PDE in Finance, Spring 2008,
`http://www.math.nyu.edu/faculty/goodman/teaching/PDEfin/index.html`
Written by Jonathan Goodman. Please report errors, sources of confusion, and typos on the class Blackboard site.
Last updated February 10, 2008.

Section 3: More general equations, qualitative properties

1 Introduction

Purely analytic methods give the solution to only a few partial differential equations. For the majority, we use computers. The qualitative theory of partial differential equations is crucial for developing computational strategies and interpreting the results. It studies questions such as existence, uniqueness, and well posedness. More detailed qualitative questions include various stability properties and the relation between different solutions of the same PDE, *comparison principles*.

The issues in the qualitative theory of PDEs are more abstract than simply finding formulas for the solutions. The techniques also are abstract. They often seek to make statements about general solutions to a particular type of PDE without being to describe that solution very precisely. Most of any modern book on PDE (Evans, John, etc) is devoted to qualitative theory. The discussion of advection here is similar to the discussion in Evans.

2 Advection and characteristics

In a fluid flow, *Advection* means something being carried along passively by the fluid, such as a dye. This also is called *convection*. I don't know the difference between advection and convection, but people always say advection in the present context. Abstractly, suppose the fluid at point x has velocity $a(x)$. Let $x(t)$ be the location at time t of a particle that is advected with the fluid. It's velocity at time t is $a(x(t))$, so it is governed by the ordinary differential equation $\partial_t x(t) = \dot{x}(t) = a(x(t))$.

Now suppose the fluid is advecting a quantity, like sugar, but not diffusing it. Let $u(x,t)$ be the density (say, of sugar) at location x at time t. We use the conservation/flux ideas from Section 1 to derive an *advection equation*, a partial differential equation that u satisfies. We assume that the quantity being advected is neither created nor destroyed, and

$$H_V(t) \;=\; \int_V u(x,t)\,dx \;,$$

is the total amount inside the *control volume*, V at time t. As in Section 1,

$$\frac{dH_V}{dt} \;=\; \int_V \partial_t u(x,t)\,dx \;=\; -\int_\Gamma F(x,t)\cdot n(x)\,dA(x)\;. \tag{1}$$

You don't have to Fourier or Fick to figure out the *advective flux*, $F(x,t)$. The rate at which material is transported across a small piece of surface is proportional to the velocity component normal to the surface, the area of the small piece of surface, and the density of the material being advected:

$$F(x,t) \cdot n(x)\, dA(x) \;=\; u(x,t)\,(a(x) \cdot n(x))\; dA(x)\;.$$

Here, u is a scalar quantity, the local density, while $a(x)$ is a velocity vector. We see that the advective flux is $F(x,t) = u(x,t)a(x)$. As in Section 1, we can use the divergence theorem to get, for any control volume,

$$0 \;=\; \int_V \partial_t u(x,t)\, dx \;+\; \int_V \operatorname{div}\,(u(x,t)a(x))\; dx\;.$$

This implies that u satisfies the partial differential equation

$$\partial_t u \;+\; \operatorname{div}\,(a(x)u(x,t)) \;=\; 0\;. \tag{2}$$

This is the advection equation for a *passive scalar* moving in velocity field $a(x)$.

The same equation applies to advection of a random point. Suppose we choose $X(0)$ according to probability density $f(x)$ and then let $X(t)$ advect with the velocity field $a(x)$. Let $u(x,t)$ be the probability density for $X(t)$. The same reasoning shows that u satisfies the advection equation (2). The solution should have two properties consistent with its interpretation as a probability density. One is that the total mass does not change with time:

$$\partial_t \int_{R^n} u(x,t)\, dx \;=\; 0\;. \tag{3}$$

This is clear from (1) if $u(x,t) \to 0$ as $x \to \infty$ fast enough so that the integral on the right goes to zero. It applies to general solutions of the advection equation. The other property is that $u(x,t) \geq 0$ for all x and t. This is related to the maximum principle and we return to it later.

