8 Heat equation: properties

We would like to solve the heat (diffusion) equation

$$u_t - ku_{xx} = 0, \tag{1}$$

and obtain a solution formula depending on the given initial data, similar to the case of the wave equation. However the methods that we used to arrive at d'Alambert's solution for the wave IVP do not yield much for the heat equation. To see this, recall that the heat equation is of parabolic type, and hence, it has only one family of characteristic lines. If we rewrite the equation in the form

$$ku_{xx} + \cdots = 0,$$

where the dots stand for the lower order terms, then you can see that the coefficients of the leading order terms are

$$A = k, \quad B = C = 0.$$

The slope of the characteristic lines are then

$$\frac{dt}{dx} = \frac{B \pm \sqrt{\Delta}}{2A} = \frac{B}{2A} = 0.$$

Consequently, the single family of characteristic lines will be given by

$$t = c.$$

These characteristic lines are not very helpful, since they are parallel to the x axis. Thus, one cannot trace points in the xt plane along the characteristics to the x axis, along which the initial data is defined. Notice that there is also no way to factor the heat equation into first order equations, either, so the methods used for the wave equation do not shed any light on the solutions of the heat equation.

Instead, we will study the properties of the heat equation, and use the gained knowledge to devise a way of reducing the heat equation to an ODE, as we have done for every PDE we have solved so far.

8.1 The maximum principle

The first properties that we need to make sure of, are the uniqueness and stability for the solution of the problem with certain auxiliary conditions. This would guarantee that the problem is wellpossed, and the chosen auxiliary conditions do not break the physicality of the problem. We begin by establishing the following property, that will be later used to prove uniqueness and stability.

Maximum Principle. If $u(x,t)$ satisfies the heat equation (1) in the rectangle $R = \{0 \le x \le l, 0 \le t \le T\}$ in space-time, then the maximum value of $u(x,t)$ over the rectangle is assumed either initially ($t = 0$), or on the lateral sides ($x = 0$, or $x = l$).

Mathematically, the maximum principle asserts that the maximum of $u(x,t)$ over the three sides must be equal to the maximum of the $u(x,t)$ over the entire rectangle. If we denote the set of points comprising the three sides by $\Gamma = \{(x,t) \in R \mid t = 0 \quad \text{or} \quad x = 0 \quad \text{or} \quad x = l\}$, then the maximum principle can be written as

$$\max_{(x,t)\in\Gamma}\{u(x,t)\} = \max_{(x,t)\in R}\{u(x,t)\}. \tag{2}$$

If you think of the heat conduction phenomena in a thin rod, then the maximum principle makes physical sense, since the initial temperature, as well as the temperature at the endpoints will dissipate through conduction of heat, and at no point the temperature can rise above the highest initial or endpoint temperature. In fact, a stronger version of the maximum principle holds, which asserts that the maximum over the rectangle R can not be attained at a point not belonging to Γ, unless $u \equiv$ constant, i.e. for nonconstant solutions the following strict inequality holds

$$\max_{(x,t)\in R\setminus\Gamma}\{u(x,t)\} < \max_{(x,t)\in R}\{u(x,t)\},$$

where $R \setminus \Gamma$ is the set of all points of R that are not in Γ (difference of sets). This makes physical sense as well, since the heat from the point of highest initial or boundary temperature will necessarily transfer to points of lower temperature, thus decreasing the highest temperature of the rod.

We finally note, that the maximum principle also implies a *minimum principle*, since one can apply it to the function $-u(x,t)$, which also solves the heat equation, and make use of the following identity,

$$\min\{u(x,t)\} = -\max\{-u(x,t)\}.$$

Thus, the minima points of the function $u(x,t)$ will exactly coincide with the maxima points of $-u(x,t)$, of which, by the maximum principle, there must necessarily be in Γ.

Proof of the maximum principle. If the maximum of the function $u(x,t)$ over the rectangle R is assumed at an internal point (x_0, t_0), then the gradient of u must vanish at that point, i.e. $u_t(x_0,t_0) = u_x(x_0,t_0) = 0$. If in addition we had the strict inequality $u_{xx}(x_0,t_0) < 0$, then one would get a contradiction by plugging the point (x_0,t_0) into the heat equation. Indeed, we would have

$$u_t(x_0,t_0) - ku_{xx}(x_0,t_0) = -ku_{xx}(x_0,t_0) > 0.$$

This contradicts the heat equation (1), which must hold for all values of x and t. Thus, the contradiction would imply that the maximum point (x_0,t_0) cannot be an internal point. However, from the second derivative test we have the weaker inequality $u_{xx}(x_0,t_0) \leq 0$ (the point would not be a maximum if $u_{xx}(x_0,t_0) > 0$), which is not enough for this argument to go through.

