$v(r, t) = q \int_0^t \frac{e^{-r^2/4\kappa \tau}}{4\pi \kappa \tau} \, d\tau. \quad (8.48)$$

We beat this integral into a standard from with the change of variables $w = r^2/4\kappa \tau$. Thus

$$v(r, t) = \frac{q}{4\pi \kappa} \int_0^\infty \frac{e^{-w}}{w} \, dw, \quad (8.49)$$

$$= \frac{q}{4\pi \kappa} E_1 \left( \eta^2 \right), \quad (8.50)$$

where $E_1$ is the exponential integral (e.g., BO section 6.2, and expint in MATLAB). Also in (8.50), $\eta \equiv r/2\sqrt{\kappa t}$ is our favourite similarity variable. Close to the rod, where $\eta \ll 1$, we use the expansion of $E_1$ given by BO in their (6.2.8): Euler’s constant is:

$$\gamma_E = 0.57721 \cdots$$

$$v(r, t) \sim \frac{q}{2\pi \kappa} \ln \frac{1}{r} + a \text{ term with } \ln t. \quad (8.51)$$

The solution is shown in the right panel of Figure 8.1. In the near field of the rod the steady $\ln r^{-1}$ solution is a dominant part of the answer, as the slowly growing $\ln t$. This slow increase in temperature is an indication that two-dimensional diffusion cannot keep up with the steady influx of heat from the rod.

8.5 Some sphere problems
8.6 Problems

Problem 8.1. One face of a slab, say \( x = 0 \), is held at uniform temperature \( u = 0 \) and the other face at \( x = \ell \) is held at \( u = u_1 \). Assume that the conductivity \( k \) is a function only of \( x \). Find the temperature profile and the steady heat flux through the slab.

Solution. In steady state the temperature depends only on \( x \) and the diffusion equation (6.1) reduces to

\[
\frac{d}{dx} k \frac{du}{dx} = 0, \quad \Rightarrow \quad \frac{du}{dx} = \frac{f}{k(x)}.
\]

(8.52)

The constant of integration \( f \) is the heat flux (Watts per square meter). To satisfy the boundary conditions we require that the integral from \( x = 0 \) to \( \ell \) of \( du/dx \) is equal to \( u_1 \), or

\[
u_1 = f \int_0^\ell \frac{dx}{k(x)}.
\]

(8.53)

With \( f \) determined by this formula, the temperature profile is

\[
u(x) = f \int_0^x \frac{dx'}{k(x')}
\]

(8.54)

If \( k \) is constant then the temperature varies linearly with \( x \) — this is the second simplest solution of Laplace’s equation in \( d = 1 \). The simplest solution is just constant \( u \). Note a small region in which \( k(x) \) is very small (an insulating layer) makes a large reduction in the heat flux \( f \) determined from (8.53).

Problem 8.2. The temperature in (8.22) is singularly large as \( r \to 0 \). This is an unphysical artifact of assuming that the radioactive pellet is a point source. So consider a new and improved model with a spherical pellet

\[
\kappa \nabla^2 u = -\frac{q}{\epsilon^3} \Delta \left( \frac{r}{\epsilon} \right),
\]

(8.55)

where \( \epsilon \) is a length (the radius of the pellet) and \( \Delta \) is the top-hat in (8.10). Solve (8.55) and verify that the central temperature \( u(0) \) is finite if \( \epsilon > 0 \).

Problem 8.3. Prove that

\[
\delta(x) = \delta(x)\delta(y)\delta(z) = \frac{\delta(r^+)}{4\pi r^2}.
\]

What’s the analog of this result in \( d = 2 \)?
Problem 8.4. What are the dimensions of $\delta(x)$ in $d = 1, 2$ and 3?

Problem 8.5. Find the spherically symmetric Green’s function defined by

$$\nabla^2 g - k^2 g = \delta(x)$$

in $d = 1, 2$ and 3.

Problem 8.6. Obtain the $d$-dimensional Laplacian Green’s function in (8.31).

Problem 8.7. (i) Show that the $d$-dimensional Green’s function in (8.37) is normalized so that its integral over all space is equal to one. (ii) Integrate $G(r,t)$ in (8.37) over all space using spherical coordinates. In these coordinates the volume element is

$$\mathrm{d}x = \Omega_d r^{d-1} \mathrm{d}r,$$

where $\Omega_d$ is the area of a $d$-dimensional unit sphere. Deduce

$$\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}, \quad \text{where } \Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, \mathrm{d}t.$$ 

Problem 8.8. (i) Assume that the solution of (8.34) and (8.35) is spherically symmetric i.e., that $G = G(r,t)$. (ii) Show that the axisymmetric diffusion equation can be written as

$$r^{d-1} G_t = \kappa \left( r^{d-1} G_r \right)_r.$$ 

Interpret this last equation in terms of flux through a sphere with area $\Omega_d r^{d-1}$ etc. (iii) Solve this PDE with the similarity method, and show that you recover (8.37).

Problem 8.9. A laser zaps a big copper sheet and instantly deposits $Q$ Joules of heat at a point $P$. Treat the sheet as two-dimensional, with density $\rho$ kilograms per square meter, and suppose that it initially has uniform temperature. Then the excess temperature $u$ satisfies (6.1). Solve this PDE with an initial condition that models the zap, and show that $Q$ Joules eventually diffuse through any circle of radius $r$ centered on $P$.

Problem 8.10. Suppose that the ansatz $u(x,y,t) = a(x,t)b(y,t)$ is a solution of the two-dimensional diffusion equation $u_t = \kappa (u_{xx} + u_{yy})$. Find the most general PDEs satisfied the functions $a$ and $b$. Generalize to dimension $d$. 

SIO203C, W.R. Young, March 21, 2011
Problem 8.11. Simplify $v(x, y, 0)$ in (8.1) in the limits $r \ll a$ and $r \gg a$. Calculate the first two or three terms so that you understand the corrections to the expected asymptotic forms $\ln r$ and $r^{-1}$.

Problem 8.12. Solve the initial value problem

$$v_t = \kappa \nabla^2 v + q \delta(x), \quad v(x, 0) = 0,$$

in $d = 1, 2$ and $3$. Write the answer in terms of exponential integrals and discuss the long-time limit. Does the $d = 1$ solution evolve to resemble the Laplacian Green’s function $g = \frac{1}{2}|x|$ as $t \to \infty$ with $x$ fixed?
Lecture 9

More on Laplace’s equation in \(d=2\)

9.1 Discrete harmonic functions

Tom and Jerry toss a fair coin. If the toss is a 'head' then Tom gives Jerry $1 and if it is a 'tail' then vice versa. Tom starts with \(x\) dollars and Jerry with \(n-x\) dollars. Find the probability that Tom wins.

Define
\[
p(x) = \text{Probability Tom wins starting with } x \text{ dollars.} \quad (9.1)
\]
Clearly \(p(0) = 0\) and \(p(n) = 1\) After the first toss Tom has either \(x-1\) dollars or \(x+1\) dollars, both with probability \(1/2\). These events are mutually exclusive and exhaustive so
\[
p(x) = \frac{1}{2}p(x-1) + \frac{1}{2}p(x+1). \quad (9.2)
\]
We have used the following idea from probability theory: suppose \(E\) is an event (Tom wins) and \(F\) and \(G\) be mutually exclusive subevents. Then
\[
\text{prob}(E) = \text{prob}(F) \times \text{prob}(E \text{ given } F) + \text{prob}(G) \times \text{prob}(E \text{ given } G) \quad (9.3)
\]
In (9.2) \(F\) is the event that the first toss is a head and \(G\) is the event that the first toss is a tail.

The function \(p(x)\) is said to be harmonic if it has the averaging property in (9.2): the solution at \(x\) is the average of the solution at the neighbouring two points.
By considering small values of \( n \) it is easy to convince oneself that the solution of the difference equation (9.2) is

\[
p(x) = \frac{x}{n}.
\]  

(9.4)

The averaging property is apparent e.g. \( 17 = \frac{(16 + 18)}{2} \). If you look at the month of May in a calendar you’ll see that each interior entry (e.g., May 23) is an average of the four entries to the North, South, East and West (e.g., May 16, 30, 24 and 22). This leads us to the two-dimensional case...

**Random walks \( d = 2 \)**

Suppose we play a similar game in two dimensions (see figure) in which a criminal random walks on an integer grid. Find

\[
p(x, y) \equiv \text{Probability of escaping starting at } (x, y) \]  

(9.5)

In this case the analog of (9.2) is

\[
p(x, y) = \frac{1}{4} \left[ p(x - 1, y) + p(x + 1, y) + p(x, y - 1) + p(x, y + 1) \right],
\]

(9.6)

where \((x, y)\) is an interior point. Once again, notice that (9.6) says that the value of \( p(x, y) \) at an interior point is the average of \( p \) at the four neighbouring points. At the boundary points \( p \) is either zero (capture) or one (escape). The solution of this problem is a two-dimensional discrete harmonic function.

**Example:** Determine \( p \) at the three interior grid points in the top panel of figure 9.1.

In this case the averaging-equations are

\[
p_1 = \frac{1}{4} p_2, \quad p_2 = \frac{1}{4} (p_1 + p_3), \quad p_3 = \frac{1}{4} (p_2 + 3).
\]

(9.7)

After some scribbling we find the solution of this \( 3 \times 3 \) linear system is

\[
p_1 = \frac{3}{36}, \quad p_2 = \frac{12}{36}, \quad p_3 = \frac{45}{36} \]

(9.8)
Figure 9.1: Two discrete grids. JoFig1.eps
The maximum principle

The averaging property in (9.6) implies that the maximum and minimum values of \( p(x) \) must be at the boundaries of the domain: an interior point cannot have a greater value of \( p \) than all four of its neighbours.

The maximum principle implies uniqueness. Suppose we have two solutions \( p_1(x, y) \) and \( p_2(x, y) \) satisfying (9.13) with the same boundary conditions. Then the difference \( v = p_2 - p_1 \) also satisfies (9.13) with \( v = 0 \) on the boundary. Thus

\[
\max(v) = \min(v) = 0, \quad \Rightarrow \quad v = 0. \quad (9.9)
\]

This is a justification for guessing solutions — as if we need justification! It may be some comfort to know that if you can find a solution than you can be sure you have the only solution.

Minimization of the Dirichlet functional

We now show that the solution of the discrete Laplace equation minimizes the Dirichlet functional

\[
D[p] = \sum_{\text{nnp}} (p_m - p_n)^2, \quad (9.10)
\]

where 'nnp' stands for 'nearest neighbour pairs' and \( p = [p_1, p_2 \cdots p_{n+m}] \) is a vector containing the value of \( p \) at the \( n \) interior points and the \( m \) boundary points.

