The Method of Characteristics

Richard Chapling v1 30 November 2018

In IA Differential Equations, you learnt to solve linear partial differential equations of the form

$$au_x + bu_y + cu = d \tag{1}$$

by setting $\xi = ax + by$ and $\eta = bx - ay$, which reduces the equation to $(a^2+b^2)u_\xi + cu = d$, with solution $u(\xi) = \frac{d}{c} + e^{-c\xi/(a^2+b^2)}A(\eta)$. This has a couple of notable features:

- Along the lines of constant η, the equation becomes essentially an ODE for ξ.
- The parametrisation (ξ, η) covers the plane, and each point has a unique coordinate.
- The function $A(\eta)$ is determined by the initial conditions.

The idea of the *Method of Characteristics* is to generalise these ideas to create a systematic way of producing solutions to certain classes of partial differential equations.

1 Characteristics for First-Order Equations

1.1 A Quick Note on Types of PDE

There are many different types of PDE,[1] so let us pause briefly to get the lie of the land. In the following, x stands for a vector of independent variables, u a dependent variable (which could also be a vector, but let's not worry about that now), and Du the collection of derivatives of u with respect to the xs.

- A PDE of order n is an equation of the form $F(x, u, Du, D^2u, \dots, D^n u) = 0$, where $D^k u$ denotes the collection of partial derivatives of u of order k. For the time being we shall only consider equations of first order.
- The equation mentioned in the previous section is *linear* with constant coefficients. A general linear first-order PDE is of the form

$$a(x) \cdot Du + b(x)u = c(x) \tag{2}$$

- Next we have *semilinear*, where the equation is still linear in Du, but the terms involving u can be nonlinear,

$$a(x) \cdot Du = c(x, u) \tag{3}$$

- Up again in generality, in *quasilinear* PDEs we allow the coefficients of Du to depend on u.

$$a(x, u) \cdot Du = c(x, u) \tag{4}$$

- Finally, a fully *nonlinear* first-order PDE is any expression

$$F(x, u, Du) = 0. \tag{5}$$

  We shall not consider the latter here, but suffice to say that they require additional equations to describe the evolution of the derivatives and their initial values, as well as the possibility that the solution will not be unique.

In this course we only consider linear PDEs in two dimensions, i.e. for first-order, those of the form

$$a(x, y)u_x + b(x, y)u_y + c(x, y)u + d(x, y) = 0. \tag{6}$$

In II Waves, characteristics are also applied to quasilinear equations, so we discuss the generalisation briefly at the end of the section.

1.2 First-Order Linear Equations

In the first section, we introduced a set of distinguished coordinates, so that the differential operator becomes differentiation with respect to a single variable. We want to extend this proceedure so that we can solve more complicated PDEs in a similar way. To do this, we construct curves that in a particular sense lie parallel to the action of the differential operator, which causes the PDE to become an ODE. We can then derive a general solution, specify constraints on the type of initial conditions we can impose, and often write down quite straightforwardly the solution with particular initial data.

We consider the differential equation (6). Let $(x(s), y(s))$ parametrise a curve in the plane. We would like the directional derivative of u on this curve to agree with the differential operator $a(x, y)u_x + b(x, y)u_y$ at a point. By the chain rule,

$$\frac{du}{ds} = \frac{dx}{ds}\frac{\partial u}{\partial x} + \frac{dy}{ds}\frac{\partial u}{\partial y}, \tag{7}$$

so for this to happen for any values of u_x and u_y, we need to satisfy the two ordinary differential equations

$$\frac{dx}{ds} = a(x(s), y(s)), \qquad \frac{dy}{ds} = b(x(s), y(s)). \tag{8}$$

These are called the *characteristic equations*; we can think of them as saying that the characteristics are *integral curves* of the vector field specified by $(a(x, y), b(x, y))$.[2] If these equations are satisfied by a curve, *on that curve* we can replace the operator $a(x, y)u_x + b(x, y)u_y$ by du/ds, which means that the original partial differential equation (6) becomes

$$\frac{du}{ds}(x(s), y(s)) + c(x(s), y(s))u(x(s), y(s)) + d(x(s), y(s)) = 0. \tag{9}$$

We have reduced the PDE to a set of three ODEs, which we expect to be both easier to solve and to have some sort of uniqueness in their solution. Because we obtain an entire family of characteristics, it is usual to give the arbitrary constants that emerge from solution of the ODEs a labelling that allows us to distinguish the characteristics; it should be emphasised that for first-order equations, this label has no significance apart from as a method to distinguish the characteristics, and hence we can choose it as we like to make things convenient (although we usually want it to be continuous and differentiable to make life simple).