The way out, is to recycle the above argument with a slight modification to the function u. Define a new function

$$v(x,t) = u(x,t) + \epsilon x^2, \tag{3}$$

where $\epsilon > 0$ is a constant that can be taken as small as one wants. Now let M be the maximum value of u over the three sides, which we denoted by Γ above. That is

$$M = \max_{(x,t)\in\Gamma}\{u(x,t)\}.$$

To prove the maximum principle, we need to establish (2). The maximum over Γ is always less than or equal to the maximum over R, since $\Gamma \subset R$. So we only need to show the opposite inequality, which would follow from showing that

$$u(x,t) \leq M, \qquad \text{for all the points } (x,t) \in R. \tag{4}$$

Notice that from the definition of v, we have that at the points of Γ, $v(x,t) \leq M + \epsilon l^2$, since the maximum value of ϵx^2 on Γ is ϵl^2. Then, instead of proving inequality (4), we will prove that

$$v(x,t) \leq M + \epsilon l^2, \qquad \text{for all the points } (x,t) \in R, \tag{5}$$

which implies (4). Indeed, from the definition of v in (3), we have that in the rectangle R

$$u(x,t) \leq v(x,t) - \epsilon x^2 \leq M + \epsilon(l^2 - x^2),$$

where we used (5) to bound $v(x,t)$. Now, since the point (x,t) is taken from the rectangle R, we have that $0 \leq x \leq l$, and the difference $l^2 - x^2$ is bounded. But then the right hand side of the above inequality can be made as close to M as possible by taking ϵ small enough, which implies the bound (4).

If we formally apply the heat operator to the function v, and use the definition (3), we will get

$$v_t - kv_{xx} = u_t - k(u_{xx} + 2\epsilon) = (u_t - ku_{xx}) - 2k\epsilon < 0,$$

since both $k, \epsilon > 0$, and u satisfies the heat equation (1) on R. Thus, v satisfies the *heat inequality* in R

$$v_t - kv_{xx} < 0. \tag{6}$$

We now recycle the above argument, which barely failed for u, applying it to v instead. Suppose $v(x,t)$ attains its maximum value at an internal point (x_0,t_0). Then necessarily $v_t(x_0,t_0)=0$, and $v_{xx}(x_0,t_0)\leq 0$. Hence, at this point we have

$$v_t(x_0,t_0)-kv_{xx}(x_0,t_0)=-kv_{xx}(x_0,t_0)\geq 0,$$

which contradicts the heat inequality (6). Thus, v cannot have an internal maximum point in R.

Similarly, suppose that $v(x,t)$ attains its maximum value at a point (x_0,t_0) on the fourth side of the rectangle R, i.e. when $t_0=T$. Then we still have that $v_x(x_0,t_0)=0$, and $v_{xx}(x_0,t_0)\leq 0$, but $v_t(x_0,t_0)$ does not have to be zero, since $t_0=T$ is an endpoint in the t direction. However, from the definition of the derivative, and our assumption that (x_0,t_0) is a point of maximum, we have

$$v_t(x_0,t_0)=\lim_{\delta\to 0+}\frac{v(x_0,t_0)-v(x_0,t_0-\delta)}{\delta}\geq 0.$$

So at this point we still have

$$v_t(x_0,t_0)-kv_{xx}(x_0,t_0)\geq 0,$$

which again contradicts the heat inequality (6).

Now, since the continuous function $v(x,t)$ must attain its maximum value somewhere in the closed rectangle R, this must happen on one of the remaining three sides, which comprise the set Γ. Hence,

$$v(x,t)\leq \max_{(x,t)\in R}\{v(x,t)\}=\max_{(x,t)\in\Gamma}\{v(x,t)\}\leq M+\epsilon l^2,$$

which finishes the proof of (5). □

8.2 Uniqueness

Consider the Dirichlet problem for the heat equation,

$$\begin{cases} u_t-ku_{xx}=f(x,t) \qquad \text{for} \quad 0\leq x\leq l, \quad t>0\\ u(x,0)=\phi(x), \\ u(0,t)=g(t), \quad u(l,t)=h(t), \end{cases} \tag{7}$$

for given functions f,ϕ,g,h. We will use the maximum principle to show uniqueness and stability for the solutions of this problem (recall that last time we used the energy method to prove uniqueness for the same problem).

Uniqueness of solutions. There is at most one solution to the Dirichlet problem (7).