The obvious way to minimize \( D \) is to make \( p \) constant: then \( D[p] = 0 \). But this trivial solution is not consistent with the boundary conditions. The functional \( D \) is a measure of how much the different elements of \( p \) vary, and minimizing \( D[p] \) makes \( p \) as close to constant as possible, consistent with the variations imposed by the boundary points. To find the true minimum we differentiate \( D \) with respect to the \( n \) interior points. For instance, suppose \( p_7 \) is an interior point. Now in \( d = 2 \) there are four terms in \( D \) containing \( p_7 \):

\[
D[p] = (p_7 - p_N)^2 + (p_7 - p_S)^2 + (p_7 - p_E)^2 + (p_7 - p_W)^2 + \cdots \quad (9.11)
\]

where \( N, S, E \) and \( W \) stands for the four nearest neighbour points to point 7. To find the minimum

\[
\frac{\partial D}{\partial p_7} = 0, \quad \Rightarrow \quad p_7 = \frac{1}{4} (p_N + p_S + p_E + p_W). \quad (9.12)
\]
Lecture 9. More on Laplace’s equation in \( d = 2 \)

Thus for every interior point we recover the discrete Laplace equation in (9.6).

From the discrete to the continuous

If we leap from the discrete to the continuous — which is justified if we have an enormous grid with lots of points — then we arrive at

\[
\begin{align*}
\text{[9.2]} & \quad \Rightarrow \quad \frac{d^2 p}{dx^2} = 0, \\
\text{[9.6]} & \quad \Rightarrow \quad p_{xx} + p_{yy} = 0.
\end{align*}
\]

(9.13)

Thus Laplace’s equation can also be regarded as a continuum approximation to this probabilistic escape problem. In the remainder of this lecture we use the escape problem to illustrate solutions of Laplace’s equation.

9.2 The Dirichlet and Neuman problems in \( d = 2 \)

We can visualize two-dimensional solutions of Laplace’s equation in a simply connected region \( R \) of the \((x, y)\) plane, as steady-state temperature distribution in sheet of metal shaped like \( R \). To prevent diffusion from making the temperature uniform, suppose that the temperature on the boundary \( \partial R \) is held in some non-uniform condition. The problem of solving Laplace’s equation \( \nabla^2 u = 0 \), subject to a boundary condition with \( u \) is specified is the Dirichlet problem.

There is a second problem, the Neuman problem, which is to solve Laplace’s equation within \( R \) with the normal derivative \( u_n \equiv \hat{n} \cdot \nabla u \) specified on \( \partial R \). In the case of heat conduction, the Neuman problem amounts to specifying the heat flux through the boundary \( \partial R \).

It is usually impossible\(^2\) to solve Laplace's equation within \( R \) if both \( u \) and \( u_n \) are specified. Again using heat conduction as an example, it is not possible to simultaneously specify both the temperature and the

---

\(^1\)Simply connected means that every closed curve in \( R \) can be contracted to a point without passing outside of \( R \). A disc is simply connected and an annulus is not simply connected.

\(^2\)If both \( u \) and \( u_n \) are specified there are no non-singular solutions.
heat flux at the boundary ∂R. If the boundary temperature is specified then the boundary heat flux is part of the answer, and vice versa.

**Example:** Suppose \( R \) is the interior of the circle \( x^2 + y^2 = a^2 \) (i.e., a disc of radius \( a \)) and on the boundary \( ∂R \), the temperature is \( u = \cos nθ \) where \( θ = \tan^{-1}(y/x) \) is the usual polar angle. Find \( u(x, y) \) inside \( R \).

The case \( n = 0 \) is trivial: \( u(x, y) = 1 \) is the solution. With \( n ≥ 1 \) we consider Laplace’s equation in polar coordinates:

\[
\begin{align*}
  u_{rr} + r^{-1}u_r + r^{-2}u_{θθ} &= 0. \quad (9.14)
\end{align*}
\]

Looking for a separable solution with

\[
  u(r, θ) = U(r) \cos nθ, \quad (9.15)
\]

we are greeted by an Euler equation (see BO):

\[
  U'' + r^{-1}U' - n^2r^{-2}U = 0. \quad (9.16)
\]

The solution is obtained with \( U = r^λ \), and we find \( λ = ±n \). Inside the disc we avoid the singular solution \( r^{-n} \), so the solution is

\[
  u = \left( \frac{r}{a} \right)^n \cos nθ. \quad (9.17)
\]

The cases \( n = 1 \) and 2 are shown in Figure ??

Solutions of the Dirichlet problem have a remarkable property. Suppose \( C \) is any closed curve lying completely in \( R \). Integrating Laplace’s equation over the area contained within \( C \), and using Gauss’s theorem, we have

\[
\int_C \nabla u \cdot \hat{n} \, dl = 0, \quad (9.18)
\]

where the line integral is around \( C \), and \( \hat{n} \) is the outward unit normal to \( C \). The physical interpretation of (9.18) is that in steady state the flux of heat into and out of a closed region must be zero.

**The maximum principle**

From the intuitive picture of steady-state heat distributions we can anticipate that \( u \) cannot have maximum within \( R \): otherwise heat would continually flow away from the hot-spot. In other words the maximum (and minimum) temperatures are on \( ∂R \). Let’s prove this maximum principle by contradiction: suppose that there is a interior point \( \mathbf{x} ∈ R \) at
Lecture 9. More on Laplace’s equation in \(d=2\)

which \(u\) achieves a local maximum. This implies that \(x\) is surrounded by an isothermal contour \(C\), whose temperature is less than \(u(x)\). The outward normal to this hypothetical isotherm \(C\) is

\[
\hat{n} = -\nabla u / |\nabla u|.
\]

(9.19)

With this expression for \(\hat{n}\), (9.18) becomes:

\[
\oint_C |\nabla u| \, d\ell = 0.
\]

(9.20)

This a contradiction: the integrand on the left is positive definite so the contour integral cannot be zero. We conclude that we cannot find an interior point of \(R\) at which the temperature is a local maximum.

The mean value theorem

Now we show that the temperature at a point inside \(R\) is the average of the temperature at surrounding points. Define an “average temperature” at a point \(x \in R\) as

\[
U(x, r) \equiv \int u(x + r) \, d\theta / 2\pi,
\]

(9.21)

where \(r \equiv r(\cos \theta, \sin \theta)\). We can pick any value of \(r\) so long as \(x + r\) lies within \(R\) for every value of \(\theta\). That is, \(x\) is surrounded by a circular contour \(C\) of radius \(r\) lying properly within \(R\).

Taking the derivative with respect to \(r\) we have

\[
\frac{\partial U}{\partial r} = \oint_C \nabla u \cdot \hat{n} \, d\theta / 2\pi.
\]

(9.22)

But from (9.18), the right hand side is zero i.e., \(U\) is independent of \(r\). Thus we can evaluate \(U\) by taking \(r = 0\) i.e.,

\[
U(x, r) = u(x).
\]

(9.23)

This is the mean value theorem for harmonic functions.

Uniqueness for the Dirichlet problem

It is easy to prove that the Dirichlet problem in a bounded domain \(R\) has a unique solution....
9.3 Laplace’s equation via analytic functions

It is very easy to manufacture solutions of Laplace’s equations in $d = 2$. To do this, we can change variables from $x$ and $y$ to

$$z \equiv x + iy, \quad \text{and} \quad z^* \equiv x - iy$$ \hfill (9.24)

Any function of $x$ and $y$ can be written equivalently as function of $z$ and $z^*$.

With a one-line substitution you can show that

$$u(x, y) = \frac{1}{2}a(z) + \frac{1}{2}b(z^*),$$ \hfill (9.25)

with $a$ and $b$ arbitrary functions, is a solution of Laplace’s equation. If $u$ is to be a real function then $a$ and $b$ aren’t independent: we must have $\Re$ is the real part.

$$u(x, y) = \frac{1}{2}a(z) + \frac{1}{2}a^*(z^*) = \Re(a(z)).$$ \hfill (9.26)

In other words, the real (and also the imaginary) parts of an analytic function are solutions of Laplace’s equations. But because of boundary conditions, it is still not easy to solve the Dirichlet problem.

**Example:** Write $f_1 = x(x^2 + y^2)$ and $f_2 = x^3 - 3y^2x$ as functions of $z$ and $z^*$.

Eliminate $x$ and $y$ using

$$x = \frac{1}{2}(z + z^*), \quad y = \frac{1}{2i}(z - z^*).$$ \hfill (9.27)

This gives

$$f_1 = \frac{1}{2}(z + z^*)zz^*, \quad f_2 = \frac{1}{2}(z^3 + z^*3).$$ \hfill (9.28)

Note $f_2$ is the real part of $z^3$ and is therefore a solution of Laplace’s equation.

**Example:** Find a bounded solution of Laplace’s equation in the half-plane $y > 0$ satisfying the boundary condition

$$u(x, 0) = \frac{1}{a^2 + x^2}.$$ \hfill (9.29)

Let’s guess that $u(x, y)$ might be equal to

$$v(x, y) = \Re \frac{1}{a^2 + z^2} = \frac{a^2 + x^2 - y^2}{(a^2 + x^2 - y^2)^2 + 4x^2y^2}.$$ \hfill (9.30)

This function satisfies the boundary condition, and also Laplace’s equation. But unfortunately $v(x, y)$ doesn’t solve the problem because there is a singularity in the upper half plane at $z = ia$: $v(0, a) = \infty$. So our guess has failed. To solve this problem you’ll have to do the exercises at the end of this lecture.
The Cauchy-Riemann equations

9.4 Laplace’s equation in the upper half plane

The Dirichlet problem

The harmonic function \( \theta \) is very handy if we want to solve the general Dirichlet problem in the upper half-plane \( y > 0 \). Just to be clear, the problem is

\[
 u_{xx} + u_{yy} = 0, \quad u(x,0) = f(x).
\]  
(9.31)

Above, \( f(x) \) is some specified boundary value. We desire a non-singular solution which is bounded as \( y \to \infty \). The Green’s function solution of this problem is given below in (9.44). On the way to this beautiful formula we admire some harmonic scenery.

We begin with discontinuous boundary data:

\[
 f(x) = \text{sgn}(x).
\]  
(9.32)

Since the boundary data does not provide a length scale, similarity reasoning suggests that the solution must have the form:

\[
 u = f \left( \frac{y}{x} \right) = g(\theta).
\]  
(9.33)

But the most general solution of Laplace’s equation which is independent of \( r \) is just \( u = a + b\theta \). Fiddling around with the boundary condition we quickly see that the solution of the problem is

\[
 u = 1 - \frac{2\theta}{\pi}.
\]  
(9.34)

The discontinuity in the \( \text{sgn}(x) \) boundary condition is removed as soon as \( y \) is a little bit positive. (This smoothing is like the erf-solution of the diffusion equation.)

Example: An escapade starts when a drunken criminal is released at \( (x,y) \) with \( y > 0 \). She keeps random walking till she arrives at \( y = 0 \). The police arrest her if she arrives at \( y = 0 \) with \( x > 0 \), while she escapes capture if she arrives at \( y = 0 \) with \( x < 0 \). Find

\[
 p(x,y) = \text{Probability of escape, starting at (x,y)}.
\]  
(9.35)
We know that this probability is determined by the Dirichlet problem
\[ p_{xx} + p_{yy} = 0, \quad p(x,0) = H(-x). \] (9.36)

The positive half of the x-axis is an absorbing boundary. The solution obtained by modifying (9.34) to fit the slightly different boundary condition with the result
\[ p = \frac{1}{2}(1-u) = \frac{\theta}{\pi} \] (9.37)

Now consider the half-plane Dirichlet problem with
\[ f(x,a) = \begin{cases} 1, & \text{if } -a < x < a; \\ 0, & \text{if } |x| > a. \end{cases} \] (9.38)

I’m sure at this point you can supply a probabilistic interpretation. And using linear superposition you can also quickly see that the solution is
\[ u = \frac{\theta^+ - \theta^-}{\pi}, \] (9.39)
where the angles \( \theta^+ \) and \( \theta^- \) are defined by
\[ \theta^\pm = \tan^{-1} \left( \frac{2y}{x \pm a} \right). \] (9.40)

This solution is illustrated in Figure 9.2.