The initial value problem would be to find a function $u(x,t)$ that satisfies (2) for all x and $t > 0$ and in addition the initial condition $u(x,0) = f(x)$. This initial value problem is well posed. We will give some of the idea of the proof later. But the solutions have many properties that are qualitatively different from solutions of the heat equation. These differences arise from the fact that advection does not destroy information the way heat conduction or diffusion does. In fact, if we have $g(x) = u(x,T)$, for some $T > 0$, we can recover $f(x)$ by running the advection equation backwards. That is, let $v(x,t)$ satisfy the advection equation with velocity field $-a(x)$ (running time backwards) and $v(x,0) = g(x)$, then $v(x,T) = f(x)$. I suppose this is not a mathematical proof, but it should seem true. The mathematical proof is to let $v(x,t) = u(x,T-t)$ and calculate that v has the required properties.

The crux of this argument is that if the initial value problem for advection equations is well posed for general velocity fields, a, then it is well posed for $-a$ as well as a. This implies that the backward evolution equation also is well

posed. For the heat equation, the equation satisfied by $v(x,t) = u(x, T-t)$ is $\partial_t v = \frac{-1}{2} \triangle v$, which is ill posed.

Another non-heat property of advection is that it is not smoothing. Instead, it preserves sharp boundaries and singularities of other kinds. You can get an intuitive picture of this by starting with all the passive material (or probability) within a volume $V(0)$ with boundary $\Gamma(0)$. Suppose $f(x) = 1$ for $x \in V(0)$ and $f(x) = 0$ for $x \notin V(0)$. Let $V(t)$ be the volume we get by advecting every point within $V(0)$ by the velocity field a up to time t. It should be plausible that $u(x,t) > 0$ for $x \in V(t)$ but $u(x,t) = 0$ for $x \notin V(t)$. Here, the initial data is discontinuous across $\Gamma(0)$ and the solution is discontinuous across $\Gamma(t)$ (the advect up to time t of $\Gamma(0)$. By contrast, a solution of the heat equation would be continuous at any time $t > 0$. This makes the numerical solution of advection equations different from numerical solution of the heat equation. For advection equations, we look for methods that are good at preserving sharp interfaces (this is hard and research is ongoing today). For the heat equation, we hope not to preserve sharp interfaces.

2.1 Characteristics

We can get more intuition about solutions of the advection equation by applying the product rule to the space derivative part to rewrite (2) as

$$\partial_t u \; + \; a(x) \cdot \nabla u(x,t) \; = \; - \operatorname{div}(a(x))\, u(x,t) \; . \tag{4}$$

The left side sometimes is called the *advective derivative* because it is the time derivative of $u(x,t)$ as observed by a particle being advected by the velocity field a. This means that if $x(t)$ satisfies $\dot{x}(t) = a(x(t))$ (i.e. x is advected by a), then the chain rule gives

$$\frac{d}{dt}\, u(x(t),t) \; = \; \partial_t u \; + \; \dot{x} \cdot \nabla u \; = \; \partial_t u \; + \; a(x) \cdot \nabla u(x,t) \; .$$

Therefore (4) may be expressed as

$$\frac{d}{dt}\, u(x(t),t) \; = \; - \operatorname{div}(a(x(t)))\, u(x(t),t) \; . \tag{5}$$

The advected paths $x(t)$ are called *characteristics* (or *characteristic curves*). We see that if $\operatorname{div}(a) = 0$ then u is constant on characteristics, because the right side of (5) is zero. More generally, if $\operatorname{div}(a) > 0$ then the value of u decreases on characteristics, while u increases if $\operatorname{div}(a) < 0$.

These facts are easy to interpret in terms of local conservation. The sign of $\operatorname{div}(a)$ determines whether a increases or decreases volume (or area in 2D or length in 1D). For example, suppose $a(x) = \pm\, x$. Then $\operatorname{div}(a) = \pm\, n$. With the plus sign, a points radially outward and stretches (increases) volumes. With the minus sign, a compresses (decreases) volumes. Take the case of $\operatorname{div}(a) > 0$ and increasing volume. The amount of material in $V(0)$ at time $t = 0$ and the amount of material in $V(t)$ at time t should be the same (Problem 1 of

Homework 3). As the volume of $V(t)$ is larger than the volume of $V(0)$, the density of material must decrease to compensate for the in increase in volume. This is consistent with (5).