How can we specify initial data for this equation? On each characteristic, the data changes based purely on the differential equation, so to guarantee consistency, we should only give the value of the function at one point on the characteristic. Normally, we specify the values on a curve, so the preceeding discussion implies

[1] There are some who think that the field of study of partial differential equations is called *'PDE', rather than 'PDEs'. The author is baffled by this, since there is clearly more than one of the things. (Is there a correlation between these people and those who think that the abbreviation of 'mathematics' is *'math'? Or people who find mass nouns confusing?)

[2] That is, curves that are parallel to the field at each point: if the vector field represents a steady flow of water, it is the curve that a piece of dust dropped into the stream would follow.

that the initial data curve must intersect each characteristic at most once. It also turns out that it is best to ask that the tangent to the initial data curve is never parallel to the vector field that gives the characteristics, for reasons of uniqueness involving the Inverse Function Theorem. A curve that satisfies both of these conditions is called *noncharacteristic* or *transverse*.

Essentially, the interpretation is that the initial data is propagated away from the initial curve by the characteristics.[3] This is consistent with the behaviour we have found, namely that the value at a point is determined only by the behaviour of the function on the characteristic through that point, starting with the initial data and evolving along it. Normally, we choose the constants of integration so that when $s = 0$, the point $(x(0), y(0))$ lies on the initial curve.

Trivial Example Consider the equation

$$u_t + cu_x = 0, \tag{10}$$

with constant coefficients.[4] We know how to solve this the old-fashioned way, but let's apply the Method of Characteristics to make sure we can use it to get an answer that agrees with the old method. The characteristic equations are

$$\frac{dx}{ds} = 1, \qquad \frac{dy}{ds} = c, \tag{11}$$

which have solutions $t(s) = s + A(r)$ and $x(s) = cs + B(r)$. If we label the lines by their intersection with $t = 0$, so that $t(0) = 0$ and $x(0) = r$, we find

$$t(s) = s, \qquad x(s) = cs + r. \tag{12}$$

We can also express s and r in terms of t and x:

$$s = t, \qquad r = x - ct. \tag{13}$$

The characteristics are thus the lines $x - ct = \text{const.}$. The PDE becomes

$$\frac{du}{ds} = 0, \tag{14}$$

and any solution to this equation is given by $u = C(r)$. Therefore

$$u(t, x) = C(x - ct) \tag{15}$$

for some function C, exactly the same as the IA result.

Example Consider the equation

$$xu_x + yu_y = 0. \tag{16}$$

We can write down the equations of the characteristics,

$$\frac{dx}{ds} = x \qquad \implies x = A(r)e^s \tag{17}$$

$$\frac{dy}{ds} = y \qquad \implies y = B(r)e^s \tag{18}$$

These are straight lines with gradient $B(r)/A(r)$. Hence we can only specify initial data on a curve that never points radially. The origin is not included (if it were it would lie on every characteristic, which would be rather unhelpful).

Example Consider the equation

$$yu_x - xu_y = u. \tag{19}$$

We can write down the equations of the characteristics,

$$\frac{dx}{ds} = y \quad \frac{dy}{ds} = -x \tag{20}$$

These must be solved simultaneously: we find that $x = A(r)\sin(s - B(r))$ and $y = A(r)\cos(s - B(r))$. The characteristics are circles centred at $x = y = 0$. The PDE becomes $du/ds = u$ on the characteristics, which has solution $u(x(s), y(s)) = C(r)e^s$ Notice that even if we specify initial data on a valid curve that is never parallel to the characteristics, u will not have the same value when we return to an initial point unless $u = 0$.

Full example Find the solution to

$$\begin{cases} u_x + xu_y + 2xu = e^{-x^2} \\ u(0, y) = ye^{-y^2}. \end{cases} \tag{21}$$

The characteristic equations are

$$\frac{dx}{ds} = 1 \qquad \frac{dy}{ds} = x, \tag{22}$$

which have solutions

$$x(s) = s + A(r), \qquad y(s) = \tfrac{1}{2}s^2 + A(r)s + B(r). \tag{23}$$

Since the initial set is $x = 0$, we choose A and B so that $x(0) = 0$ and $y(0) = r$, which gives

$$x(s) = s, \qquad y(s) = \tfrac{1}{2}s^2 + r. \tag{24}$$

We can invert this to obtain s and r in terms of x and y, which we will require later:

$$s = x \qquad r = y - \tfrac{1}{2}x^2. \tag{25}$$

Along a characteristic, the PDE becomes

$$\frac{du}{ds} + 2su = e^{-s^2}. \tag{26}$$

This is easy to integrate using the integrating factor e^{s^2}, and we find that

$$u = (s + C(r))e^{-s^2}. \tag{27}$$

Putting $s = 0$ gives $C(r) = u(x(0), y(0)) = u(0, r) = re^{-r^2}$, so

$$u = (s + re^{-r^2})e^{-s^2}, \tag{28}$$

and replacing s and r by their expressions in terms of x and y finally gives

$$u(x, y) = \left(x + (y - \tfrac{1}{2}x^2)e^{-(y-x^2/2)^2}\right)e^{-x^2}. \tag{29}$$