Now we take a \( \delta \)-limit in which \( a \to 0 \) in (9.38). We also have to divide by \( 2a \) so that the boundary condition is
\[ \lim_{a \to 0} \frac{f(x,a)}{2a} = \delta(x). \] (9.41)
The solution of this new problem — the Green’s function of the upper half plane Dirichlet problem — is

\[
g(x, y) = \lim_{a \to 0} \frac{1}{2a\pi} \left( \tan^{-1} \left( \frac{y}{x - a} \right) - \tan^{-1} \left( \frac{y}{x + a} \right) \right),
\]
\[
\partial_x \theta = -\frac{\gamma}{x^2 + y^2}
\]
\[
g(x, y) = \frac{y}{\pi} \frac{1}{x^2 + y^2}.
\]

With the Green’s function in hand, the solution of the half-plane Dirichlet problem posed in (9.51) is

\[
u(x, y) = \int_{-\infty}^{\infty} \frac{y f(\xi)}{(x - \xi)^2 + y^2} \frac{d\xi}{\pi}.
\]

**Example:** Consider a big sheet of metal, which we take to be the half-plane \(y > 0\). Suppose that the temperature at \(y = 0\) is prescribed to be

\[f(x) = \cos kx.
\]

Find the temperature in the sheet.

The fastest way to solve this Dirichlet problem is separation of variables: substituting \(u(x, y) = U(y) \cos kx\) we have

\[U'' - k^2 U = 0, \quad \Rightarrow \quad U = A e^{-ky} + B e^{ky}.
\]

We take \(k > 0\) and discard the exponentially growing solution. Thus the upper half plane solution is

\[u(x, y) = \cos kx e^{-ky} = \Re e^{ikz}.
\]

We might have guessed

\[v(x, y) = \Re \cos kx = \cos kx \cosh ky
\]

as a solution of this problem. Indeed, \(v\) satisfies the boundary condition on \(y = 0\), and \(v\) is a harmonic function. But \(v\) is not bounded as \(y \to \infty\), and so our guess is wrong.

Heat flows into the plate where the temperature is high and flows out where it is cold. For instance, heat fluxes into the plate through the boundary segments where \(\cos kx > 0\) and out of the plate where \(\cos kx < 0\). The net flow of heat through \(y = 0\) is zero — what goes in at one place comes out somewhere else. The thermal disturbance decays at large distances from the boundary, and the decay length, or penetration distance, is \(k^{-1} = \lambda/2\pi\) where \(\lambda\) is the wavelength on the boundary.
In solving this problem we haven’t used the Green’s function formula in (9.44). So it’s sporting to independently check that
\[ \cos kx e^{-ky} = \int_{-\infty}^{\infty} \frac{y \cos k\xi}{(x - \xi)^2 + y^2} \frac{d\xi}{\pi}. \] (9.49)

After a noticeable pause on my computer, mathematica verifies the formula above with the command
\[
\text{Assuming}\{\{k > 0, x > 0, y > 0\}, \\
\text{Integrate}[y \cos[k*a]/((x - a)^2 + y^2), \{a, -\infty, \infty\}]\}
\]

One way to do the integral yourself, is to notice that
\[ \frac{1}{(x - \xi)^2 + y^2} = \int_{0}^{\infty} e^{-\beta[(x-\xi)^2+y^2]} \, d\beta. \] (9.50)

Using this trick one can move the denominator in (9.49) into an exponential. Then switching the order of integration produces standard Gaussian integrals. Dutifully performing this exercise will convince you that separation of variables is the best way to solve this problem.

The Neuman problem

The Neuman problem is
\[ u_{xx} + u_{yy} = 0, \quad u_y(x, 0) = g(x). \] (9.51)

Above, \( g(x) \) is some specified boundary value. We desire a non-singular solution which is bounded as \( y \to \infty \). However our desires are likely frustrated if
\[ \int_{-\infty}^{\infty} g(x) \, dx \neq 0. \] (9.52)

Using the heat-condition analogy, the condition above means there is a net flux of heat into the sheet. In steady state, this has to be conducted away to \( r = \infty \) and this probably requires \( u \sim \ln r \). So perhaps we have to reduce our Neuman expectations to saying that \( u(x, y) \) grows no faster than \( \ln(r) \).

Example: Use the Dirichlet solution in (9.39) to construct a Neuman solution. Suppose we calculate
\[ g(x) \equiv \hat{\partial}_y \left( \frac{\theta^- - \theta^+}{\pi} \right), \quad \text{evaluated at } y = 0. \]

Once we have this \( g(x) \) we can amaze our friends by exhibiting a soluble Neuman problem....
9.5 Minimization of the Dirichlet functional

The solution of Laplace’s equation in the domain $\mathcal{R}$ minimizes the Dirichlet functional

$$D[v] = \int_{\mathcal{R}} |\nabla v|^2 \, dA.$$  \hspace{1cm} (9.53)

Here $v(x, y)$ is any function whose boundary value is the same as that of the harmonic function $u(x, y)$.

**Example:** As an example consider

$$v(x, y) = e^{-\alpha y} \cos kx$$ \hspace{1cm} (9.54)

in the half-plane $y > 0$. The ‘trial function’ $v(x, y)$ is always equal to $\cos kx$ on the boundary $y = 0$. Now we calculate the Dirichlet functional

$$D[v] = \frac{1}{4} L \left( \frac{k^2}{\alpha} + \alpha \right).$$ \hspace{1cm} (9.55)

($L = 2\pi/k$ is one wavelength of the periodic-in-$x$ solution.) Sure enough, the minimum is achieved if $\alpha = k$, and this makes $v$ the solution of Laplace’s equation obtained previously by separation of variables. ■

Minimizing the Dirichlet functional is a good example of the calculus of variations. Suppose the minimum value of $D[v]$ is achieved at $u(x, y)$, and consider how the functional changes if $v = u + \delta v$. Here $\delta v(x, y)$ is a small variation away from the optimal $u(x, y)$. We insist that $u$ and $v$ satisfy the same boundary condition, and therefore $\delta v$ must vanish on the boundary $\partial \mathcal{R}$. Now

$$D[u + \delta v] = D[u] + 2 \int_{\mathcal{R}} \nabla u \cdot \nabla \delta v \, dA + \int_{\mathcal{R}} |\nabla \delta v|^2 \, dA.$$ \hspace{1cm} (9.56)

At the optimal $u$, the first variation must vanish for all possible $\delta v$’s. But integrating by parts

$$2 \int_{\mathcal{R}} \nabla u \cdot \nabla \delta v \, dA = \oint_{\partial \mathcal{R}} \delta v \nabla u \cdot \hat{n} \, d\ell - \int_{\mathcal{R}} \delta v \nabla^2 u \, dA.$$ \hspace{1cm} (9.57)

To make the first variation zero, the final integral must be zero for all $\delta v$, including $\delta v$’s which are non-zero only in the close neighbourhood of an arbitrary point in $\mathcal{R}$. The only way to do this is to make the integrand zero at every point, which is achieved by $\nabla^2 u = 0$. 

SIO203C, W.R. Young, March 21, 2011
9.6 The disk: Poisson’s formula

Consider a circular disk which heated nonuniformly only at the edge \( r = a \). For instance, the problem might be

\[
\nabla^2 u = 0, \quad u(a, \theta) = \sqrt{\theta}.
\]

That is, the top half of the disk \((0 < \theta < \pi)\) is hot and the lower half \((-\pi < \theta < 0)\) is cold. We have a handy Fourier series representation of the boundary condition

\[
\sqrt{\theta} = \frac{4}{\pi} \left[ \sin \theta + \frac{1}{3} \sin 3\theta + \frac{1}{5} \sin 5\theta + \cdots \right].
\]

But, more generally, we consider Laplace’s equation with some arbitrary boundary condition \( u(a, \theta) = f(\theta) \) where

\[
f(\theta) = \sum_{m=-\infty}^{\infty} f_m e^{im\theta}, \quad \text{with} \quad f_m = \oint e^{-im\theta'} f(\theta') \frac{d\theta'}{2\pi}.
\]

With linear superposition in view, we first consider the simpler problem

\[
\nabla^2 u_m = 0, \quad u_m(a, \theta) = e^{im\theta}.
\]

We can guess the solution is separable

\[
u_m(r, \theta) = e^{im\theta} U_m(r) \Rightarrow U_m'' + r^{-1} U_m' - m^2 r^{-2} U_m = 0.
\]

I’m sure you immediately recognize that the ODE above is an Euler equation so that the substitution

\[
U_m = r^q,
\]

produces the general solution:

\[
q = \pm m, \quad \text{or} \quad U_m = Ar^m + Br^{-m}.
\]

The temperature is unlikely to be singular at \( r = 0 \) and so if \( m > 0 \) we must take \( B = 0 \). We secure the \( r = a \) boundary condition by taking \( A = a^{-m} \) and so:

\[
u_m(r) = \left( \frac{r}{a} \right)^m e^{im\theta}, \quad \text{if} \ m > 0.
\]
If \( m < 0 \), then the non-singular solution is
\[
  u_m(r) = \left( \frac{r}{a} \right)^{-m} e^{im\theta}, \quad \text{if } m < 0. \tag{9.66}
\]
We should also mention \( m = 0 \). In this case the two solutions are \( u \) is a constant and \( u \propto \ln r \). We reject the logarithm because of its singularity at \( r = 0 \).

With the boundary condition in (9.60) we can write down the solution
\[
  u(r, \theta) = \sum_{m=-\infty}^{\infty} f_m \left( \frac{r}{a} \right)^{|m|} e^{im\theta}. \tag{9.67}
\]
The \(|m|\) above takes care of both \( m > 0 \) and \( m < 0 \). Something glorious happens if we recall that \( f_m \) is given by the \( \theta' \)-integral in (9.60) and we press on by exchanging integration and summation:

\[
  u(r, \theta) = \sum_{m=-\infty}^{\infty} \left[ \oint e^{-im\theta'} f(\theta') \frac{d\theta'}{2\pi} \right] \left( \frac{r}{a} \right)^{|m|} e^{im\theta},
\]
\[
  = \oint f(\theta') \left[ \sum_{m=-\infty}^{\infty} e^{im(\theta-\theta')} \left( \frac{r}{a} \right)^{|m|} \right] \frac{d\theta'}{2\pi},
\]
\[
  = \oint f(\theta') \left[ \frac{(a^2 - r^2)}{a^2 + r^2 - 2ar \cos(\theta - \theta')} \right] \frac{d\theta'}{2\pi}. \tag{9.68}
\]
This is Poisson’s integral — we can get the solution everywhere inside the disk from the boundary function \( f(\theta) \).