The characteristic form (4) of the advection equation also shows that solving the partial the partial differential equation is equivalent to solving ordinary differential equations. To solve the initial value problem for (4) we first compute the characteric curves by solving $\dot{x} = a(x)$, then we compute the values of u on the characteristic curves by solving

$$\dot{M} = A(t)M \ , \tag{6}$$

where $M(0) = f(x)$ is the given initial condition, $A(t) = -\text{div}(a(x(t))$, and the solution is $u(x(t), t) = M(t)$. There is a separate function $M(t)$ for each characteristic curve. Problem 2 of Homework 3 has a complete example to make this more concrete.

2.2 Qualitative theory, well posedness, estimates

The method of characteristics also allows us to prove that the initial value problem for (2) is well posed. Well posedness means that the solution to the initial value problem exists, is unique, and depends continuously on the initial data in a reasonable sense. First existence: the characteristic construction above yields a $u(x, t)$ that satisfies (4), which is equivalent to (2). A subtle point is that the characteristic construction does not show right away that $u(x, t)$ is differentiable. A straightforward but involved calculation (see below) shows that if the components of a are twice differentiable with all second partials bounded, and if f is differentiable with all first partials continuous, then the characteristic construction gives a u that has all first partials continuous and bounded. Given that, it is easy to verify that they satisfy the PDE.

The uniqueness is similar but easier. If u is a solution, then it must be the same as the one constructed by the method of characteristics. Indeed, we derived (5) assuming only that u is a classical solution to (2) (i.e. has continuous first partial derivatives that satisfy the PDE).

Continuous dependence on initial data (the last part of well posedness) relies on semi-quantitative *estimates* about general solutions. In PDE theory, an estimate is not an educated guess at the value of the solution, but an inequality that limits what the solution may do. Suppose $u(x, t)$ has initial data $u(x, 0) = f(x)$, and $v(x,)$ has initial data $v(x, 0) = g(x)$. Let $w = u - v$ and $h = f - g$. Since the PDE satisfies the superposition principle (is linear), w is the solution to the initial value problem with initial data h. Therefore, bounding (estimating) $u - v$ in terms of $f - g$ is the same as bounding w in terms of h.

One such bound is automatic from the equation. It follows immediately[1] from the equation that

$$\|w(\cdot, t)\|_{L^1} \ = \ \int_{R^n} |w(x, t)| \ dx \ = \ \int_{R^n} |h(x)| \ dx \ = \ \|h\|_{L^1} \ .$$

[1]This might take some thought, but the positive and negative parts of w are advected separately. Because characterics never cross, there is no cancellation.

This means that the average value of $|u(x,t) - v(x,t)|$ (averaging over x for a definite t value) is the same as the average of $|f(x) - g(x)|$. This is a form of continuity: as g approaches f, v approaches u, at least in the average deviation (L^1 norm) sense.

Another measure of closeness of functions is the *sup norm*:

$$\|f\|_{L^\infty} \;=\; \sup_x |f(x)| \;.$$

We write sup for *supremum* rather than max for *maximum* because the supremum might not be *attained*. For example, if $f(x) = 1 - e^{-x^2}$, then $\|f\|_{L^\infty} = 1$ although there is no x with $|f(x)| = 1$. If we assume that first partials of a are uniformly bounded so that

$$\sup_x |\operatorname{div}(a(x))| \;=\; K \;<\; \infty \;,$$

Then the characteristic equation (6) implies that if $M(t) = w(x(t),t)$, then $|M(t)| \leq e^{Kt} \sup |h(x)|$. Since this inequality applies to every value of $w(x,t)$, we have

$$\|u(\cdot,t) - v(\cdot,t)\|_{L^\infty} \;\leq\; e^{Kt} \|f - g\|_{L^\infty} \;. \tag{7}$$

This says that for any fixed t, the maximum difference between the solutions at time t is bounded in terms of the maximum difference of initial data. This is a different form of continuity.