[3] The implicit use of temporal terminology is noteworthy here: in the second-order case we will see that having characteristics is quite closely linked to having a variable that behaves like time.

[4] This is often called the *transport equation*.

Full example Find the solution to

$$\begin{cases}(x-y)u_x + (x+y)u_y = \alpha u \\ u(x,0) = x^2, \qquad\qquad x \geqslant 0\end{cases} \tag{30}$$

We first write down the characteristic equations,

$$\frac{dx}{ds} = x - y, \qquad \frac{dy}{ds} = x + y, \tag{31}$$

which have solution

$$\begin{aligned} x(s) &= e^s(A(r)\cos s + B(r)\sin s) \\ y(s) &= e^s(-B(r)\cos s + A(r)\sin s). \end{aligned} \tag{32}$$

Since the initial data is given on the line $y = 0$, it is sensible to choose A and B so that $x(0) = r$ and $y(0) = 0$. This implies that $A(r) = r$ and $B(r) = 0$, which gives the characteristic parametrisation as

$$x(s) = re^s \cos s, \qquad y(s) = re^s \sin s. \tag{33}$$

These equations are not invertible in closed form, but we do have

$$\tan s = x/y, \qquad r^2 e^{2s} = x^2 + y^2 \tag{34}$$

which we will need later.

Along a characteristic, the PDE becomes

$$\frac{du}{ds} = \alpha u. \tag{35}$$

This is easy to solve, and gives

$$u = C(r)e^{\alpha s}. \tag{36}$$

Putting $s = 0$ implies that $C(r) = u(x(0), y(0)) = u(r, 0) = r^2$, so the solution is

$$u = r^2 e^{\alpha s}. \tag{37}$$

We see that data was only specified on the positive axis because any characteristic intersecting the positive axis goes on to intersect the negative one. Moreover, the solution can be written as $(x^2 + y^2)e^{(\alpha-2)\arctan(y/x)}$, and since the arctangent is multivalued, the solution is multivalued unless $\alpha = 2$.

1.3 More General First-Order Equations

1.3.1 Semilinear PDEs

If the PDE is only semilinear, we can calculate the characteristics in the same way, but the difference comes when we try to solve the ODE for u, which is no longer linear. As with semilinear ODEs, it is now possible that the solution may blow up in finite time. A simple example is

$$a(x,y)u_x + b(x,y)u_y = 1 + u^2, \tag{38}$$

which becomes

$$\frac{du}{ds} = 1 + u^2 \tag{39}$$

on a characteristic, and hence has solutions of the form $u = \arctan(s + C(r))$.

This may mean that u can only be expressed implicitly, but the characteristics themselves are unaffected.

1.3.2 Quasilinear PDEs

Now the characteristics can also depend on u itself. This can cause characteristics to cross, in which case the solution breaks down and we have shocks or wave breaking, depending on the system being modelled.[5]

Example The inviscid Burgers' equation is the simplest equation of this sort:

$$\begin{cases} u_t + uu_x = 0 \\ u(0,x) = \phi(x). \end{cases} \tag{40}$$

The characteristic equations are clearly

$$\frac{dt}{ds} = 1, \qquad \frac{dx}{ds} = u(t(s), x(s)). \tag{41}$$

We see that while we can solve and find $t(s) = s$, we need to know about u to find $x(s)$, although we can still impose $x(0) = r$. Fortunately, along a characteristic the PDE becomes

$$\frac{du}{ds} = 0, \tag{42}$$

so we find that $u(t(s), x(s)) = u(0, r) = \phi(r)$, and hence we can integrate the x equation,

$$x(s) = s\phi(r) + r. \tag{43}$$

These characteristics can intersect: if there are $r_1 > r_1$ so that $t(\phi(r_1) - \phi(r_2))/(r_1 - r_2) = -1$. If $\phi(r)$ is continously differentiable, the mean value theorem implies that this occurs if and only if there is an r_3 between r_2 and r_1 so that $\phi'(r_3) = -1/t$. We conclude that

- If $\phi'(r)$ is ever negative, the characteristics will cross and a shock will form
- This occurs at time t^*, where $1/t^* = -\min_r \phi'(r)$.