Evaluating (9.69) at \( r = 0 \), we see that the central temperature \( u(0, \theta) \) is equal to the average of \( f(\theta) \). Once again, this is the mean value theorem for harmonic functions.

### 9.7 References

For much more on the discrete Laplace equation, see the little book:

9.8 Problems

Problem 9.1. Let \( m(x) \) be the expected number of coin-tosses before someone wins in Tom and Jerry’s game. (i) Show that
\[
m(x) = \frac{1}{2} m(x - 1) + \frac{1}{2} m(x + 1) + 1.
\]
(ii) Solve this difference equation using the appropriate boundary conditions at \( x = 0 \) and \( N \).

Problem 9.2. In Tom and Jerry’s game, suppose that the probability of a head is \( a \) and the probability of a tail is \( b \) with \( a + b = 1 \). (i) Show that the generalization of (9.2) is
\[
p(x) = ap(x - 1) + bp(x + 1).
\]
(ii) Solve this difference equation (with the appropriate boundary conditions at \( x = 0 \) and \( n \)) by looking for solutions of the form
\[
p = r^x.
\]
(See page 41 of BO for a discussion of constant coefficient difference equations.) (iii) Develop an approximate second-order ODE, analogous to (9.13), assuming that \( n \) is a large \( p(x) \) changes only slightly when \( x \) changes by \( \pm 1 \). (iv) Solve the ODE and numerically compare the approximate solution with the exact solution of the difference equation for selected values of \( a, b \) and \( n \).

Problem 9.3. Find the escape probabilities for the criminal random walking on the lattice in the bottom panel of figure 9.1.

Solution. I got
\[
p_1 = p_5 = \frac{2}{47}, \quad p_2 = p_6 = \frac{3}{47}, \quad p_3 = p_7 = \frac{7}{47}, \quad p_4 = p_8 = \frac{18}{47}.
\]

Problem 9.4. Why is there a minus sign on the right of (9.19)?

Problem 9.5. Is the product of two harmonic functions harmonic?

Problem 9.6. Calculate the integral
\[
\oint_C x^3 - 3xy^2 \, dl,
\]
where \( C \) is the circle \((x - 1)^2 + (y - 3^{-1/2})^2 = 7\).
Problem 9.7. Find a vector field $\mathbf{v}$ for which:

\[
\text{Area enclosed by any closed curve } C = \oint_{C} \mathbf{v} \cdot \mathbf{n} \, d\ell.
\]

Problem 9.8. If $C$ is a closed contour calculate the contour integrals

\[
J_1 = \oint_{C} \mathbf{x} \cdot \mathbf{n} \, d\ell, \quad J_2 = \oint_{C} (x \mathbf{x} - y \mathbf{y}) \cdot \mathbf{n} \, d\ell.
\]

Problem 9.9. Real solutions of Laplace’s equation can be written in the form (9.26). Verify that the functions below are solutions of Laplace’s equation, and find $a(z)$:

\[
u_1 = e^{-ky} \cos kx, \quad u_2 = x^4 - 6x^2y^2 + y^4, \quad u_3 = \frac{x + 1}{(x + 1)^2 + y^2}.
\]

Calculate the area integrals

\[
I_n = \iint_{C} u_n \, d^2x,
\]

where $C$ is the unit circle.

Problem 9.10. Consider an annular sheet of metal with the temperature of the inner boundary ($r = a$) fixed at temperature $u_0$ and the outer boundary ($r = b > a$) at temperature $u_1$ (see Figure 9.3). (i) Solve
Lecture 9. More on Laplace’s equation in $d=2$

Laplace’s equation and thus find $u(x)$. (ii) Apply Gauss’s theorem to the two shaded areas in Figure 9.3 and thus calculate

$$J_n \equiv \oint_{C_n} \nabla u \cdot \hat{n} \, d\ell.$$

**Problem 9.11.** An ensemble of random walkers start at a radius $r$ in the annular domain shown in Figure 9.3. Those that first reach the inner boundary at $r = a$ are captured, while those that first reach the outer boundary at $r = b$ escape. Find the probability of escape.

**Problem 9.12.** (i) Show that both

$$u_1 \equiv \cos kx \cosh ky, \quad \text{and} \quad u_2 \equiv \cos kx e^{-ky}$$

are UHP solutions of the Dirichlet problem $\nabla^2 u = 0$ with $u(x,0) = \cos kx$. (ii) Why doesn’t this example contradict the uniqueness proof.

**Problem 9.13.** Consider heat conduction in a sheet of metal $\mathcal{R}$ which is uniformly heated. Denote the area of the sheet by $A$. Suppose that the temperature at the boundary is $u = 0$. Poisson’s equation for the steady-state temperature distribution, $u(x,y)$, is

$$0 = \kappa \nabla^2 u + \frac{h}{\text{heating}},$$

where the constant $h$ is the uniform heating. (i) Calculate the flux of heat through the boundary of the sheet. (ii) Define an “average temperature” at a point $x$ as

$$U(x,r) \equiv \int u(x + r) \frac{d\theta}{2\pi},$$

where $r = r(\cos \theta, \sin \theta)$. On physical grounds, do you expect $U(x,r)$ to be greater or less than $u(x)$? (iii) Prove that

$$U(x,r) = u(x) - \frac{hr^2}{4\kappa}.$$

**Problem 9.14.** Let $\phi(x,y)$ be the angle subtended by a line segment in the plane at the point $x = (x,y)$. Show that $\phi$ is harmonic (except at the end-points of the segment).

**Problem 9.15.** (i) Solve the Dirichlet-Poisson problem

$$\nabla^2 u = e^{-y} \cos kx, \quad u(x,0) = 0,$$
Lecture 9. More on Laplace’s equation in $d=2$

in the upper half plane. (ii) Solve the Neuman-Poisson problem
\[ \nabla^2 u = e^{-y} \cos kx, \quad u_y(x, 0) = 0. \]

Make sure you discuss the case $k = 0$.

**Problem 9.16.** (i) Consider heat conduction in a circular disk of metal, $0 < r < a$. Solve Laplace’s equation for the steady-state temperature distribution:
\[ \nabla^2 u = 0, \]
assuming that the boundary condition at $r = a$ is
\[ u_1(a, \theta) = 1, \quad u_2(a, \theta) = \cos \theta, \quad u_3r(a, \theta) = 1, \quad u_4r(a, \theta) = \cos \theta. \]
In one of the four cases the problem “obviously” has no solution — so don’t be surprised if you can only solve three of the four. (ii) After you have identified the case with no steady solution, continue to find an unsteady solution of the diffusion equation
\[ u_t = \kappa \nabla^2 u \] (9.70)
with the offending boundary condition.

**Problem 9.17.** Consider heat conduction in a circular disk of metal, $0 < r < a$. Show that the steady state Poisson’s equation distribution:
\[ \nabla^2 u = -1, \]
with the no-flux boundary condition
\[ u_r(a, \theta) = 0 \] (9.71)
has no solution. Find a solution of the full diffusion equation
\[ u_t - \kappa \nabla^2 u = 1, \]
with the “no-flux” boundary condition (9.71).

**Problem 9.18.** (i) Find a two real solutions of Laplace’s equation by fiddling around with the complex function
\[ A(z) = \frac{1}{z}, \quad z = x + iy. \]

(ii) Using the solutions from part (i) as a building block, find a bounded solution of Laplace’s equation in the half-plane $y > 0$ with the boundary condition
\[ u(x, 0) = \frac{1}{a^2 + x^2}. \]
Lecture 9. More on Laplace’s equation in d=2

Problem 9.19. (i) Find an approximate solution \( v(x, y) \) of the half-plane \((y > 0)\) Dirichlet problem

\[ \nabla^2 u = 0, \quad u(x, 0) = \frac{1}{x^2 + a^2}, \]

by minimizing the Dirichlet function with an optimal choice of \( q \) in the test function

\[ v(x, y) = \frac{e^{-qy}}{x^2 + a^2}. \]

(ii) If you’ve also solved the previous problem, compare the exact and approximate values of the Dirichlet functional i.e., calculate \( D[u] \) and \( D[v] \). To check your answer show that \( D[v] = \sqrt{2}D[u] \). (iii) Bonus if you can find a simple test function \( w(x, y) \) with \( D[w] < D[v] \).

Problem 9.20. (i) Show that if \( u(x, y) \) is a solution of either the Dirichlet or Neuman problems in a region \( R \), then

\[ D[u] = \oint_{\partial R} u \nabla \cdot \hat{n} \, d\ell. \]

\[ D[u] = \int |\nabla u|^2 \, dA. \]

(ii) Suppose \( v(x, y) \) satisfies the same Neuman or Dirichlet boundary condition as the harmonic function \( u(x, y) \) on \( \partial R \). Starting from

\[ \int_R |\nabla (u - v)|^2 \, dA \geq 0, \]

show that \( D[v] \geq D[u] \).

Problem 9.21. (i) Solve Laplace’s equation \( \nabla^2 u = 0 \) for \( r < a \) with the boundary condition \( u(a, \theta) = x^4 \). (ii) With the boundary condition \( u(a, \theta) = \sin^3 \theta \).

Problem 9.22. Suppose \( u(r, \theta) \) is the solution of Laplace’s equation on a disc with \( r < a \), with boundary value \( u(a, \theta) = f(\theta) \). Find

\[ \tilde{u}(r) \equiv \oint u(r, \theta) \frac{d\theta}{2\pi} \]

in terms of \( f(\theta) \).

Problem 9.23. Suppose \( u(r, \theta) \) is the solution of Laplace’s equation on a disc with \( r < a \), with boundary value \( u(a, \theta) = f(\theta) \), and \( v(r, \theta) \) is defined by the Poisson equation

\[ \nabla^2 v = u, \quad v(a, \theta) = 0. \]
Lecture 9. More on Laplace’s equation in d=2

Find

\[ \tilde{v}(r) \equiv \oint v(r, \theta) \frac{d\theta}{2\pi} \]

in terms of \( f(\theta) \).

**Problem 9.24.** Show that if \(|\rho| < 1\)

\[ \sum_{m=-\infty}^{\infty} \rho|m|e^{im\phi} = \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos \phi}. \]

**Problem 9.25.** Consider \( f(\theta) = 1 \) in Poisson’s integral. From the solution of the diffusion equation deduce that

\[ \oint \frac{d\alpha}{A + B \cos \alpha} = \frac{2\pi}{(A^2 - B^2)^{1/2}}. \]

Find other results of this nature by choosing \( f(\theta) \) so that you can easily solve the diffusion equation. (This problem is adapted from CP.)

**Problem 9.26.** Derive a formula analogous to Poisson’s integral for Laplace’s equation exterior to the disc (i.e., \( r > a \)) with the Dirichlet boundary condition \( u(a, \theta) = f(\theta) \).

**Problem 9.27.** Derive a formula analogous to Poisson’s equation assuming that the boundary condition is fixed-flux

\[ u_r(a, \theta) = f(\theta). \]

(Notice that \( f(\theta) \) must satisfy a “solvability condition” if this problem is to have a solution.)

