Introduction to Partial Differential Equations
Exercise sheet 10

Ross Ogilvie

4th November, 2024

30. Special solutions of the heat equation

(a) Solutions of PDEs that are constant in the time variable are called "steady-state" solutions. Describe steady-state solutions of the inhomogeneous heat equation. *(1 point)*

(b) Consider the heat equation $\dot{u} - \Delta u = 0$ on $\mathbb{R}^n \times \mathbb{R}^+$ with smooth initial condition $u(x, 0) = h(x)$. Suppose that the Laplacian of h is a constant. Show that there is a solution whose time derivative is constant. *(1 point)*

(c) Consider "translational solutions" to the heat equation on $\mathbb{R} \times \mathbb{R}^+$ (ie $n = 1$). These are solutions of the form $u(x, t) = F(x - bt)$. Find all such solutions. *(2 points)*

(d) If u is a solution to the heat equation, show for every $\lambda \in \mathbb{R}$ that $u_\lambda(x, t) := u(\lambda x, \lambda^2 t)$ is also a solution to the heat equation. *(2 points)*

Solution.

(a) Steady-state solutions are defined by $\dot{u} = 0$. The heat equation then reduces to a Poisson equation: $0 - \Delta u = f$.

(b) If the time derivative of u is constant, then it follows that $u(x, t) = u_0(x) + u_1 t$. From the initial condition we must have $u_0 = h$. Putting this into the heat equation gives

$$u_1 - \Delta h = 0.$$

So choose the constant $u_1 = \Delta h$.

(c) Putting this into the heat equation gives

$$-bF' - F'' = 0.$$

Integrating gives

$$bF + F' = C.$$

This has the solution $F(y) = C + A \exp(-by)$. So we get the solution $u(x, t) = C + A \exp(-bx + b^2 t)$.

(d) This follows because $\partial_t(u(\lambda x, \lambda^2 t)) = \lambda^2 \dot{u}(\lambda x, \lambda^2 t)$ and $\partial_j^2(u(\lambda x, \lambda^2 t)) = \lambda^2 (\partial_j^2 u)(\lambda x, \lambda^2 t)$.

31. The Fourier transform

In this question we expand on some details from Section 4.1. Recall that the Fourier transform of a function $h(x) : \mathbb{R}^n \to \mathbb{R}$ is defined to be a function $\hat{h}(k) : \mathbb{R}^n \to \mathbb{R}$ given by

$$\hat{h}(k) = \int_{\mathbb{R}^n} e^{-2\pi i k \cdot x} h(x)\, dx.$$

Lemma 4.3 shows that it is well-defined for Schwartz functions.

(a) Give the definition of a Schwartz function. *(1 point)*

(b) Argue that $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \exp(-x^2)$ is a Schwartz function. *(2 points)*

(c) Consider

$$I^2 = \left(\int_{\mathbb{R}} e^{-x^2}\, dx \right)^2 = \left(\int_{\mathbb{R}} e^{-x^2}\, dx \right) \left(\int_{\mathbb{R}} e^{-y^2}\, dy \right) = \int_{\mathbb{R}^2} e^{-x^2-y^2}\, dx\, dy.$$

By changing to polar coordinates, compute this integral. *(1 point)*

(d) Prove the rescaling law for Fourier transforms: if $h(x) = g(ax)$ then *(1 point)*

$$\hat{h}(k) = |a|^{-n} \hat{g}(a^{-1}k).$$

(e) Prove the shift law for Fourier transforms: if $h(x) = g(x - a)$ then *(1 point)*

$$\hat{h}(k) = e^{-2\pi i a \cdot k} \hat{g}(k).$$

(f) Show that δ is a tempered distribution. *(2 points)*

(g) Compute the Fourier transform of δ. *(2 points)*

(h) Try to compute the Fourier transform of 1. What is the difficulty?

(2 points)

Solution.

