Mathematics 428, Spring 2007

Introduction to the Method of Characteristics

Suppose $\Omega$ is an open bounded subset of $\mathbb{R}^2$, and $g$ is a smooth function defined on $\partial \Omega$.

Suppose $a$ and $b$ are some real constants, and $u(x,y)$ is a continuously differentiable solution of a first-order linear PDE

$$au_x(x,y) + bu_y(x,y) = 0, \quad \text{for } (x,y) \in \Omega,$$

and $u(x,y) = g(x,y)$ on the boundary of $\Omega$.

Consider a trajectory in $\mathbb{R}^2$ parametrized by

$$(x(t), y(t)) = (ta, tb) + (x_0, y_0).$$

The claim is that $u(x(t), y(t))$ is constant as long as $(x(t), y(t))$ is inside $\Omega$.

Indeed, differentiating with respect to $t$, by the chain rule,

$$\frac{d}{dt} [u(x(t), y(t))] = u_x(x(t), y(t))a + u_y(x(t), y(t))b = 0,$$

where the last equality is true simply because $u$ is a solution of (1).

Thus, to find $u(x,y)$ for any $(x,y) \in \Omega$, you need to

1. take a straight line parallel to a vector $(a, b)$ and passing through $(x,y)$;
2. find its intersection with the boundary of $\Omega$ — e.g., at some point $(x_0, y_0) \in \partial \Omega$;
3. take $u(x,y) = g(x_0, y_0)$.

Of course, there are several caveats here to consider. First of all, if that straight line intersects $\partial \Omega$ at several points, then the function $g$ should better have the same value at all of those points. (Otherwise boundary conditions are said to be "inconsistent" or "incompatible".)

Secondly, you might think that it works only because the PDE (1) is so simple. If you slightly generalize it to

$$au_x(x,y) + bu_y(x,y) = f(x,y), \quad \text{for } (x,y) \in \Omega,$$

the same argument shows that, along the trajectory (2), we have

$$\frac{d}{dt} [u(x(t), y(t))] = u_x(x(t), y(t))a + u_y(x(t), y(t))b = f(x(t), y(t)).$$
So, returning to the 3 steps outlined above, the first two stay the same, but the third is now:

let \( t = (x - x_0)/a \) (or, equivalently, \( t = (y - y_0)/b \)) and set

\[
    u(x,y) = g(x_0, y_0) + \int_0^t [f(x_0 + as, y_0 + bs)] ds.
\] (4)

(We note that the compatibility of boundary conditions becomes a bit trickier too.)

The lines parallel to \((a, b)\) are called "characteristics" and people often say that "the information is traveling along the characteristics". In these examples it merely means that we could integrate something along the characteristic curve to obtain \( u(x,y) \) for every point on that curve.

We can push this logic one step further, by considering

\[
a(x,y)u_x(x,y) + b(x,y)u_y(x,y) = f(x,y), \quad \text{for} \ (x,y) \in \Omega, \tag{5}
\]
i.e., \( a \) & \( b \) are no longer constants, but some smooth functions of \( x \) and \( y \). A similar trick works, but the trajectory that we need to consider is more complicated. Suppose the trajectory \((x(t), y(t))\) is such that

\[
x'(t) = a(x(t), y(t)), \\
y'(t) = b(x(t), y(t)),
\] (6)

for all values of \( t \). (these are called "the characteristic ODEs")

Then, by the chain rule,

\[
\frac{d}{dt}[u(x(t), y(t))] = u_x(x(t), y(t))a(x(t), y(t)) + u_y(x(t), y(t))b(x(t), y(t)) = f(x(t), y(t)),
\]

where the last equality follows from the fact that \( u \) solves (5).

Thus, given a point \((x,y) \in \Omega\) we now need to solve the "back to the boundary" initial value problem

\[
v'(s) = -a(v(s), w(s)), \\
w'(s) = -b(v(s), w(s)), \\
v(0) = x, \\
w(0) = y
\] (7)

until we (hopefully) cross the boundary of \( \Omega \) at some point \((x_0, y_0)\) at some time \( t \).
This guarantees that if we solve the original "forward from the boundary" system

\[
\begin{align*}
x'(s) &= a(x(s), y(s)), \\
y'(s) &= b(x(s), y(s)), \\
x(0) &= x_0, \\
y(0) &= y_0
\end{align*}
\]  

we will have \(x(t) = x\) and \(y(t) = y\).