**Problem 9.28.** Solve Laplace’s equation in the disc (radius \( a \)) with the boundary condition \( u(a, \theta) = \text{sqr}(\theta) \).

**Solution.** GRI says:

\[ u(r, \theta) = \frac{2}{\pi} \tan^{-1} \left( \frac{2ar \sin \theta}{a^2 - r^2} \right). \] (9.72)

This looks to be right!
Lecture 10

Shocks and conservation laws

Now that we know a bit about diffusion we return to the nonlinear advection equation
\[ u_t + uu_x = 0, \quad (10.1) \]
and its diffusive relative, Burgers’ equation:
\[ u_t + uu_x = \nu u_{xx}. \quad (10.2) \]

10.1 Multivalued solutions and shock propagation

What do multivalued solutions mean? The mathematics is fine: as we saw in earlier lecture, a multivalued function \( u(x, t) \) can be a valid mathematical solution of a nonlinear PDE. But in physical problems the appearance of a shock usually signals the failure of the continuum approximation. In lecture 1 we started with the integral conservation law
\[ \frac{d}{dt} \int_a^b \rho(x, t) \, dx + f(b, t) - f(a, t) = 0, \quad (10.3) \]
and then we took the limit \( a \to b \). We assumed that \( \rho \) is continuous and \( f \) differentiable. This assumption is essential in obtaining the differential statement of the conservation law:
\[ \rho_t + f_x = 0. \quad (10.4) \]
The transition from (10.3) to (10.4) is not possible once the solution becomes nonsmooth e.g., by forming a shock. On physical grounds, after
the shock forms, we expect that further evolution involves the dynamics of discontinuities called *shock waves*.

To determine the motion of a shock wave we insist that the integral form of the conservation law is valid at all time. We also assume that if \( \rho(x,t) \) is discontinuous at \( x = s(t) \) then there are still well defined left and right limits, \( \rho(s+,t) \) and \( \rho(s-,t) \). Figure 10.1 — see figure 10.1. To determine \( s(t) \) we break-up (10.3) like this:

\[
\frac{d}{dt} \left[ \int_s^a \rho(x,t) \, dx + \int_s^b \rho(x,t) \, dx \right] + f(b,t) - f(a,t) = 0.
\]

(10.5)

Then, recalling Leibnitz’s rule for differentiating integrals, we have

\[
\int_s^a \rho_t(x,t) \, dx + \int_s^b \rho_t(x,t) \, dx + \dot{s} \left[ \rho(s-,t) - \rho(s+,t) \right] + f(b,t) - f(a,t) = 0.
\]

(10.6)

Now we let \( a \) and \( b \) approach \( s(t) \). The first two terms in (10.6) vanish because \( \rho \) is bounded. Therefore we get

\[
-\dot{s}[[\rho]] + [[f]] = 0, \quad \text{or} \quad \dot{s} = \frac{[[f]]}{[[\rho]]}.
\]

(10.7)

Equation (10.7) is called the *shock condition*; it relates the speed of a shock to the conditions immediately behind and in front of the shock. The jump in \( \rho, [[\rho]] \), is the *strength* of the shock and the spacetime curve \( x = s(t) \) is the *shock path*. 

**Figure 10.1: Derivation of the shock condition in (10.7). shockCond.png**
10.2 Two examples of shock propagation

Consider the nonlinear advection equation

\[ u_t + uu_x = 0, \quad u(x, 0) = H(-x). \tag{10.8} \]

The shock is already present in the initial condition at \( x = 0 \). Since \( f = u^2/2 \) the shock speed is

\[ \dot{s} = \frac{\frac{1}{2}u_+^2 - \frac{1}{2}u_-^2}{u_+ - u_-} = \frac{1}{2}(u_+ + u_-). \tag{10.9} \]

The initial condition has \( u_+ = 0 \) and \( u_- = 1 \) so the shock speed is \( \dot{s} = 1/2 \) and the shock path is \( s = t/2 \). The characteristic diagram is shown in figure 10.2. Notice particularly that characteristics flow into the shock from both sides. This is an essential feature of shock waves: shocks swallow characteristics and prevent the solution becoming multivalued.

Now consider a slightly different nonlinear advection equation

\[ v_t + v^2 v_x = 0, \quad v(x, 0) = H(-x). \tag{10.10} \]

Again, the shock forms instantly because the initial condition is discontinuous. In this case, with \( f = v^3/3 \), the shock speed is

\[ \dot{s} = \frac{\frac{1}{3}v_+^3 - \frac{1}{3}v_-^3}{u_+ - u_-} = \frac{1}{3}(v_+^2 + v_+ v_- + v_-^2). \tag{10.11} \]

Since the initial condition has \( v_+ = 0 \) and \( v_- = 1 \) the shock speed is now \( \dot{s} = 1/3 \) and the shock path is \( s = t/3 \).
Suppose we change variables in (10.10) by introducing \( v = \sqrt{w} \). Since
\[
v_t = \frac{1}{2} w_t / \sqrt{w} \quad \text{and} \quad v^2 v_x = \frac{1}{2} \sqrt{w} w_x,
\]
we get
\[
v_t + v^2 v_x = 0 \Rightarrow w_t + w w_x = 0. \tag{10.13}
\]
Also \( w(x,0) = v^2(x,0) = H(-x) \). It seems that the PDE for \( w(x,t) \) is the same as the PDE for \( u(x,t) \) in (10.8): we are tempted to conclude that \( w = u \) and therefore \( v = \sqrt{u} \)? But this can’t be true — the \( v \)-shock travels with speed \( \dot{s} = 1/3 \), while the \( u \)-shock travels with speed \( \dot{s} = 1/2 \)? The point is that the shock condition (10.7) provides additional physical information which removes the mathematical ambiguity which occurs if the density is discontinuous. In other words, we can’t guess the correct shock condition by staring at a naked PDE: the integral form of the conservation law provides fundamental physical information which is simply lost in the transition from (10.3) to (10.4).

### 10.3 Shocks from continuous initial conditions

**A kinky initial condition**

In our earlier example (10.8) the shock was already present in the initial condition. Now let’s consider the kinky initial condition from the previous lecture:
\[
u = \begin{cases} 
1 - |x|, & \text{if } |x| < 1, \\
0, & \text{if } |x| > 1.
\end{cases} \tag{10.14}
\]

This initial condition is continuous but nondifferentiable at three points. We know that a shock forms at \((x_s,t_s) = (1,1)\) (see figure 3.2). Let’s now find the shock path and strength.

Before the shock forms, the solution is shown in the upper panel of figure 10.3. At \( t = 1 \) the right face of the triangular wave becomes vertical and the shock appears. This vertical face (a discontinuity in \( u \)) then propagates to the right, into the region where \( u = 0 \). Thus, ahead of the shock \( u(s^+,t) = 0 \). Behind the shock we can continue to use the solution in the upper panel of figure 10.3
\[
u(x,t) = \frac{1 + x}{1 + t}, \tag{10.15}
\]
Figure 10.3: Upper panel: the steepening profile at $t = 1/2$ before shock formation. Middle panel: the characteristic diagram showing the shock path. Bottom panel: the triangular wave. kinkShock.eps
so that \( u(s^-, t) = (s + 1)/(1 + t) \). Ahead of the shock \( u(s^+, t) = 0 \). Therefore the shock condition is

\[
\dot{s} = \frac{1}{2} \frac{s + 1}{t + 1}.
\] (10.16)

We have to solve this ODE for \( s(t) \) subject to the initial condition that \( s(1) = 1 \). Multiplying by the integrating factor \( 1/\sqrt{1 + t} \) we have

\[
\frac{d}{dt} \left( \frac{s}{\sqrt{1 + t}} \right) = \frac{1}{2(1 + t)^{3/2}}.
\] (10.17)

Integrating from \( t' = 1 \) to \( t' = t \), and applying \( s(1) = 1 \), we find

\[
s(t) = \sqrt{2(1 + t)} - 1.
\] (10.18)

The shock path is the heavy curve in the characteristic diagram shown in the middle panel of figure [10.3]. Characteristics flow into both sides of the shock. The solution is the triangular wave shown in the bottom panel of figure [10.3].

The strength of the shock is

\[
[[u]] = \frac{1 + s}{1 + t} = \sqrt{\frac{2}{1 + t}}.
\] (10.19)

Thus we can easily see that the area under the triangular wave in the bottom panel of figure [10.3] is constant

\[
\int_{-\infty}^{\infty} u(x, t) \, dx = \frac{1}{2} \times [s(t) + 1] \times [[u]] = 1.
\] (10.20)

Of course this must be the case since \( u(x, t) \) is a conservative density. On the other hand, we can also compute the ‘energy’ of the solution

\[
\int_{-\infty}^{\infty} u^2 \, dx = \int_{-1}^{s(t)} \frac{(1 + x)^2}{(1 + t)^2} \, dx = \frac{4}{3} \frac{1}{\sqrt{2(1 + t)}} \quad \text{(if } t > 1\text{).}
\] (10.21)

Thus once, the shock forms, when \( t > 1 \), the energy decreases monotonically. How does the energy evolve before \( t = 1 \)?

**Smooth initial condition and Whitham’s geometric construction**

Now suppose the initial condition is \( u(x, 0) = F(x) \) where \( F \) is a completely differentiable function, such as our earlier example, \( F(x) = 1/(1 + \)
x^2). In that example the shock forms at \((x_s, t_s, \xi_s) = (\sqrt{3}, 8\sqrt{3}/9, 1/\sqrt{3})\).

We also know that this is a generic property of the evolution of smooth initial humps: nonlinear steepening produces a shock at \((x_s, t_s, \xi_s)\) and these three unknowns are determined from

\[ F''(\xi_s) = 0, \quad 1 + t_s F'(\xi_s) = 0, \quad x_s = \xi_s + t_s F(\xi_s). \quad (10.22) \]

Where does the shock go after that? At time \(t\) the shock is swallowing two characteristics, \(\xi_-(t)\) behind the shock and \(\xi_+(t)\) ahead of the shock. The velocities before and after the shock are therefore

\[ u(s^-, t) = F(\xi_-), \quad \text{and} \quad u(s^+, t) = F(\xi_+). \quad (10.23) \]

Thus to determine the shock path we must solve the following system

\[ s(t) = \xi_+ + t F(\xi_+), \quad s(t) = \xi_- + t F(\xi_-), \quad (10.24) \]

and

\[ \dot{s} = \frac{1}{2} [F(\xi_-) + F(\xi_+)] \quad (10.25) \]

for the three unknowns \([s(t), \xi_-(t), \xi_+(t)]\). This is a much more difficult problem than our earlier examples because we cannot easily solve for \(\xi_-\) and \(\xi_+\) in terms of \(s(t)\). (Try it with \(F(\xi) = 1/(1 + \xi^2)\) and see how far you can get!)

Fortunately, Whitham (1974) has formulated a set of geometric arguments which reveal the structure of the solution without getting involved in algebraic details. There are two principles. First there is the equal-area rule: on the multi-valued solution at time \(t\) the shock is located by by drawing a vertical line which cuts off equal area lobes (see figure 10.4).