(a) A Schwartz function is a smooth functions whose partial derivatives (of all orders) decay faster than the reciprocal of any polynomial. For such functions, for all multi-indices α and $k \in \mathbb{N}$

$$\lim_{|x| \to \infty} |x|^k \partial^\alpha f(x) = 0.$$

Clearly because this decays to zero, $\sup |x|^k |\partial^\alpha f(x)|$ exists. Conversely, if this supremum exists, then

$$\lim_{|x| \to \infty} |x|^k |\partial^\alpha f(x)| = \lim_{|x| \to \infty} |x|^{-1} |x|^{k+1} |\partial^\alpha f(x)| \leq \lim_{|x| \to \infty} |x|^{-1} \sup |x|^{k+1} |\partial^\alpha f(x)| = 0.$$

Therefore these two conditions are equivalent. The second version, $\sup |x|^k |\partial^\alpha f(x)| < \infty$, is often more useful.

(b) This is a function of one variable, so we do not need to use multi-indices.

$$\begin{aligned} f &= e^{-x^2}, \\ f' &= -2xe^{-x^2} \\ f'' &= -2e^{-x^2} + 4x^2 e^{-x^2}. \end{aligned}$$

It is clear that higher derivatives derivative have the form $P(x)e^{-x^2}$ where $P(x)$ is a polynomial. The exponential factor is dominant for large $|x|$, so $f^{(n)}$ decays faster than any polynomial. Thus f is Schwartz.

(c) After the substitution into polar coordinates we have

$$I^2 = \int_0^{2\pi} \int_0^{\infty} \exp(-r^2)\, r\, dr\, d\theta = \int_0^{2\pi} \Big[-\frac{1}{2} \exp(-r^2) \Big]_0^{\infty} d\theta = \pi.$$

Thus $I = \sqrt{\pi}$.

This integral has a long history. As best as I can tell, de Moivre was the first to prove something equivalent to this calculation. In 1733 he showed that the binomial distribution in probability theory could be approximated by an exponential term

$$\binom{n}{\frac{n}{2} + d} 0.5^n \approx Ce^{-2d^2/n}$$

and gave a numerical value for the constant. This is equivalent an approximation for the factorial, to which Stirling is given credit for calculating exact constant

$$\lim_{n \to \infty} \frac{n!}{n^{n+0.5} e^{-n}} = \sqrt{2\pi}$$

using the Wallis product for π (1656)

$$\frac{\pi}{2} = \left(\frac{2}{1} \cdot \frac{2}{3} \right) \left(\frac{4}{3} \cdot \frac{4}{5} \right) \left(\frac{6}{5} \cdot \frac{6}{7} \right) \cdots$$

De Moivre and Stirling were contemporaries, Stirling had just written a big book on infinite series, and de Moivre (according to a quote I've seen) gives Stirling credit, but it's not entirely clear to me how involved Stirling actually was. Perhaps de Moivre was just being modest. My sources for these historical remarks are Lee, Pearson, 1924, and Stahl, 2006.

In any case, since the sum of probabilities must add to 1, and the sum can be approximated by an integral, this implies the result. The first person to explicitly state the result was Laplace in 1774. Gauss' name seems be attached to this integral because of his role in popularising e^{-x^2} as a model of measurement errors in astronomy and other sciences. The modern, most common, proof is due to Poisson, simplifying a method of his PhD advisor Laplace. This is the method above and it's 'one from the book'.

In the script we needed a similar integral to compute the Fourier transform of the Gaussian curve, except in multiple variables and along a complex line. The integrals separate, so it is sufficient to consider the problem in one variable, namely that $\int_{ik+\mathbb{R}} \exp(-x^2)\, dx = \sqrt{\pi}$ for any k. Let us present some arguments for this.

Argument from complex analysis: For the integral along a line in the complex plane, it is of course natural to use a result of complex analysis. Consider

$$g(k) = \int_{\mathbb{R}} \exp -[\pi ik + y]^2 \, dy.$$

This is an analytic function in the variable $k \in \mathbb{C}$. For all $k \in i\mathbb{R}$, the transformation $z = \pi ik + y$ is a real transformation, so it does indeed reduce to $g(k) = \int_{\mathbb{R}} \exp(-z^2)\, dz = \sqrt{\pi}$. The unique continuation property says that two (in this case complex) analytic functions that agree on a sequence and its limit must agree everywhere. Think of the right hand side as the constant function $\sqrt{\pi}$, which is analytic. Hence $g(k) = \sqrt{\pi}$ for all k, not just imaginary k.

