The heat equation is the prototypical parabolic PDE:

$$\boxed{u_t = \Delta u.} \tag{1}$$

This equation describes the isotropic diffusion of a quantity that might, for example, be heat in a solid or concentration of salt in a motionless body of water.

The history begins with the work of Joseph Fourier around 1807. In a remarkable memoir, Fourier invented both the equation (1) and the method of Fourier analysis for its solution. For definiteness, let us consider the one-dimensional problem $u_t = u_{xx}$ for $x \in \mathbb{R}$ with initial data $u_0(x)$. The Fourier transform decomposes u_0 into its components at various wave numbers k:

$$u_0(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathrm{i}kx}\hat{u}_0(k)\mathrm{d}k, \tag{2}$$

where $\hat{u}_0(k)$ is defined by the integral

$$\hat{u}_0(k) = \int_{-\infty}^{\infty} e^{-\mathrm{i}kx}u_0(x)\mathrm{d}x. \tag{3}$$

The evolution of each component $e^{\mathrm{i}kx}$ under (1) is a trivial matter—it decays at the rate e^{-k^2t}. Superposition gives us the evolution of the general initial function u_0:

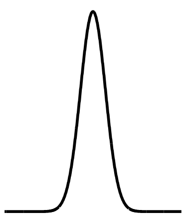
$$u(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathrm{i}kx-k^2t}\hat{u}_0(k)\mathrm{d}k.$$

Replacing $\hat{u}_0(k)$ by its integral (3) and applying an identity for the Fourier transform of a Gaussian yields the formula

$$u(x,t) = \frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{\infty} e^{-(x-s)^2/4t}u_0(s)\mathrm{d}s. \tag{4}$$

In a bounded domain in $\mathbb{R}$ or $\mathbb{R}^n$, an analogous treatment of the heat equation would go by separation of variables, leading to solutions of the form $e^{\alpha_j t}\phi_j(x)$, where the functions $\phi_j(x)$ are eigenfunctions of the Laplacian operator for Ω ($\rightarrow$ *ref*).

Fig. 1: Gaussian kernel of height $(4\pi t)^{-1/2}$, width $O(t^{1/2})$



Equation (4) asserts that the solution to (1) at time t is the convolution of the initial data u_0 with the *Gaussian kernel* $e^{-(x-s)^2/4t}/\sqrt{4\pi t}$, whose integral is 1. Heat is conserved ($\|u(t)\|_1 = \|u_0\|_1$ for all $t>0$), but it diffuses over a range of order $\sqrt{t}$, the width of the Gaussian. At time t, any structures of wavelengths shorter than $O(\sqrt{t})$ will have been smoothed away. Since the tail of the Gaussian is never zero, on the other hand, a small amount of information propagates unboundedly fast, in contrast to the situation for a hyperbolic PDE.

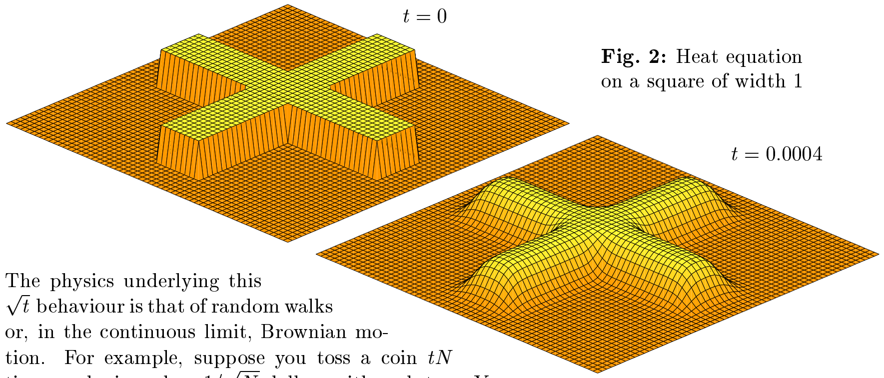


Fig. 2: Heat equation on a square of width 1

The physics underlying this $\sqrt{t}$ behaviour is that of random walks or, in the continuous limit, Brownian motion. For example, suppose you toss a coin tN times and win or lose $1/\sqrt{N}$ dollars with each toss. Your profit follows a binomial distribution that converges to the Gaussian $e^{-x^2/4t}/\sqrt{4\pi t}$ in the limit $N \rightarrow \infty$. Arbitrarily large profits are possible, but anything much bigger than $\sqrt{t}$ is very unlikely. This $\sqrt{t}$ effect is at the root of much of the field of statistics, and it was the basis of Einstein's epochal paper on Brownian motion in his *annus mirabilis* 1905.

It seems obvious that the solution to (1) should be unique, but in fact it is not. There are other solutions besides (4) in which an infinite amount of heat floods in from infinity just after $t=0$. For example, if $g(t) = \exp(-t^{-2})$, then the power series $\sum_{k=0}^{\infty}(\mathrm{d}^kg(t)/\mathrm{d}t^k)\,x^{2k}/(2k)!$ converges for each $t>0$ to an analytic function of x that satisfies (1) with initial data $u_0 = 0$. However, uniqueness for (1) is achieved if we require that $|u(x,t)|$ is bounded as $|x| \rightarrow \infty$. In fact, it is enough to require that $u(x,t)$ or $-u(x,t)$ is bounded, and thus, for example, if $u_0(x) \ge 0$, then (1) has a unique solution with $u(x,t) \ge 0$.

The heat equation is the canonical smoothing process, and as a application of this property we can prove the *Weierstrass approximation theorem:* a continuous function f on $[-1,1]$ can be approximated to within any error ε by a polynomial p. Take f, extended continuously to a function of compact support on $\mathbb{R}$, as initial data u_0 for (1). Let a sufficiently small time t elapse so that $|u(x,t) - u_0(x)| < \varepsilon/2$ everywhere. The function $u(x,t)$ is an entire function of x, i.e., analytic for all x. Now take $p(x)$ to be a truncation of the power series for $u(x,t)$ to enough terms so that $|\,p(x) - u(x,t)| < \varepsilon/2$ for all $x \in [-1,1]$.

References

H. S. Carslaw and J. C. Jaeger, *Conduction of heat in solids*, Oxford University Press, 1986.

J. Crank, *The mathematics of diffusion*, Oxford University Press, 1979.

A. Einstein, *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Ann. Phys. **17** (1905), 549–560.

J. Fourier, *La théorie analytique de la chaleur*, Paris, 1822.

F. John, *Partial differential equations*, 4th ed., Springer, 1986.

D. V. Widder, *The heat equation*, Academic Press, 1975.