What can we say about the actual function? We know that $u(t(s), x(s)) = \phi(r)$, which gives the implicit equation

$$\phi(r) = u(t, t\phi(r) + r) \tag{44}$$

for u. If we instead consider u as an independent variable, we can write $r = x - ut$ on a characteristic, so we find that an alternative expression is

$$u = \phi(x - ut), \tag{45}$$

which is normally easy to solve by iteration.

1.4 Summary of (Semi)Linear First-Orders

Given a differential equation

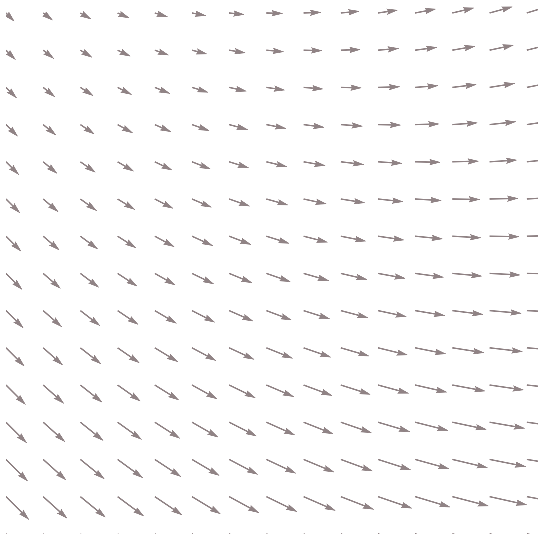
$$au_x + bu_y + f(u) = 0, \tag{46}$$

where a, b are functions of x, y,

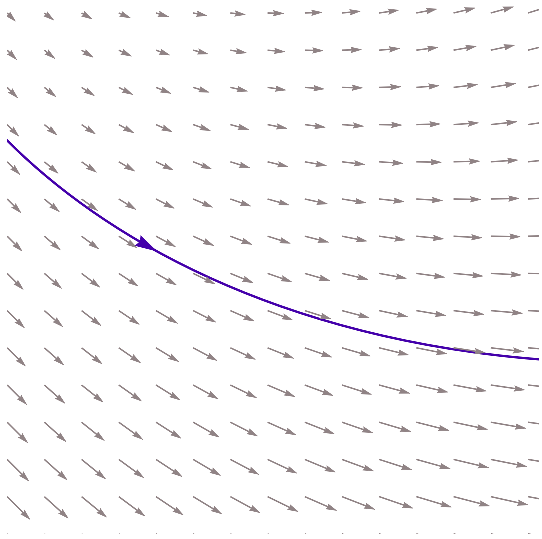
1. Associated to the differential operator $a\partial_x + b\partial_y$ in the equation is a vector field (a, b). (Figure 1a)
2. The integral curves of this vector field (Figure 1b) are called *characteristics*. The differential operator becomes a single derivative on these curves by the chain rule, so the PDE becomes an ODE. The equations
$$\frac{dx}{ds} = a, \qquad \frac{dy}{ds} = b, \qquad \frac{du}{ds} + cu + d = 0 \tag{47}$$
are called the *characteristic equations*.
3. We create a parametrisation of the space of independent variables by labelling the characteristics continuously, providing a transformation $(r, s) \mapsto (x, y)$ (Figure 1c).
4. Solving the characteristic equations implies that the value at a point depends only on the values along the characteristic before that point, via a first-order ODE. Therefore, initial data should only specify one value on each curve, so the initial data curve should only intersect each curve once, and to use the Inverse Function Theorem to obtain a unique solution, we also require that the initial data curve never has the same tangent direction as the vector field (Figure 1d).

[5] While such effects are defects of the model, real systems can exhibit these phenomena, as you know if you've ever heard a sonic boom (sadly rather less likely now that Concorde is no longer in service).