Argument from Schmidt: In previous years Prof Schmidt gave an very concrete argument using power series. We begin with the observation

$$\int_{\mathbb{R}} e^{-(\pi ik+y)^2+(\pi ik)^2}\, dy = \int_{\mathbb{R}} e^{-2\pi iky-y^2}\, dy.$$

Let us then investigate the following integral for real values of ω. By the same algebraic manipulations, we have

$$\int_{\mathbb{R}} e^{-(\omega+y)^2+\omega^2}\, dy = \int_{\mathbb{R}} e^{-2\omega y-y^2}\, dy.$$

The left hand side is easy to manipulate

$$\int_{\mathbb{R}} e^{-(\omega+y)^2+\omega^2}\, dy = e^{\omega^2}\int_{\mathbb{R}} e^{-(\omega+y)^2}\, dy = e^{\omega^2}\int_{\mathbb{R}} e^{-z^2}\, dz = \sqrt{\pi}e^{\omega^2} = \sum_{l=0}^{\infty}\frac{\sqrt{\pi}}{l!}\omega^{2l}$$

On the right hand side, we expand this into a power series in ω:

$$\int_{\mathbb{R}} e^{-2\omega y-y^2}\, dy = \int_{\mathbb{R}} e^{-y^2}\sum_{l=0}^{\infty}\frac{(-2\omega y)^l}{l!}\, dy = \sum_{l=0}^{\infty}\left(\int_{\mathbb{R}} e^{-y^2}\frac{(-2y)^l}{l!}\, dy\right)\omega^l$$

These two power series are equal, so their coefficients must be equal. In other words

$$\int_{\mathbb{R}} e^{-y^2}\frac{(-2y)^l}{l!}\, dy = \begin{cases}\frac{\sqrt{\pi}}{(l/2)!} & l \text{ even}\\ 0 & l \text{ odd}\end{cases}$$

We can now return to the imaginary case. Again, making the power series expansion in k

$$\begin{aligned}\int_{\mathbb{R}} e^{-2\pi iky-y^2}\, dy &= \int_{\mathbb{R}} e^{-y^2}\sum_{l=0}^{\infty}\frac{(-2\pi iky)^l}{l!}\, dy\\ &= \sum_{l=0}^{\infty}\left(\int_{\mathbb{R}} e^{-y^2}\frac{(-2y)^l}{l!}\, dy\right)(\pi ik)^l\\ &= \sum_{l=0}^{\infty}\frac{\sqrt{\pi}}{l!}(\pi ik)^{2l}\\ &= \sqrt{\pi}\sum_{l=0}^{\infty}\frac{(-\pi^2k^2)^l}{l!}\\ &= \sqrt{\pi}\exp(-\pi^2k^2).\end{aligned}$$

This also gives the result. Note here that at the core of the argument was "two equal power series have equal coefficients". This is how one proves that analytic functions have the unique continuation property. So the above two proofs are not as different as they first appear.

Argument from differentiating: I've saved the simplest method until last. Consider the function $g(k)$ again and differentiate

$$\begin{aligned} g(k) &= \int_{\mathbb{R}} e^{-[\pi ik+y]^2}\,dy \\ g'(k) &= \int_{\mathbb{R}} -2[\pi ik+y]\times \pi i\times e^{-[\pi ik+y]^2}\,dy \\ &= \pi i\int_{\mathbb{R}} -2[\pi ik+y]\times e^{-[\pi ik+y]^2}\,dy \\ &= \pi i\int_{\mathbb{R}} \frac{\partial}{\partial y} e^{-[\pi ik+y]^2}\,dy = \pi i e^{-[\pi ik+y]^2}\Big|_{y=-\infty}^{y=\infty} = 0. \end{aligned}$$

Thus $g(k)$ is a constant function, and is equal to $g(0)$. But this is then exactly the real integral that we know the value of.