Once we have \((x_0, y_0)\) and \(t\), we can compute \(u(x, y)\) as before, using the generalization of formula (4):

\[
u(x, y) = g(x_0, y_0) + \int_0^t [f(x(s), y(s))] ds.
\]  

Here we did not have a nice closed form for the characteristic curves available to us a priory, but could find them by solving a system of ODEs. Moreover, if \(a(x, y)\) and \(b(x, y)\) were "nice" (i.e., Lipschitz-continuous and never became zero simultaneously inside \(\Omega\)), we knew that there is only one characteristic passing through every point – this is important, since a different value of \(u(x, y)\) could be obtained by integrating along each characteristic passing through \((x, y)\).

It is not hard to see that almost the same ideas work for an "almost linear first-order" PDE

\[
a(x, y)u_x(x, y) + b(x, y)u_y(x, y) = f(x, y, u(x, y)), \quad \text{for } (x, y) \in \Omega,
\]  

where the right hand side of the equation is allowed to depend on \(u\). Indeed,

\[
\frac{d}{dt}[u(x(t), y(t))] = u_x(x(t), y(t))a(x(t), y(t)) + u_y(x(t), y(t))b(x(t), y(t)) = f(x(t), y(t), u(x(t), y(t))),
\]

and the only part of the above that needs to be changed is the formula (9). Instead, we define \(z(t) = u(x(t), y(t))\) and compute it by solving an additional initial value problem:

\[
\begin{align*}
z'(s) &= f(x(s), y(s), z(s)), \\
z(0) &= g(x_0, y_0).
\end{align*}
\]  

That simple life is over once you allow \(a\) or \(b\) to depend upon \((u(x, y))\).

[The PDEs in that category include "hyperbolic conservation laws"; one typical example is the traffic-flow equation, where the average speed of cars depends on their current congestion level at that point of the road.]
For such general \textit{quasi-linear} first-order PDEs
\[ a(x,y,u(x,y))u_x(x,y) + b(x,y,u(x,y))u_y(x,y) = f(x,y,u(x,y)), \quad \text{for } (x,y) \in \Omega. \tag{12} \]
the same trick “sort of works” if you start at the point \((x_0, y_0)\) on the boundary, but you need to solve for \(u\) simultaneously (since the characteristic’s direction depends upon it):
\[
\begin{align*}
x'(s) &= a(x(s), y(s), z(s)), \\
y'(s) &= b(x(s), y(s), z(s)), \\
z'(s) &= f(x(s), y(s)),
\end{align*}
\tag{13}
\]
and the initial conditions are
\[ x(0) = x_0; \quad y(0) = y_0; \quad z(0) = q(x_0, y_0). \]
However, here two ugly complications make our life unpleasant:
\begin{enumerate}
\item characteristics can cross: i.e., starting from some other boundary point \((x_1, y_1)\), you might also reach \((x, y)\) after some time \(t_1\) and the corresponding \(z(t_1)\) will be quite different from the \(z(t)\) produced starting from \((x_0, y_0)\).
\item given \((x, y) \in \Omega\), we don’t know how to find any \((x_0, y_0) \in \partial \Omega\) that will eventually lead to it; the previous “backward in time” trick will not work here because we also need to have \(u(x, y)\) to figure out the correct \(a\) and \(b\).
\end{enumerate}
“Viscosity solution”-type tricks allow you to deal with the first difficulty (they simply identify which of the characteristics passing through \((x, y)\) should we trust).
The second difficulty has no clear-cut easy solution in general, but you can try to address it as a two-point boundary value problem.

With a bit more effort (and many additional caveats), the method of characteristics is even applicable for fully non-linear first-order PDEs such as the Eikonal equation shown in the first lecture.
\[\text{[Note: most likely we will not attempt that generalization in this course; please see Chapter 3 of Evans’ textbook if you are curious about the subject.]}\]

So why can’t we declare a victory? (at least in the linear and almost linear case?)
First of all, our recipes for computing \(u(x,y)\) are produced in the spirit of “if a smooth solution on \(\overline{\Omega}\) existed, then here is what it would have to be equal to at the point \((x,y)\)".
Can we compute such a value for every point in $\Omega$? [Note the word "hopefully" in the caveat after the formula (7).] Will thus defined function actually solve the PDE? This is simple to verify directly for the equation (1), since the analytic formula for $u$ is easy to derive (though still depends on the exact shape of $\Omega$). For all the other PDEs discussed above, if $a$, $b$, $f$, and $g$ are some generic smooth functions and $\partial \Omega$ is not particularly simple, we have no idea how to produce an analytic formula for the characteristic passing through $(x,y)$ nor for the corresponding point $(x_0,y_0) \in \partial \Omega$. Thus, in general, a direct verification is not an option, and we need to find a different way to prove that what we are computing is actually a solution.

Chapters 2 and 3 of Zachmanoglou & Thoe will help us to address these issues.