The second principle is the chordal construction: To find \(\xi_-\) and \(\xi_+\) at time \(t\) draw a chord of slope \(-1/t\) on the graph of \(F(\xi)\). Translate this line vertically till it cuts-off equal-area lobes (see figure 10.5).
One interesting result is that as \( t \to \infty \) the hump evolves into a triangular wave, essentially identical to the bottom panel in figure 10.3. Thus with Whitham’s rules we have a complete understanding of how a hump evolves even without suffering through the solution of (10.24) and (10.25).

### 10.4 Burgers’ equation and travelling waves

Perhaps you are alarmed by the ambiguity inherent in PDE models? Fortunately in most physical applications it is easy to understand which variables are conservative. Our favourite equations usually contain dissipative terms. For instance, as an analog of fluid mechanics, Burgers introduced the model equation

\[
  u_t + uu_x = \nu u_{xx}.
\]

If the viscosity \( \nu \) is very small we might be tempted to drop the diffusive term on the right hand side. Then we get the advection equation as an approximation of Burgers’ equation. But the advection equation by itself is ambiguous because, for instance,

\[
  u_t + uu_x = 0 \quad \iff \quad \left( \frac{1}{2} u^2 \right)_t + \left( \frac{1}{3} u^3 \right)_x = 0.
\]

The different forms have different shock conditions. But if the original problem was (10.14) then the correct shock condition is obtained using the flux \( f = u^2/2 \) in the \( u \)-equation (not the \( u^2 \)-equation).

If we do form the \( u^2 \)-equation from Burgers, here is what happens:

\[
  \left( \frac{1}{2} u^2 \right)_t + \left( \frac{1}{3} u^3 \right)_x = \nu (uu_x)_x - \nu u_x^2.
\]

Figure 10.5: Whitham’s chordal construction. chord.eps
Considering a humplike initial condition, so that the fluxes all vanish at \( x = \pm \infty \), we see that Burgers’ equation dissipates ‘energy’

\[
E(t) \equiv \frac{1}{2} \int_{-\infty}^{\infty} u^2 \, dx , \quad \text{and from (10.28):} \quad \dot{E} = -\nu \int_{-\infty}^{\infty} u_x^2 \, dx . \tag{10.29}
\]

It turns out that \( \dot{E} \) is nonzero even in the limit \( \nu \to 0 \) (see the problems).

The traveling wave solution

We can understand the role of the right hand side in (10.14) by looking for a traveling wave solution of Burgers’ equation. We have in mind a smooth solution, as shown in figure 10.6 in which \( u \) varies gradually between \( u(-\infty, t) = L \) to \( u(+\infty, t) = R \). We introduce the traveling wave guess

\[
u = U(z) \, , \quad z \equiv x - ct , \tag{10.30}
\]

into (10.26):

\[
-cU' + UU' = \nu U'' \Rightarrow A - cU + \frac{1}{2} U^2 = \nu U' . \tag{10.31}
\]

Requiring that \( u(-\infty, t) = L \) and \( u(+\infty, t) = R \), we get

\[
A = \frac{RL}{2} \quad \text{and} \quad c = \frac{R + L}{2} . \tag{10.32}
\]

At this point the expression for the speed, \( c \), should remind you of the shock condition for the advection equation in (10.9)!

We press on by rearranging (10.31) as

\[
\nu U' = -\frac{1}{2}(R-U)(U-L) \Rightarrow \frac{1}{U-L} \left( \frac{1}{R-U} + \frac{1}{U-L} \right) = -\frac{dz}{2\nu} . \tag{10.33}
\]

Integrating

\[
\frac{U - L}{R - U} = e^{\mu z} , \quad \text{where} \quad \mu \equiv \frac{L - R}{2\nu} . \tag{10.34}
\]

Solving for \( U(z) \) we finally have

\[
u(x,t) = \frac{L + Re^{\mu(x-ct)}}{1 + e^{\mu(x-ct)}} . \tag{10.35}
\]

The width of the transition region is

\[
\mu^{-1} = \frac{2\nu}{L - R} . \tag{10.36}
\]
Thus as $\nu \to 0$ the solution becomes a discontinuous shock. In other words, as $\nu \to 0$, the transition between $u = L$ and $u = R$ is an internal boundary layer. If we have a gentleman’s agreement that the details of this boundary layer are of no interest then we can set the right hand side of (10.26) to zero and solve the advection equation while using

$$\dot{s} = \frac{1}{2} u_+ + \frac{1}{2} u_-$$  \hspace{1cm} (10.37)

to track the location of the shocks. We are arguing heuristically that if $\nu$ is very small then $u(x,t)$ will vary slowly relative to the shock thickness in (10.36).

**The Cole-Hopf transformation**

First substitute $u = \phi_x$ into Burgers’ equation (10.26). We can then integrate with respect to $x$:

$$\phi_t + \frac{1}{2} \phi_x^2 = \nu \phi_{xx}.$$

(10.38)

Now we substitute $\phi = \alpha \ln \theta$; notice

$$\phi_x = \alpha \frac{\partial x}{\partial} \theta, \quad \phi_{xx} = \alpha \frac{\partial x}{\partial} \theta - \alpha \left( \frac{\partial x}{\partial} \theta \right)^2 = \alpha \frac{\partial x}{\partial} \theta - \frac{1}{\alpha} \phi_x^2.$$

(10.39)

We find then

$$\alpha \frac{\partial x}{\partial} \theta + \left( \frac{1}{2} + \frac{\nu}{\alpha} \right) \phi_x^2 = \nu \alpha \frac{\partial x}{\partial} \theta.$$

(10.40)
Lecture 10. Shocks and conservation laws

Picking $\alpha = -2\nu$ we destroy the nonlinear term and find that $\theta$ satisfies the heat equation

$$\frac{\partial \theta}{\partial t} = \nu \frac{\partial^2 \theta}{\partial x^2}. \quad (10.41)$$

To summarize, the substitution

$$u = \phi_x = -2\nu \frac{\partial x}{\theta} \quad (10.42)$$

enables us to move back and forth between Burgers’ equation and the heat equation.

As an example, it is easy to guess that

$$\theta = x^2 + 2\nu t \quad (10.43)$$

is a solution of the heat equation. Thus it seems that

$$u = -2\nu \frac{2x}{x^2 + 2\nu t} \quad (10.44)$$

must be a not-so-obvious solution of Burgers’ equation.

10.5 Problems

**Problem 10.1.** Assuming that $u$ is a conserved density, find the location and the strength of the shock which forms from

$$u_t + uu_x = 0, \quad u(x, 0) = -\frac{x}{1 + x^2}.$$

**Problem 10.2.** Assuming that $u$ is a conserved density, find the locations and the strengths of the shocks which forms from

$$u_t + uu_x = 0, \quad u(x, 0) = -\sin x.$$

**Problem 10.3.** Consider

$$u_t + (u^2)_x = 0, \quad u(x, 0) = H(-x)\sqrt{-x}.$$

(Notice the wave speed is $2u$ in this problem.) Assume $u(x, t)$ is a conserved density. (i) Show that

$$\int_0^\infty u(x, t) \, dx = \int_0^t u^2(0, t') \, dt'.$$
(ii) Draw the characteristic diagram and find where the shock first forms. (iii) Solve the PDE behind and ahead of the shock. (iv) Find an ODE which determines the shock position, $s(t)$, and solve this equation. (I found it was easy to guess the solution of the ODE.) (v) Check your answer by verifying the result from part (i).

**Problem 10.4.** Suppose that the initial condition $F(x)$ is a smooth function. Investigate the solution of (10.24) and (10.25) close to the initial condition $(x, t, \xi) = (x_s, t_s, \xi_s)$. Show that if $t = t_s + \tau$, with $\tau \ll 1$, then

$$s(t) \approx x_s + F_s \tau + \sqrt[3]{\frac{8 F_s F_s^3}{3 F_{ss}}} \tau^{3/2} + \ldots$$

where $F_s \equiv F(\xi_s)$, $F_s' \equiv F'(\xi_s)$ etc.

**Problem 10.5.** Consider the nonlinear signaling problem with $x > 0$ and $t > 0$:

$$u_t + uu_x = 0, \quad u(x, 0) = 1, \quad u(0, t) = 1 + t.$$  

(i) Sketch the characteristic diagram and show that a shock forms at $(x_s, t_s) = (1, 1)$. (ii) Assuming that $u(x, t)$ is a conserved density, show that the shock path is $s(t) = (t + 3)(3t + 1)/16$, with $t > 1$.

**Problem 10.6.** Consider the signaling problem with $x > 0$ and $t > 0$:

$$u_t + uu_x = 0, \quad u(x, 0) = 0, \quad u(0, t) = a(t).$$

Let us use three different models for the boundary condition at $x = 0$:

$$a_1(t) = 1, \quad a_2(t) = \frac{1}{1 + t}, \quad a_3(t) = t.$$  

(i) Assuming that $u(x, t)$ is a conserved density, show that:

$$\int_0^\infty u(x, t) \, dx = \frac{1}{2} \int_0^t a^2(t') \, dt'.$$

(ii) Construct a single-valued solution using the shock condition to ensure that the global conservation law above is satisfied. (iii) Find another $a(t)$ for which you can calculate $s(t)$.

**Solution.** The location of the shock, $x = s(t)$, in the three cases is

$$s_1(t) = t/2, \quad s_2(t) = \sqrt{1 + t} - 1, \quad s_3(t) = 3t^2/16.$$
Problem 10.7. Consider Burgers’ equation
\[ u_t + uu_x = \nu u_{xx}, \]
with the kinky initial condition
\[ u = \begin{cases} 1 - |x|, & \text{if } |x| < 1, \\ 0, & \text{if } |x| > 1. \end{cases} \]
If we take the limit \( \nu \to 0 \) we obtain the problem discussed in section 10.3. (i) Find the thickness of the internal boundary layer which smoothes out the shock at \( x = s(t) \). (ii) We saw in (10.21) that \( E(t) \propto t^{-1/2} \) as \( t \to \infty \). On the other hand, from (10.29)
\[ \dot{E} = -\nu \int_{-\infty}^{\infty} u_x^2 \, dx. \]
(10.45)
Explain carefully how \( E(t) \propto t^{-1/2} \) is consistent with (10.45) as \( \nu \to 0 \).

Problem 10.8. The hamburger equation
\[ u_t + u^2 u_x = \epsilon u_{xx}, \quad \lim_{x \to +\infty} u(x,t) = 0, \quad \lim_{x \to -\infty} u(x,t) = 1, \]
has a travelling wave solution in which \( u(x,t) = U(\xi) \) with \( \xi = x - ct \). (i) Find the value of \( c \) and show that
\[ -U + U^3 = 3\epsilon U' \]
where the prime denotes an \( \xi \)-derivative. (ii) Solve the ODE above and give an explicit expression for \( U(\xi) \).