Pay attention to the difference between the argument here and some of the arguments of Section 1. There we represent the solution of the heat equation as a convolution integral of the initial data with the fundamental solution. Here we find the solution of the advection equation in terms of the solutions of many ordinary differential equations. The difference is that here we showed that every solution of the advection equation may be described using characteristics. There we showed that convolution with the fundamental solution produces a solution, but not that every solution has that form. This means that the convolution solution of the heat equation does not yet prove that the solution is unique. There are two possible routes to uniqueness for (the initial value problem for) the heat equation. We may be able to prove that any solution of the heat equation initial value problem is given by the convolution. Or we may find a direct proof of some kind of estimate like the L^1 or L^∞ estimates we derived for the advection equation.

Also notice the difference between natural and less natural estimates. The most natural estimates are almost obvious from the nature of the problem. Here, that's the L^1 estimate that is obvious from the original conservation of stuff derivation. Less natural is the L^∞ estimate. If the equation had been

$$\partial_u \;+\; a(x) \cdot \nabla u \;=\; 0 \;, \tag{8}$$

then the L^∞ bound would have been obvious from the method of characteristics:

$$\frac{d}{dt}\, u(x(t),t) \;=\; 0 \;.$$

This would imply that (the values of u do not change with time, only the places where those values are taken move around with the characteristics)

$$\|u(\cdot,t)\|_{L^\infty} \;=\; \|u(\cdot,t)\|_{L^\infty} \;. \tag{9}$$

The actual estimate (estimate means upper bound) for the advection equation (2) replaces no growth (9) with exponential growth (7). This may be a large *quantitative* difference between the two problems, but it is a small *qualitative* difference. In both cases the estimate shows that the initial value problem is well posed in the L^∞ norm.

2.2.1 Bla bla

A very smart friend of mine used to talk about the *bla bla* of mathematics. This is the part where you quote theorems and perform mundane algebra to verify that the fundamental operations you did to derive something are correct. It is not particularly interesting but consumes a fair amount of time when you're trying to prove something. Here is some bla bla to go with the above discussion. We are going to prove that if a is twice differentiable and f is once differentiable then the characteristic construction gives a u that is once differentiable and satisfies the PDE in the classical sense. You don't have to read it.

I don't even have to write it.

3 Second order parabolic equations

A general second order parabolic PDE is an equation of the form

$$\partial_t u \;=\; \sum_{i,j=1}^{n} a_{ij}(x)\partial_{x_i}\partial_{x_j} u \;+\; \sum_{i=1}^{n} b_i(x)\partial_{x_i} u \;+\; c(x)u \;+\; f(x,t)\;. \tag{10}$$

The terms on the right side are the *second order terms*, the *first order terms*, the *undifferentiated* term, and the *inhomogeniety* (or *forcing*, or *source* term. The a_{ij}, the b_i and the c are the *coefficients*. *Constant coefficients* means that all of them are independent of x and t. As written, the coefficients depend on space but not time. We also could allow time dependent coefficients, which arise in some applications[2]. The equation is *parabolic* and *non-degenerate* if the $n \times n$ matrix a is positive definite at every point. It is usual to assume that the second order coefficient matrix is symmetric: $a_{ij} = a_{ji}$. If not, we can replace a_{ij} with the symmetrized $\frac{1}{2}(a_{ij} + a_{ji})$ without changing the solutions.

The equations may be manipulated to put the differentiation outside the coefficients. This is the *divergence form* of the equation. The trick is the "differentiation by parts" identity $f\partial_x g = \partial_x(fg) - (\partial_x f)g$. Applying this to the first order term in (10) gives

$$\sum_{i=1}^{n} b_i(x)\partial_{x_i} u \;=\; \sum_i \partial_{x_i}\left(b_i(x)\,u\right) \;-\; \left(\sum_i \partial_{x_i} b_i(x)\right) u\;.$$

[2]See Jim Gatheral's book on volatility surface modeling.