Indeed, arguing from the inverse, suppose that there are two functions, u, and v, that both solve the inhomogeneous heat equation and satisfy the initial and Dirichlet boundary conditions of (7). Then their difference, $w=u-v$, satisfies the homogeneous heat equation with zero initial-boundary conditions, i.e.

$$\begin{cases} w_t-kw_{xx}=0 \qquad \text{for} \quad 0\leq x\leq l, \quad t>0\\ w(x,0)=0, \\ u(0,t)=0, \quad u(l,t)=0, \end{cases} \tag{8}$$

But from the maximum principle, we know that w assumes its maximum and minimum values on one of the three sides $t=0$, $x=0$, and $x=l$. And since $w=0$ on all of these three sides from the initial and boundary conditions in (8), we have that for $x\in[0,l], t>0$

$$0\leq w\leq 0 \quad \Rightarrow \quad w(x,t)\equiv 0.$$

Hence,

$$u-v=w\equiv 0, \qquad \text{or} \qquad u\equiv v,$$

and the solution must indeed be unique.

Notice again that all of the above arguments hold for the case of the infinite interval $-\infty<x<\infty$ as well. In this case one ignores the effect of the infinitely far endpoints and considers an IVP. And the maximum principle simply asserts that the maximum of the solutions must be attained initially. We will use this in the next lecture when deriving the solution for the IVP for the heat equation on the entire real line $x\in\mathbb{R}$.

8.3 Stability

Stability of solutions with respect to the auxiliary conditions is the third ingredient of well-posedness, after existence and uniqueness. It asserts that *close* auxiliary conditions lead to *close* solutions. There are, however, different ways of measuring closeness of functions, which initial and boundary data, as well as the solutions are.

Consider two solutions, u_1, u_2, of the heat equation (1) for $x \in [0, l], t > 0$, which satisfy the following initial-boundary conditions

$$\begin{cases} u_1(x,0) = \phi_1(x), \\ u_1(0,t) = g_1(t), \quad u_1(l,t) = h_1(t), \end{cases} \qquad \begin{cases} u_2(x,0) = \phi_2(x), \\ u_2(0,t) = g_2(t), \quad u_2(l,t) = h_2(t). \end{cases} \tag{9}$$

Stability of solutions means that *closeness* of ϕ_1 to ϕ_2, g_1 to g_2 and h_1 to h_2 implies the closeness of u_1 to u_2. Notice that the difference $w = u_1 - u_2$ solves the heat equation as well, and satisfies the following initial-boundary conditions

$$\begin{cases} w_1(x,0) = \phi_1(x) - \phi_2(x), \\ w(0,t) = g_1(t) - g_2(t), \quad w(l,t) = h_1(t) - h_2(t). \end{cases}$$

But then the maximum and minimum principles imply

$$-\max_{(x,t)\in\Gamma}\{|w(x,t)|\} \le \max_{\substack{0\le x\le l \\ t\ge 0}}\{w(x,t)\} \le \max_{(x,t)\in\Gamma}\{|w(x,t)|\},$$

and hence, the absolute value of the difference $u_1 - u_2$ will be bounded by

$$\begin{aligned} \max_{\substack{0\le x\le l \\ t\ge 0}}\{|u_1(x,t) - u_2(x,t)|\} = \max_{\substack{0\le x\le l \\ t\ge 0}}\{|w(x,t)|\} &\le \max_{(x,t)\in\Gamma}\{|w(x,t)|\} \\ &= \max_{\substack{0\le x\le l \\ t\ge 0}}\{|\phi_1(x) - \phi_2(x)|, |g_1(t) - g_2(t)|, |h_1(t) - h_2(t)|\}. \end{aligned}$$

Thus, the smallness of the maximum of the differences $|\phi_1 - \phi_2|$, $|g_1 - g_2|$ and $|h_1 - h_2|$ implies the smallness of the maximum of the difference of solutions $|u_1 - u_2|$. In this case the stability is said to be in the *uniform* sense, i.e. smallness is understood to hold uniformly in the (x, t) variables.

An alternate way of showing the stability is provided by the energy method. Suppose u_1 and u_2 solve the heat equation with initial data ϕ_1 and ϕ_2 respectively, and zero boundary conditions. This would be the case for the problem over the entire real line $x \in \mathbb{R}$, or if $g_1 = g_2 = h_1 = h_2 = 0$ in (9). In this case the energy method for the difference $w = u_1 - u_2$ implies that $E[w](t) \le E[w](0)$ for all $t \ge 0$, or

$$\int_0^l [u_1(x,t) - u_2(x,t)]^2\,dx \le \int_0^l [\phi_1(x) - \phi_2(x)]^2\,dx, \qquad \text{for all} \quad t \ge 0.$$

Thus the closeness of ϕ_1 to ϕ_2 in the sense of the square integral of the difference implies the closeness of the respective solutions in the same sense. This is called stability in the *square integral* (L^2) sense.

8.4 Conclusion

As expected, the method of characteristics is inefficient for solving the heat equation. We then need to find an alternative method of reducing the equation to an ODE. But before embarking on this path, we first study the properties of the heat equation, which will serve as beacons in the later reduction to an ODE. Today we established the maximum principle for the heat equation, which immediately implied the uniqueness and stability for the solution. Next time we will look at the invariance properties of the equation and derive the solution using these properties.