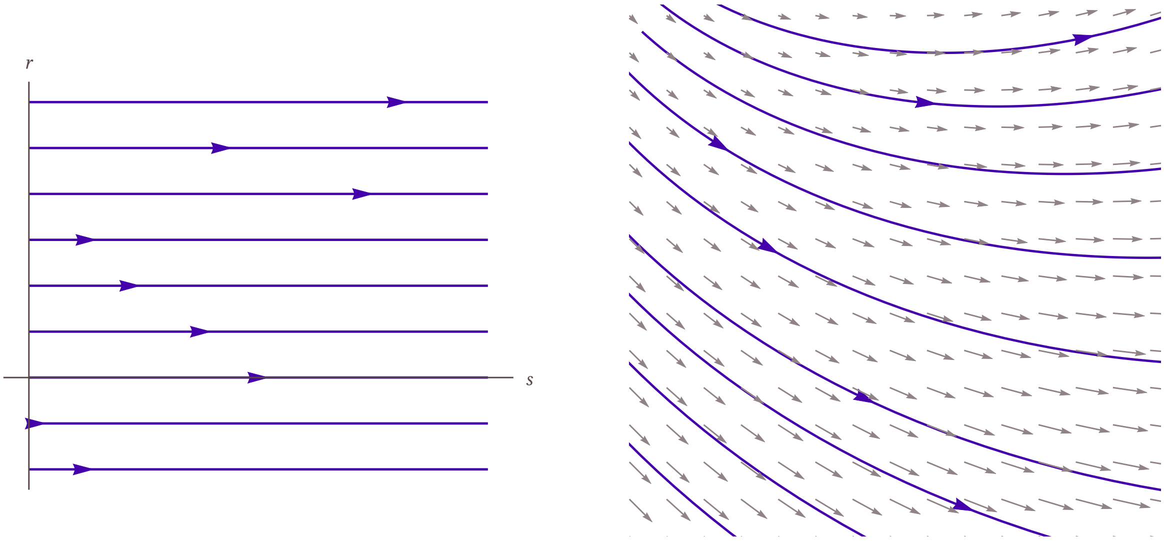
Figure 1: Construction of characteristics



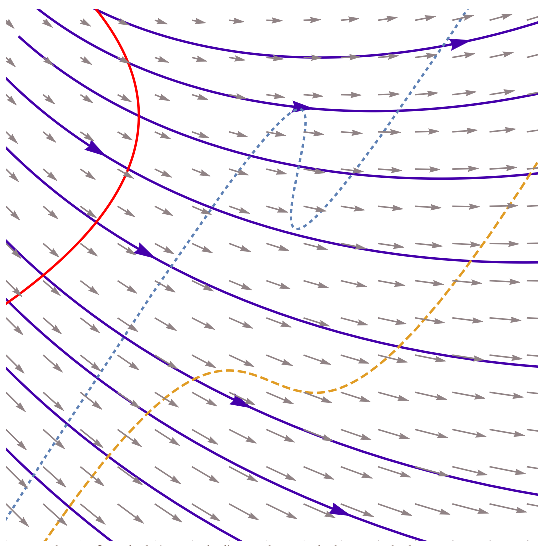
(a) Vector field defined by differential operator



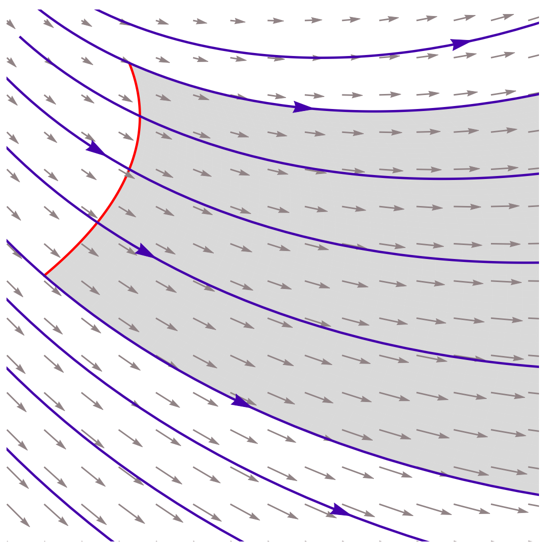
(b) An integral curve of the vector field



(c) Parametrisation of space of independent variables by characteristics



(d) Examples of valid (■, solid) and invalid initial data curves: one (■, dotted) crosses some characteristics more than once, while the other (■, dashed) is tangent to the vector field at a point.



(e) Region where solution is determined by initial data (shaded)

5. All of this means that we know the value of u at a point if there is a single characteristic from somewhere the initial curve to that point. If there are no characteristics, we know nothing, and if there is more than one, we need more information (Figure 1e).

2 Characteristics for Second-Order Linear Equations

The idea here is quite different from first-orders: we can have different numbers of characteristics depending on the nature of the differential equation.

2.1 Classification of Second-Order Linear Equations

In two dimensions, we have considered three PDEs: Laplace's equation, the heat/diffusion equation, and the wave equation, each of which have very different types of solution. It turns out that *all* second-order linear PDEs in two dimensions fall into one of these categories.

As you know, solutions to the three archetypes act quite differently: in the wave equation, singularities propagate, whereas in the diffusion equation, everything is rapidly smoothed out and decays, for example. Solutions to Laplace's equation and the heat equation have no internal maxima, whereas the wave equation can.