The last term on the right is $\mathrm{div}(b)u$. Manipulating the second order term will be easier if we use the *Einstein summation convention*, which is not to write the summation signs that really should be there. The second order term is

$$\begin{aligned} a_{ij}(x)\partial_{x_i}\partial_{x_j}u &= \partial_{x_i}\left(\, a_{ij}(x)\,\partial_{x_j}u\,\right) \,-\, (\,\partial_{x_i}a_{ij}(x)\,)\,\partial_{x_j}u \\ &= \partial_{x_i}\partial_{x_j}\left(\, a_{ij}(x)u\,\right) \,-\, \partial_{x_i}\left(\,\left(\,\partial_{x_j}a_{ij}(x)\,\right)\,u\,\right) \\ &- \partial_{x_j}\left(\,(\,\partial_{x_i}a_{ij}(x)\,)\,u\,\right) \,+\, \left(\partial_{x_j}\partial_{x_i}a_{ij}(x)\right)u\;. \end{aligned}$$

With this, the equation (10) may be rewritten

$$\partial_t u \;=\; \partial_{x_i}\partial_{x_j}\left(\, a_{ij}(x)u\,\right) \,+\, \partial_{x_i}\left(\,\widetilde{b}_i(x)u\right) \,+\, \widetilde{c}u + f\;, \tag{11}$$

where (implicit summation over j)

$$\widetilde{b}_i(x) \;=\; b_i(x) \;-\; \partial_{x_j}\left(\, a_{ij}(x) + a_{ji}(x)\,\right)\;,$$

and (summation over i and j)

$$\widetilde{c}(x) \;=\; c(x) \;-\; \partial_{x_i}b_i(x) \,+\, \partial_{x_i}\partial_{x_j}a_{ij}(x)\;.$$

The point of all this is that solutions of (10) and (11) have similar qualitative behavior even if the quantitative behavior is different.

3.1 Maximum principle

The *maximum principle* states that the maximum of the solution is attained on the boundary. As with superposition, it is not a principle, but a property that many PDEs have and many others do not. We may use it to show the uniqueness and continuous dependence parts of well-posedness. There are many other estimates, such as L^1 estimates, that also do this. Maximum principles are particularly useful for nonlinear Hamilton Jacobi Bellman equations that arise in optimal investment and stochastic control problems.

We start with the simplest example, the maximum principle for harmonic functions, and work up to more complicated ones. A function $u(x)$ with $x \in R^n$ is *harmonic* if . It is *subharmonic* if $\triangle u > 0$ and *superharmonic* if $\triangle u < 0$. As we will see, a subharmonic function is less than the harmonic function with the same boundary values and a superharmonic function is greater.

The basic idea is as follows. Suppose x_* is a local maximum: $u(x_*) = \max u(x)$. In order to be a local maximum, the Hessian matrix, $D^2u(x_*)$ must be negative semi-definite. If the matrix $D^2u(x_*)$ had a positive eigenvalue, then x_* would not be a local maximum. Therefore, there are two choices, either $D^2u(x_*)$ is the zero matrix or it has a negative eigenvalue. The point is that $\triangle u(x)$ is the trace of $D^2u(x)$, so $\triangle u(x_*) = 0$ is incompatible with $D^2u(x_*)$ having no positive eigenvalues, except for the possibility that $D^2u(x_*) = 0$.

The formal statement of the *weak maximum principle* for harmonic functions is that if V is a bounded open set in R^n whose boundary is Γ (notation from

previous sections), and if $u(x)$ is continuous with second partials continuous in V and u itself is continuous in the closure of V (which is $V \cup \Gamma$), then the maximum of u is attained on the boundary. That is, there is an $x_* \in \Gamma$ so that $u(x_*) = \max u(x)$, where the max is over all $x \in V \cup \Gamma$. A theorem of analysis states that the maximum of a continuous function on a bounded closed set (such as $V \cup \Gamma$) is attained; the maximum is the supremum. The point is to show that this maximum value is attained somewhere on the boundary. For example, let V be the unit ball, $u(x) = -|x|^2$, and Γ the unit sphere. The maximum of u is $u(0) = 0$ but the maximum of u on Γ is -1, which is less. This can not happen if u is harmonic or subharmonic.