Problem 10.9. (i) Show by substitution that if \( u(x,t) \) is a solution of Burgers’ equation then so is \( v(x,t) = u(x - ct, t) + c \). (ii) Show that the problem
\[ u_t + uu_x = \nu u_{xx}, \]
with boundary conditions
\[ \lim_{x \to -\infty} u(x,t) = L, \quad \lim_{x \to -\infty} u(x,t) = R, \]
can be reduced to the parameterless form:
\[ U_t + UU_X = U_{XX}, \quad \text{with} \quad \lim_{X \to \pm \infty} U(X,T) = \pm 1. \]
(iii) Use this transformation to deduce that the shock thickness is given by (10.36) without solving any PDEs.
Problem 10.10. Consider the cheeseburger equation

\[ u_t + uu_x = -\nu u_{xxx}, \]

where \( \nu > 0 \). (i) Explain the sign of the rhs. (ii) Slavishly follow the Burgers' travelling wave calculation — starting in (10.30) — and see how far you can get in finding a travelling wave solution of the cheeseburger equation. Make sure that show that the wave speed is given by \( c = (R + L)/2 \). (iii) Use the method in problem 10.9 to reduce the cheeseburger equation to a parameterless form. Deduce the thickness of the cheeseburger shock in terms of \( R, L \) and \( \nu \).

Problem 10.11. Construct some polynomial solutions of the heat equation, beginning like

\[ \theta = x^3 + \ldots, \quad \text{and} \quad \theta = x^4 + \ldots \]

Use the Cole-Hopf transformation to generate some less obvious solutions of Burgers equation.

Problem 10.12. Show that

\[ \theta(x, t) = A e^{\kappa q^2 t - qx} + B e^{\kappa t p^2 - px} \]

is a solution of the diffusion equation. Use the Cole-Hopf transformation to find a solution of Burgers equation.

Problem 10.13. Consider the problem of finding a travelling-wave solution, \( u(x, t) = u(z) \), with \( z \equiv x - ct \), of the elastic-medium equation

\[ u_{tt} = u_{xx} + u_x u_{xx} + u_{xxxx}, \quad \lim_{x \to -\infty} u = L, \quad \lim_{x \to \infty} u = R. \]

(i) Show that \( v = u_z \) satisfies

\[ v_z^2 = (c^2 - 1) v^2 - \frac{1}{3} v^3. \]

(ii) Solve the ODE above by substituting

\[ v = a \text{sech}^2 (pz) \]

and determining \( a \) and \( p \) in terms of the wave speed, \( c \). (iii) Finally, determine \( c \) by ensuring that \( u \) has the correct limiting behaviour as \( x \to \pm \infty \).
Lecture 11

The Lighthill-Whitham traffic model

A Google search using “Lighthill” and “Whitham” indicates that one of the most popular applications of quasilinear PDEs is the theory of traffic flow developed by Lighthill & Whitham.

11.1 Formulating the model

$\rho(x,t)$ is the density of cars on a highway; with this definition $\rho(x,t)$ cannot become multivalued (or negative). We start with the differential form of the conservation law

$$\rho_t + f_x = 0, \quad (11.1)$$

and argue that the flux is

$$f = \rho u(\rho), \quad (11.2)$$

where $u(\rho)$ is the average speed of cars if the traffic density is $\rho$. We make the assumption that drivers slow down as the density increases, and that they stop completely once the density exceeds $\rho_J$, where the subscript J stand for "jam".

A simple model with the behaviour above is

$$u = \begin{cases} u_m \left(1 - \frac{\rho}{\rho_J}\right), & \text{if } \rho < \rho_J, \\ 0, & \text{if } \rho > \rho_J. \end{cases} \quad (11.3)$$
Lecture 11. The Lighthill-Whitham traffic model

Figure 11.1: The fundamental diagram of road traffic: if we plot the flux $f$ as a function of $\rho$ then both $u = f/\rho$ and $c = df/d\rho$ are the slopes of the lines shown above. Heavy traffic is when the density is greater than the optimal density. fundDiag.eps

The maximum speed, characteristic only of low densities, is $u_m$. With this assumption about human behaviour, the relation between flux to density is

$$f = u_m \rho \left(1 - \frac{\rho}{\rho_J}\right), \quad \text{provided that } \rho < \rho_J. \quad (11.4)$$

The flux is a maximum,

$$f_m = \frac{1}{4} u_m \rho_J, \quad (11.5)$$

at $\rho = \rho_J/2$. According Haberman, in the Lincoln Tunnel the maximum flux is observed to be about 1600 cars per hour at a speed of around 20 miles per hour and a density of 80 cars per mile. If $\rho$ is greater than this optimal density we say that the traffic is heavy and if $\rho$ is less than the optimal then the traffic is light. The maximum flux is the capacity of the road (cars per second).

To summarize: we use the simple model in (11.4) and rewrite the conservation law (11.1) as

$$\rho_t + c(\rho) \rho_x = 0 \quad c(\rho) \equiv f_\rho = u_m \left(1 - \frac{2 \rho}{\rho_J}\right). \quad (11.6)$$

The Lincoln Tunnel connects New York to New Jersey under the Hudson River. The tunnel is about 1.5 miles long. Construction was funded by the New Deal’s Public Works Administration, and started in 1934.
Lecture 11. The Lighthill-Whitham traffic model

\( c(\rho) \) is the speed with which small disturbances propagate through a stream of traffic with uniform density \( \rho \). Notice that \( c(\rho) \) is different from the speed of a car — that’s \( u(\rho) \). Figure 11.1 displays both \( c(\rho) \), \( u(\rho) \) and \( f(\rho) \) graphically. This is the “fundamental diagram” of road traffic.

**Small disturbances on a uniform stream**

We study small disturbances propagating through a uniform stream (density \( \rho_0 \)) by linearizing (11.6)

\[
\rho = \rho_0 + \eta, \quad \Rightarrow \quad \eta_t + c(\rho_0)\eta_x + O(\eta^2) = 0. \tag{11.7}
\]

Neglecting the quadratic terms, the disturbance \( \eta \) satisfies the linear wave equation and consequently:

\[
\eta = \eta(x - c_0 t). \tag{11.8}
\]

The speed \( c_0 \equiv c(\rho_0) \) can be either positive or negative depending on whether the traffic is either light (\( \rho_0 < \rho_j/2 \)) or heavy (\( \rho_0 > \rho_j/2 \)).

Notice that \( u(\rho) > c(\rho) \). This agrees with the experience of highway drivers: one overtakes disturbances in the traffic stream. (You can see the wave of red brake lights moving towards you.) And in heavy traffic the density wave actually moves backwards relative to the road (i.e., \( c(\rho_0) < 0 \)) even though no car is moving backwards.

These remarks emphasize that there are two important velocities \( c(\rho) \) and \( u(\rho) \).

**11.2 The green light problem**

Suppose that there is a block of cars stopped at a red light located at \( x = 0 \). At \( t = 0 \) the light turns green. What happens? The initial density in this situation is

\[
\rho(x, 0) = F(x) \equiv \begin{cases} \rho_j, & \text{if } -\ell < x < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{11.9}
\]

The initial condition is discontinuous at two points: \( x = 0 \) and \( x = -\ell \).
The leading expansion fan

Let us begin our analysis of this problem by first taking \( \ell = \infty \). Then we have to interpret the behaviour of the discontinuity at the light, \( x = 0 \). We can write down our favourite implicit solution of (11.6) with the initial condition in (11.9) as

\[
\rho(x,t) = F[x - c(\rho)t].
\]  

We also draw the characteristics of the PDE in (11.6). These are straight lines in the \((x,t)\) plane on which \( \rho \) is constant:

\[
\xi = x - c(\rho)t
\]  

At \( t = 0 \) \( x = \xi \) and \( c = \pm u_m \) — see figure 11.2. There is a wedge-shaped region, originating at \((x,t) = (0,0)\), and within the wedge there are no characteristics.

To understand how to fill in the wedge we soften the discontinuous initial condition by smoothing out the jump. For example, instead of a discontinuous \( F(x) \), let us use

\[
F(x,\epsilon) = \frac{1}{1 + e^{x/\epsilon}},
\]  

and then take the limit \( \epsilon \to 0 \). Now when we can plot the characteristics (see figure 11.2) we see that the wedge is filled by an expansion fan or rarefaction wave. Inside the fan

\[
c(\rho) = x/t + O(\epsilon),
\]  

and outside of the fan, \( c = \pm u_m \).

In other words, as \( \epsilon \to 0 \), the solution of

\[
\rho(x,t) = F[x - c(\rho)t,\epsilon].
\]  

is

\[
\rho = \rho_0 + \epsilon \rho_1 + \epsilon^2 \rho_2 + \cdots
\]  

where the leading order term is

\[
\rho_0(x,t) = \begin{cases} 
\rho_j, & x < -u_m t, \\
\frac{1}{2} \rho_j \left(1 - \frac{x}{tu_m}\right), & -u_m t < x < u_m t, \\
0, & x > u_m t.
\end{cases}
\]
Figure 11.2: The expansion fan. expFan.png
We are encouraged because $\rho_0$ is independent of the detailed structure of $F(x, \epsilon)$.

Notice that $\rho(0, t) = \rho_j/2$ so that the flux at the light is $f = f(\rho_j/2) = f_m$, where $f_m$ is the maximum flux. The result that $f(\rho(0, t)) = f_m$ is independent of the detailed form of $f(\rho)$: when the light turns green the piled-up traffic flushes through the intersection at the maximal rate, $f_m$. To prove this, denote the position at which $f = f_m$ by $x(t)$ and the corresponding density $\rho_m$. Then

$$f[\rho(x_m, t)] = f_m, \text{ and } \rho[x_m(t), t] = \rho_m. \quad (11.17)$$

Also $c(\rho_m) = 0$ because the flux is a maximum at $\rho_m$. Thus evaluating the conservation equation

$$\rho_t + c \rho_x = 0, \quad (11.18)$$
on the trajectory $(x_m(t), t)$ we get

$$\rho_t(x_m, t) = 0. \quad (11.19)$$

But differentiating the second relation in (11.17) we have

$$\rho_t(x_m, t) + \frac{dx_m}{dt} \rho_x(x_m, t) = 0. \quad (11.20)$$

From (11.19) and (11.20) we conclude that either

$$\frac{dx_m}{dt} = 0, \quad \text{or} \quad \rho_x(x_m, t) = 0. \quad (11.21)$$

Thus as long as there is a density gradient at the light ($\rho_x \neq 0$) the flux at the light is $f_m$.

**The trailing shock**

Now, back to (11.9) and suppose $\ell$ is finite. Then the final car remains stationary until the time

$$t_* = \ell / u_m \quad (11.22)$$

at which the expansion fan reaches the end of the line, $x = -\ell$. The arrival of the fan at $x = -\ell$ sets the hitherto stationary rear shock into
Lecture 11. The Lighthill-Whitham traffic model

Figure 11.3: The forward expansion fan and the trailing shock. The shock is stationary at \( x = -\ell \) till the fan arrives at \( t = t^* \).

motion. We denote the position of this trailing shock by \( x = s(t) \). Evidently

\[
-\ell = s(t^*).
\]  

(11.23)

To figure out the subsequent position of the shock we use

\[
\dot{s} = \frac{f(s^+, t) - f(s^-, t)}{\rho(s^+, t) - \rho(s^-, t)}.
\]  

(11.24)

\[
\dot{s} = u_m \left( 1 - \frac{\rho_+ + \rho_-}{\rho_f} \right)
\]  

In the fan

\[
\rho(x, t) = \frac{\rho_f}{2} \left( 1 - \frac{x}{u_m t} \right).
\]  

The integrating factor is \( 1/\sqrt{t} \).