A general second-order linear equation in two variables is

$$\begin{aligned} &a(x,y)u_{xx} + 2b(x,y)u_{xy} + c(x,y)u_{yy} \\ &+ d(x,y)u_x + e(x,y)u_y + f(x,y)u + g(x,y) = 0. \end{aligned} \tag{48}$$

Once again, let us consider the case of constant coefficients for inspiration. As before, we look at how the form of the equation can change when we change variables. The simplest change of variables we know is linear, so let

$$\xi = \alpha x + \beta y, \qquad \eta = \gamma x + \delta y, \tag{49}$$

and write $w(\xi,\eta) = u(x(\xi,\eta), y(\xi,\eta))$. Then the derivatives become

$$u_x = \frac{\partial \xi}{\partial x} w_\xi + \frac{\partial \eta}{\partial x} w_\eta = \alpha w_\xi + \gamma w_\eta \tag{50}$$

$$u_y = \frac{\partial \xi}{\partial y} w_\xi + \frac{\partial \eta}{\partial y} w_\eta = \beta w_\xi + \delta w_\eta \tag{51}$$

$$u_{xx} = \alpha^2 w_{\xi\xi} + 2\alpha\gamma w_{\xi\eta} + \gamma^2 w_{\eta\eta} \tag{52}$$

$$u_{xy} = \alpha\beta w_{\xi\xi} + (\alpha\delta + \beta\gamma) w_{\xi\eta} + \gamma\delta w_{\eta\eta} \tag{53}$$

$$u_{yy} = \beta^2 w_{\xi\xi} + 2\beta\delta w_{\xi\eta} + \delta^2 w_{\eta\eta} \tag{54}$$

And then the differential equation becomes

$$\begin{aligned} 0 =& (a\alpha^2 + 4b\alpha\beta + c\beta^2) w_{\xi\xi} \\ &+ 2(2a\alpha\gamma + b(\alpha\beta + \gamma\delta) + 2\beta\delta) w_{\xi\eta} \\ &+ (a\gamma^2 + 4b\gamma\delta + c\delta^2) w_{\eta\eta} \\ &+ (d\alpha + e\beta) w_\xi + (d\gamma + e\delta) w_\eta + fw + g \end{aligned} \tag{55}$$

You may recognise the relationships in the coefficients of the second derivatives: this is what happens when we change variables in a quadratic form, namely

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \mapsto \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^T ! \tag{56}$$

So in the constant coefficients case, we can use the theory of quadratic forms to inform us about how such transformations affect the second derivative terms. In particular, we know that the eigenvalues of a quadratic form are real, and Sylvester's law of inertia implies that the signature of the quadratic form (i.e. how many positive, negative and zero eigenvalues it has) is invariant under any transformation we make. Since quadratic forms can be diagonalised, there is a change of variables so that it can be written in one the forms

$$u_{\xi\xi} + u_{\eta\eta}, \qquad u_{\xi\xi} - u_{\eta\eta}, \qquad u_{\xi\xi}, \tag{57}$$

depending on the nature of the eigenvalues: these are refered to as elliptic, hyperbolic and parabolic cases, after the quadrics that their respective quadratic forms.[6] However, as you noted last year, the middle form is not quite so useful as the equivalent form $u_{\xi\eta}$, which is easier to integrate, so we shall adopt this as the hyperbolic canonical form.

Does this still apply when the coefficients are not constant? Yes, but it is slightly messier. Suppose now that a, b, c are all now functions of x and y, and that we take the general nondegenerate substitution $(\xi(x,y), \eta(x,y))$ (i.e. so that its Jacobian matrix

$$J = \begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} \tag{58}$$

is nonsingular). By the Inverse Function Theorem, this has an inverse $(x(\xi,\eta), y(\xi,\eta))$, and we set $w(\xi,\eta) = u(x(\xi,\eta), y(\xi,\eta))$. We now find that when we calculate the derivatives of u, they acquire extra first-order terms from differentiating the substitution:

$$u_x = \xi_x w_\xi + \eta_x w_\eta \tag{59}$$

$$u_y = \xi_y w_\xi + \eta_y w_\eta \tag{60}$$

$$u_{xx} = \xi_x^2 w_{\xi\xi} + 2\xi_x\eta_x w_{\xi\eta} + \eta_x^2 w_{\eta\eta} + \xi_{xx} w_\xi + \eta_{xx} w_\eta \tag{61}$$

$$\begin{aligned} u_{xy} =& \xi_x\xi_y w_{\xi\xi} + (\xi_x\eta_y + \xi_y\eta_x) w_{\xi\eta} + \eta_x\eta_y w_{\eta\eta} \\ &+ \xi_{xy} w_\xi + \eta_{xy} w_\eta \end{aligned} \tag{62}$$

$$u_{yy} = \xi_y^2 w_{\xi\xi} + 2\xi_y\eta_y w_{\xi\eta} + \eta_y^2 w_{\eta\eta} + \xi_{yy} w_\xi + \eta_{yy} w_\eta. \tag{63}$$