The clever trick to get around the technicality above is to consider a sequence of strictly subharmonic functions $u_\epsilon(x) = u(x) + \epsilon |x|^2$. Let $M_\epsilon = \max u_\epsilon(x)$, with the maximum over $V \cup \Gamma$. We know that $M_\epsilon \to M_0$ as $\epsilon \to 0$ (think this through). We also know that there are points $x_\epsilon \in \Gamma$ so that $M_\epsilon = u_\epsilon(x_\epsilon)$. It is easy to see that $M_\epsilon \to \max_\Gamma u(x)$. You have to use the fact that $|x|$ is bounded on Γ so the maximum difference between u_ϵ and u goes to zero as $\epsilon \to 0$. This implies that the maximum of $u(x)$ is achieved by at least one point $x_* \in \Gamma$. The proof uses the assumption that u is continuous on $V \cup \Gamma$ and that the second partials of u are continuous within V. The latter implies that u_ϵ cannot have a local maximum within V.

This maximum principle implies that the solution to the Dirichlet boundary value problem for the Laplace equation is unique, if it exists, at least within the class of functions for which the maximum principle holds. Indeed, suppose u_1 and u_2 both are solution of $\triangle u = f$ for $xinV$ and $u(x) = g(x)$ for $x \in \Gamma$. Then $w(x) = u_2(x) - u_1(x)$ is harmonic ($\triangle w = 0$) for $x \in V$, and $w(x) = 0$ for $x \in \Gamma$. This implies that $\max w(x) = 0$, i.e. $w(x) \leq 0$ for all $x \in V$. Similarly, $-w(x)$ is harmonic and $-v(x) = 0$ for $x \in \Gamma$. This implies that $w(x) > 0$ for all $x \in V$. Together, we learn that $w(x) = 0$ for all $x \in V$, which is the statement that $u_1(x) = u_2(x)$ for all $x \in V$.

A similar argument applies to solutions of the heat equation. Suppose $u(x,t)$ satisfies $\partial_t u = \frac{1}{2} \triangle u + f(x,t)$ with $f(x,t) \leq 0$. Let $V \subset R^n$ be a bounded open set with boundary Γ. Choose a $T > 0$ and let W be the set of (x,t) values with $x \in V$ and $0 \leq t \leq T$. Let Δ be part of the boundary of W consisting of the sides ((x,t) with $x \in \Gamma$) and the bottom $(x,0)$ for $x \in V \cup \Gamma$). The weak maximum principle for the heat equation states that if u is differentiable in W and continuous in $W \cup \Delta$, then $\max_{W \cup \Delta} u(x,t) = \max_\Delta u(x,t)$. That is, the maximum is attained either at the initial time or somewhere on the boundary. A consequence of this maximum principle is a uniqueness theorem. If we have two solutions to the heat equation with the same initial and boundary conditions, then the solutions are the same. This is a big part of the proof that the initial boundary value problem for the heat equation is well posed.

The proof follows the proof above. The first step proves the statement when f is strictly negative. Suppose (x_*, t_*) is a local max of u, then $\partial_t u \geq 0$ (why?) and $D^2 u$ has no positive eigenvalues at (x_*, t_*). This contradicts f being strictly negative. Note that this allows the max to be at the bottom but not at the top. If we only know $f \leq 0$, we consider $u_\epsilon(x,t) = u(x,t) - \epsilon |x|^2$ as before. This is

the proof.

3.2 Other comparison principles

A comparison principle is a statement that if u and v are two solutions to a PDE and $u \leq v$ some places, then $u \leq v$ elsewhere. For example, if u and v are solutions of the heat equation with $u(x,0) \leq v(x,0)$ for all x, then $u(x,t) \leq v(x,t)$ for all $t > 0$ and all x. We can prove this by looking at $w(x,t) = u(x,t) - v(x,t)$. This is a solution of the heat equation (superposition) and satisfies $w(x,0) \leq 0$ for all x. The maximum principle states that the maximum of $w(x,t)$ over all x and $t > 0$ is achieved with $t = 0$. This implies that $w(x,t) \leq 0$ for all x and $t > 0$. The same can be said for two solutions of the Laplace equation: if $u(x) \leq v(x)$ for $x \in \Gamma$, then $u(x) \leq v(x)$ for $x \in D$.

Conversely, comparison principles imply maximum principles. Suppose, for example, that we know a comparison principle for solutions of the Laplace equation on D with boundary Γ. Let $M = \max_{x \in \Gamma} u(x)$. Let $v(x) = M$ for all x. This (constant) function satisfies the PDE $\triangle v = 0$ and $v(x) \geq u(x)$ for $x \in D$. Therefore, $u(x) \leq v(x) = M$ for all $x \in D$. This is the maximum principle.