Behind the shock \( \rho^- = f^- = 0 \) and so (11.24) reduces to

\[
\dot{s} = u_m \left[ 1 - \frac{\rho(s^+, t)}{\rho_f} \right].
\]  

(11.25)

As one intuitively anticipates, the shock moves with the same speed as the final car.

Ahead of the shock we can use the expansion fan solution in (11.16) to get \( \rho(s^+, t) \). Thus (11.25) becomes an ODE for the shock position

\[
\dot{s} = \frac{u_m}{2} \left( 1 + \frac{s}{tu_m} \right), \quad \Rightarrow \quad \frac{ds}{dt} = \frac{1}{2} \frac{u_m}{\sqrt{t}}.
\]  

(11.26)

Integrating (11.26) from \( t^* \) to \( t \) we get the shock position as

\[
s(t) = \ell \left( \frac{t}{t^*} - 2\sqrt{\frac{t}{t^*}} \right), \quad \text{BYU.}
\]  

(11.27)

The shock, and the final car, passes the traffic light at \( t = 4t^* \).
11.3 The red light problem

Suppose we have a uniform stream of traffic with density $\rho_0$ and suddenly we turn on a red light at $x = 0$. We assume that the first car stops at the red light instantly and the following cars stop, building a jam with density $\rho_J$.

The problem is therefore

$$\rho_t + c(\rho)\rho_x = 0, \quad \rho(x,0) = \rho_0 < \rho_J, \quad f(0,t) = 0. \quad (11.28)$$

There are two shocks propagating away from $x = 0$. The left-moving shock has position

$$L(t) = -\frac{\rho_0}{\rho_J} u_m t. \quad (11.29)$$

This shock goes backwards into the incident traffic stream, leaving a jam in its wake.

The right-moving shock is at

$$R(t) = u_m \left( 1 - \frac{\rho_0}{\rho_J} \right) t. \quad (11.30)$$

The last guy is congratulating himself on making it through, and the speed of his car, $u(\rho_0)$, is also the speed on the shock.

We use the “standard model”:

$$f = u_m \rho \left( 1 - \frac{\rho}{\rho_J} \right)$$

throughout this lecture. The shock speed is then

$$\dot{s} = u_m \left( 1 - \frac{\rho_+ + \rho_-}{\rho_J} \right).$$

SIO203C, W.R. Young, March 21, 2011
11.4 Red light and green light

Now suppose we have the situation in figure 11.4 and the light turns back to green at \( t = \tau \). What happens?

The key is to realize that there is an expansion fan solution centered on the point \((x, t) = (0, \tau)\). This fan solution is

\[
\rho_F(x, t) = \frac{\rho_f}{2} \left( 1 - \frac{x}{(t - \tau) u_m} \right), \quad \text{if} \quad |x| < u_m(t - \tau).
\]  

The fan first interacts with the left shocks at point A in figure 11.5:

\[
(s_L, \tau_L) = \left( \frac{-\rho_0 u_m \tau}{\rho_j - \rho_0}, \frac{\rho_f \tau}{\rho_j - \rho_0} \right), \quad \tag{11.32}
\]

And the fan reaches the right shock at point B in figure 11.5:

\[
(s_R, \tau_R) = \left( \frac{u_0 \rho_f \tau}{\rho_0}, \frac{\rho_f \tau}{\rho_0} \right). \quad \tag{11.33}
\]

You can obtain these results clearly by drawing the characteristic diagram and figuring out where the edges of the expansion fan at \( x = \pm u_m t \) first hit the shocks in (11.29) and (11.30).

Once the fan starts to deflect the shocks the shock speed is given by

\[
\dot{s} = \frac{u_F \rho_F - u_0 \rho_0}{\rho_F(s, t) - \rho_0} = u_m \left( 1 - \frac{\rho_0 + \rho_F}{\rho_j} \right). \quad \tag{11.34}
\]

Thus to determine the position of both the left and the right shock we must integrate the ODE:

\[
\dot{s} - \frac{1}{2} \frac{s}{t - \tau} = u_m \left( \frac{1}{2} - \frac{\rho_0}{\rho_j} \right). \quad \tag{11.35}
\]

The solution of (11.35) is

\[
s(t) = u_m \left( 1 - 2 \frac{\rho_0}{\rho_j} \right) (t - \tau) + B \sqrt{t - \tau}, \quad \tag{11.36}
\]

where \( B \) is the constant of integration. We can determine \( B \) using the initial conditions in (11.32) and (11.33).
Figure 11.5: Trajectories of the shocks in the red light and green light problem. Before the expansion fan arrives the shocks move with constant speed (these are the heavy line segments). After the arrival of the expansion fan (at points $A$ and $B$) the shock paths curve. The expansion fan produced by the Green light is centered on the point $(x, t) = (0, \tau)$. In the solution above $\rho_0 = 3\rho_1/8$ and the shock finally makes it through the light at $t = 16\tau$. redgreen.eps
For the left hand shock we find after some brutal algebra that

\[ L(t) = u_m \left(1 - 2 \frac{\rho_0}{\rho_j}\right)(t - \tau) - 2u_m \sqrt{\left(1 - \frac{\rho_0}{\rho_j}\right) \frac{\rho_0}{\rho_j} \tau(t - \tau)} , \quad \text{if } t > \tau_L \]  

(11.37)

For the right hand shock we find

\[ R(t) = u_m \left(1 - 2 \frac{\rho_0}{\rho_j}\right)(t - \tau) + 2u_m \sqrt{\left(1 - \frac{\rho_0}{\rho_j}\right) \frac{\rho_0}{\rho_j} \tau(t - \tau)} , \quad \text{if } t > \tau_R \]  

(11.38)

If the traffic is light (\(\rho_0 < \rho_j/2\)) then the left shock reverses direction and eventually goes through the intersection (as shown in figure 11.5). This happens at

\[ \tau_c = \frac{\tau}{(1 - 2\rho_0/\rho_j)^2} \]  

(11.39)

After the clearance time, \(\tau_c\), the congestion behind the light has dissipated. An observer sitting at the light notices that at \(t = \tau_c\) the density drops from \(\rho_j/2\) back to \(\rho_0\). The clearance time can be distressingly long e.g., in figure 11.5 \(\tau_c = 16\tau\) so a stoppage lasting five minutes produces a jam requiring an hour and a quarter to clear.

11.5 References

My discussion of traffic closely follows


11.6 Problems

Problem 11.1. The expansion fan (11.16) is a similarity solution in which \(x\) and \(t\) appear only in the combination \(\eta = x/u_mt\). Substitute \(\rho(x, t) = R(\eta)\) into

\[ \rho_t + c(\rho)\rho_x = 0 \]

and show that you obtain (11.13).
Problem 11.2. Calculate the function \( \rho_1 \) in the expansion (11.15). Obtain an explicit expression for \( \rho_1 \), assuming that \( F \) is given by (11.12). Is \( \varepsilon \rho_1 \ll \rho_0 \) for all \( x \) as \( \varepsilon \to 0 \) i.e., is the expansion uniformly valid?

Problem 11.3. Solve the expansion-fan problem

\[
\begin{align*}
  u_t - u^2 u_x &= 0, \\
  u(x,0) &= H(-x).
\end{align*}
\]

Sketch both the characteristic diagram and \( u(x,1) \).

Problem 11.4. Solve the characteristic diagram and \( u(x,1) \).

Problem 11.5. Paint flowing down a wall has a thickness \( \eta(x,t) \) governed by

\[
\eta_t + \eta^2 \eta_x = 0.
\]

A stripe of paint is applied at \( t = 0 \) with

\[
\eta(x,0) = \begin{cases} 
  1, & 0 < x < 1, \\
  0, & \text{otherwise}.
\end{cases}
\]

Find \( \eta(x,t) \).

Problem 11.6. For the green light solution shown in figure 11.3 draw \( \rho(x,t) \) as a function of \( x \) at \( t = 0, 1/2, 1, 2, 4 \) and 5. Compute the strength of the trailing shock as a function of time.

Problem 11.7. Solve the traffic problem

\[
\rho_t + c(\rho)\rho_x = 0, \quad \rho(x,0) = \rho_J H(x).
\]

Draw the characteristic diagram.

Problem 11.8. Consider a more general traffic flow model with the flux function

\[
f(\rho) = u_m \rho \left[ 1 - (\rho/\rho_J)^n \right],
\]

where \( n > 0 \) is a parameter. The special case \( n = 1 \) is the standard model. Solve the green light problem, starting at (11.9), using this more general model. (You can check your answer by taking \( n = 1 \) and recovering the results from the lecture.) Suppose the last car (ar \( x = -\ell \)) starts moving at \( t_* \) and passes through the light \( (x = 0) \) at \( t = t_{**} \). Find the ratio \( t_{**}/t_* \) as a function of \( n \) and discuss the limits \( n \to 0 \) and \( n \to \infty \).
Problem 11.9. Suppose that
\[ f(\rho) = 120\sqrt{\rho}(1 - \sqrt{\rho}), \]
where \(\rho\) is cars per mile. At \(t = 0\) (Noon) the caravan is 2 miles long. The airport is 15 miles away and the velocity of the lead car is
\[ U = 120t, \text{ } t \text{ measured in hours past noon.} \]
The position of the lead car is therefore \(X(t) = 60t^2\) and so it reaches the airport at half past noon. The plane takes off at 12:35PM. Does the last car in the caravan arrive in time for the occupants to board the plane?

Problem 11.10. Draw \(\rho(x, t)\) as a function of \(x\) for for the solution in figure [11.5] (Note that in this figure \(\rho_0 = 3\rho_J/8\).) Choose values of \(t/\tau\) to illustrate how the structure of the solution changes with time.

Problem 11.11. Repeat problem 1 assuming that \(\rho_0 = 5\rho_J/8\) (heavy traffic). The analog of figure [11.5] is figure [11.6]

Problem 11.12. Let us develop an alternative theory for the speed of an \(\rho_J\)-plug behind a red light. Suppose the cars are all initially moving at \(u(\rho_0)\) and consider car \(n + 1\) behind the light. The initial distance between this car and the light is then \(n/\rho_0\). (i) How far will this car be from the light when it is forced to stop? Call this stopping distance \(X_n\). (ii) Assuming instant deceleration from \(u_0\) to rest, how long it take car \(n + 1\) to stop? Call this stopping time \(T_n\). (iii) Eliminate \(n\) between \(X_n\) and \(T_n\) and show that the line of stopped vehicles moves backwards with the speed \(U_L\) in (11.29).

Problem 11.13. Give a simpler derivation of the clearance time in (11.39). First show that
\[ \text{Number of cars through the light between } \tau \text{ and } \tau_c = f_m(\tau_c - \tau). \]
Also show that
\[ \text{Number of cars slowed by the light} = f(\rho_0)\tau_c. \]
Equating these two numbers show that
\[ \tau_c = \frac{f_m}{f_m - f_0}. \quad (11.40) \]
Show that for the standard model (11.40) reduces to (11.39).
Figure 11.6: Shock trajectories if $\rho_0 = 5\rho_f/8$ (heavy traffic). rgprob.eps