Putting all these together gives the transformed equation in the form

$$0 = Aw_{\xi\xi} + 2Bw_{\xi\eta} + Cw_{\eta\eta} + Dw_\xi + Ew_\eta + Fw + G, \tag{64}$$

where the new coefficients are

$$A = a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2 \tag{65}$$

$$B = a\xi_x\eta_x + b(\xi_x\eta_y + \xi_y\eta_x) + c\xi_y\eta_y \tag{66}$$

$$C = a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2 \tag{67}$$

$$D = d + a\xi_{xx} + 2b\xi_{xy} + c\xi_{yy} \tag{68}$$

$$E = e + a\eta_{xx} + 2b\eta_{xy} + c\eta_{yy} \tag{69}$$

$$F = f, \qquad G = g. \tag{70}$$

Do not bother remembering these formulae, but remember how to derive them. We therefore see that once again the new quadratic form is the transformation of the old quadratic form using the Jacobian as the matrix, namely

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix} = J \begin{pmatrix} a & b \\ b & c \end{pmatrix} J^T, \tag{71}$$

and looking at the determinant, $AC - B^2 = (\det J)^2(ac - b^2)$, so the sign of the determinant is invariant, consistent with Sylvester's Law of Inertia. Therefore the sign of the quantity $\Delta = ac - b^2$ is the important thing to consider to classify second-order linear PDEs: we see that a PDEs is elliptic at a point if $\Delta > 0$, parabolic if $\Delta = 0$, and hyperbolic if $\Delta < 0$. The same equation may have different type at different points, as shown by the Tricomi equation $yu_{xx} + u_{yy} = 0$, which is hyperbolic for $y < 0$, parabolic for $y = 0$, and elliptic for $y > 0$.

Why have we said nothing about the first derivative terms? It turns out that the second derivative terms are much more significant in the behaviour of solutions to the equation (to justify this is beyond the scope of this course, let alone this handout).

We now reduce the three types of equation to its canonical form, and consider the features of each.

[6] In more dimensions, hyperbolic means having one negative eigenvalue. If the form has no zero eigenvalues and more than one negative and one positive eigenvalue, it is called *ultrahyperbolic*, which can only happen in 4 or more dimensions.

2.2 Hyperbolic

Suppose that $\Delta < 0$. Then we want to find ξ and η so that A and C vanish, i.e. so that

$$a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2 = a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2 = 0. \quad (72)$$

This will reduce the PDE (48) to its canonical form

$$w_{\xi\eta} + Dw_\xi + Ew_\eta + Fw = G \quad (73)$$

after dividing by $2B$ and relabelling.

The equations for ξ and η are identical quadratic equations, so in fact it suffices to consider both roots of one. We can assume that we have labelled our coordinates so that $a \neq 0$, and then ξ solves the equation if

$$\xi_x - \frac{-b + \sqrt{b^2 - ac}}{a}\xi_y = 0 \quad \text{or} \quad \xi_x - \frac{-b - \sqrt{b^2 - ac}}{a}\xi_y = 0. \quad (74)$$

Labelling the roots of $a\mu^2 + 2b\mu + c = 0$ as $\mu_\pm$, we can therefore dictate arbitrarily that our new coordinates satisfy

$$\xi_x - \mu_+\xi_y = 0 \quad \text{and} \quad \eta_x - \mu_-\eta_y = 0. \quad (75)$$

(Is this a legitimate change of coordinates? Yes: $\det J = (\mu_+ - \mu_-)\xi_y\eta_y = 2\sqrt{-\Delta}\xi_y\eta_y \neq 0$.)

At last we see why this is part of the method of characteristics: to find ξ and η, we can apply the Method for first-order equations. For example, for ξ we see that the equations are

$$\frac{dx}{ds} = 1 \qquad \frac{dy}{ds} = -\mu_+, \quad (76)$$

and on this line $d\xi/ds = 0$, so ξ is constant. In view of the chain rule, we see that we can write $dy/dx = -\mu_+$ to obtain the solution for y as a function of x. Writing the solution in the form $f(x, y) = C$, C the constant of integration, we may then simply set $\xi = f(x, y)$. Exactly the same may be done for η.