Another common comparison principle is that if $u \geq 0$ in some places then $u \geq 0$ elsewhere. This would be a consequence of a maximum principle, if there were one. But sometimes it may apply even when the maximum principle does not. For example, consider

$$\partial_t u(x,t) \;=\; \frac{1}{2} \triangle u \;+\; c(x) u \;.$$

If $c > 0$, the solutions can grow (e.g. if u and c are independent of x). Still, if you review the proof of the maximum principle for the heat equation, you will be able to prove that if $u(x,0) > 0$ for all x then $u(x,t) > 0$ for all x and $t > 0$. This does not depend on the sign of c. Let x_* and $t_* > 0$ be the first place $u = 0$. This is a local minimum in x, and has $\partial_t u \leq 0$, and has $c(x_*) u(x_*, t_*) = 0$. This brings us exactly back to our earlier argument. A negative x can make u smaller, but it cannot make u cross zero.

3.3 Representations of the solution

One way to prove uniqueness for first order conservation equations was to prove the method of characteristics formula. This formula follows directly from the equation. We can do the same for the initial value problem for the heat equation. We showed that the integral with the Green's function gives a solution. We also proved a uniqueness theorem using the maximum principle, which shows that this is the only solution. Still, it is useful to have a direct derivation of the integral representation from the equation. The duality argument is useful in other situations.

Suppose $u(x,t)$ satisfies the heat equation and $v(x,t)$ satisfies the backward heat equation $\partial_t v = \frac{-1}{2} \triangle v$. In the application, we will take $v(x,t) = G_0(x - y, T - t)$ for some $T > 0$ and some value of y. If $u(x,t)$ is defined for $t \geq 0$,

our formulas will make sense in the range $0 \le t \le T$. The main fact is that the integral

$$I(t) \;=\; \int_{R^n} u(x,t)v(x,t)\,dx$$

is independent of time. You can check this by differentiating I with respect to t, using the u and v equations, and integrating by parts. You must use the hypothesis that the boundary terms (called surface terms in more than one dimension) vanish as you take the boundary to infinity. This can be a bit technical, which is not the purpose of this class.

However, following Kohn's notes, I want to illustrate that you can prove uniqueness (and the integral representation formula) even for solutions of the heat equation that tend to infinity as $x \to \infty$. The maximum principle might have trouble applying if the maximum is always infinity. Kohn gives a (standard) example of a function that satisfies the heat equation for all x and $t > 0$ and has $u(x,0) = 0$ but $u(x,t) > 0$ for $t > 0$. This shows that the uniqueness theorem is not quite as simple as I have been hinting. These bad solutions grow as fast as e^{x^2} for large x. One can prove a uniqueness theorem under a hypothesis that rules out such growth but still allows a simple exponential. More precisely, suppose $u(x,t)$ is continuous for all x and $t \ge 0$ and has two continuous derivatives for $t > 0$ and that

$$|u(x,t)| \;\le\; Ce^{a|x|}\;,\;\; |\nabla u(x,t)| \;\le\; Ce^{a|x|} \tag{12}$$

for some finite C and positive a. Then (check this)

$$\frac{d}{dt}\int_{|x|\le R} u(x,t)G_0(x-y,T-t)\,dx \;=\; \int_{|x|=R} \partial_n u G_0 - u\partial_n G_0\,dA(x)\;.$$

Now fix $T > 0$ and y and suppose $0 < t < T$. Then the right side goes to zero as $R \to \infty$ because the u part grown at most exponentially while the G_0 part decays like $e^{-b|x|^2}$.

If we set $t = 0$, we get that the integral is equal to

$$\int u(x,0)G_0(x-y,T)\,dx\;.$$

If we take the limit $t \to T$ we see that we get $u(y,T)$. (more details coming?) This is the fundamental solution formula

$$u(y,T) \;=\; \int G_0(x-y,t)u(x,0)\,dx\;.$$

Now we have derived it rather than simply checking that it gives a solution.