We end up with *two* sets of characteristics for a hyperbolic equation. This is closely connected with the wave equation taking *two* initial conditions, whereas the other equations only take one: two sets of characteristics gives us two archetypal solutions, which require two initial conditions to prescribe unambiguously.

2.3 Parabolic

The canonical form is

$$w_{\xi\xi} + Dw_\xi + Ew_\eta + Fw + G = 0. \quad (77)$$

For parabolic equations, we know that $ac - b^2 = 0$. A convenient way to write this is, again on the assumption that $a \neq 0$, is $c = b^2/a$. We decide to set $C = 0$, since it makes no difference: this means solving a single characteristic equation,

$$\eta_x - \eta\xi_y = 0, \quad (78)$$

where μ is a solution to $0 = a\mu^2 + 2b\mu + b^2/a = a(\mu + b/a)^2$, i.e. $\mu = -b/a$, so in terms of x and y, the equation becomes

$$\frac{dy}{dx} = -\mu = \frac{b}{a}. \quad (79)$$

If the solution to this is $f(x, y) = C$, we can set $\eta(x, y) = f(x, y)$. What about the other variable? Anything will work, providing that $\xi_x\eta_y - \xi_y\eta_x \neq 0$, i.e. the transformation is nondegenerate. This freedom is often quite useful, since it may enable us to simplify the equation even further.

Thus there is only one set of characteristics for a parabolic equation; this is essentially why only on initial condition can be specified at each initial point. Information propagates along characteristics, but may be influenced by neighbouring information via the $w_{\xi\xi}$ and w_ξ terms: this is what causes the diffusive effects in the diffusion equation, for example.

2.4 Elliptic

The canonical form is

$$w_{\xi\xi} + w_{\eta\eta} + (\text{lower order terms}) = 0, \quad (80)$$

so we want to have $A = C$ and $B = 0$, i.e.

$$a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2 = a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2 \quad (81)$$

$$a\xi_x\eta_x + b(\xi_x\eta_y + \xi_y\eta_x) + c\xi_y\eta_y = 0. \quad (82)$$

For elliptic equations, there are no real characteristics, so we have to do something else. Notice that the first equation can be rewritten as

$$a(\xi_x^2 - \eta_x^2) + 2b(\xi_x\xi_y - \eta_x\eta_y) + c(\xi_y^2 - \eta_y^2) = 0 : \quad (83)$$

this looks like the real part of a complex combination of ξ and η derivatives. Indeed, the other equation looks like an imaginary part, and we find that if we put $\phi = \xi + i\eta$, both equations can be represented as

$$a\phi_x^2 + 2b\phi_x\phi_y + c\phi_y^2 = 0. \quad (84)$$

We know that $a\mu^2 + 2b\mu + c = 0$ has two complex conjugate roots $\mu_\pm = -b/a \pm i\sqrt{\Delta}/a$ in this case, so we can consider the complex factorisation

$$a(\phi_x - \mu_+\phi_y)(\phi_x - \mu_-\phi_y) = 0. \quad (85)$$

We can then solve the complex characteristic equations

$$\frac{dy}{dx} = -\mu_\pm, \quad (86)$$

which gives solutions in the form $f_\pm(x, y)$, and then we can choose $\phi_\pm(x, y) = f_\pm(x, y) = C$. We can finally return to the real world by identifying

$$\xi = \tfrac{1}{2}(\phi_+ + \phi_-), \qquad \eta = \tfrac{1}{2i}(\phi_+ - \phi_-) \quad (87)$$

(essentially because the characteristic equations are complex conjugates, so their solutions can be chosen to be complex conjugates, to be consistent with our original definitions), whence we obtain the desired canonical form. While it would be possible to avoid complex functions in this derivation, we would gain little and make the calculations rather more complicated.

3 References

The author used the following resources in preparing this handout:

- Characteristics for first-order equations is nicely done, in a progression similar to the one employed above, in the *First-Order Equations* notes for Julie Levandosky's course, here: `https://web.stanford.edu/class/math220a/lecturenotes.html`. (The interpretation of the geometry involved is slightly different, however.[7])

- Classification of second-order equations and their reduction to standard forms is nicely covered in Kris Wysocki's notes, found here: `https://www.math.psu.edu/wysocki/M412/Math412.html` (Week 5 covers this, and has several examples).

[7] One *can* interpret the characteristics problem for a first-order ODE in terms of calculating a surface from its normals and a curve in it, but because this interpretation does not extend to more general problems, and provides no understanding of what is happening in the second-order case, we have not discussed it.