(d) We make the coordinate transformation $y = ax$. The Jacobian of this transformation is $|a|^n$. Thus

$$\hat{h}(k) = \int_{\mathbb{R}^n} e^{-2\pi ik\cdot x} g(ax)\,dx = \int_{\mathbb{R}^n} e^{-2\pi ik\cdot(a^{-1}y)} g(y)\,|a|^{-n}dy = |a|^{-n}\hat{g}(a^{-1}k).$$

(e) We make the coordinate transformation $y = x - a$. The Jacobian of this transformation is 1. Thus

$$\hat{h}(k) = \int_{\mathbb{R}^n} e^{-2\pi ik\cdot x} g(x-a)\,dx = \int_{\mathbb{R}^n} e^{-2\pi ik\cdot(y+a)} g(y)\,dy = e^{-2\pi ik\cdot a}\hat{g}(k).$$

(f) Let ϕ_m be a sequence of test functions converging to zero in $\mathscr{S}$. In particular, for $\alpha = 0, l = 0$ we must have $\rho_{0,0}(\phi_m) = \sup|\phi_m| \to 0$. This allows us to conclude that $\delta(\phi_m) = \phi_m(0) \to 0$.

(g) Similar to the previous part, if $\phi_m \to \phi$ in $\mathscr{S}$ then in particular $\|\phi - \phi_m\|_\infty \to 0$ and $\phi_m(0) \to \phi(0)$. This shows that the delta distribution acts on Schwartz functions also by evaluation at 0. Therefore

$$\hat{\delta}(\phi) = \delta(\hat{\phi}) = \hat{\phi}(0) = \int_{\mathbb{R}^n} e^{-2\pi i 0\cdot x}\phi(x)\,dx = \int_{\mathbb{R}^n} 1\,\phi(x)\,dx.$$

In other words, $\hat{\delta} = F_1$.

Another way to deduce this is to use Lemma 4.4, that the Fourier transforms turns a convolution into a product. Then we have for all $u \in \mathscr{S}$

$$\hat{u} = \mathcal{F}(\delta * u) = \hat{\delta}\hat{u}.$$

This is only possible if $\hat{\delta} = 1$.

(h) From the previous part and the inverse Fourier transform, we see that $\hat{1} = \delta$. But let's see what happens if we try this directly, like the question asks.

Firstly we should check that 1 is a tempered distribution. As in an earlier part let ϕ_m be a sequence of test functions converging to zero in $\mathscr{S}$, so we know $\rho_{0,0}(\phi_m) = \sup_x |\phi_m(x)| \to 0$

and $\rho_{n,0}(\phi_m) = \sup_x |x|^{2n}|\phi_m(x)| \to 0$. Hence

$$|F_1(\phi_m)| = \left|\int_{\mathbb{R}^n} 1\,\phi_m(x)\,dx\right| = \left|\left(\int_{B(0,1)} + \int_{\mathbb{R}^n\setminus B(0,1)}\right)\phi_m(x)\,dx\right|$$
$$\leq \rho_{0,0}(\phi_m)\,\omega_n + \int_{\mathbb{R}^n\setminus B(0,1)} |x|^{-2n}|x|^{2n}|\phi_m(x)|\,dx$$
$$\leq \rho_{0,0}(\phi_m)\,\omega_n + \rho_{2n,0}(\phi_m)\int_{\mathbb{R}^n\setminus B(0,1)} |x|^{-2n}\,dx.$$

Because

$$\int_{\mathbb{R}^n\setminus B(0,1)} |x|^{-2n}\,dx = \int_1^\infty r^{-2n}\,n\omega_n r^{n-1}\,dr = n\omega_n\int_1^\infty r^{-n-1}\,dr$$
$$\leq n\omega_n\int_1^\infty r^{-2}\,dr = n\omega_n,$$

we conclude that $|F_1(\phi_m)| \to 0$ as $m \to \infty$. Hence $F_1 \in \mathscr{S}'$.

Now we come to the actual Fourier transform. By definition

$$\widehat{F_1}(\phi) = F_1(\hat{\phi}) = \int_{\mathbb{R}^n}\hat{\phi}(k)\,dk = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{-2\pi i k\cdot x}\phi(x)\,dx\right)\,dk.$$

At this point, we would like to interchange the order of integration, but the inner integral of

$$\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{-2\pi i k\cdot x}\,dk\right)\phi(x)\,dx$$

is not well-defined. We 'hand-waved' in the beginning of Section 4.1 that such integral are zero for $k \neq 0$ because of periodicity, but this is not really a valid argument. And for $k = 0$ the inner integral is ∞. This is reminiscent of the trying to view the delta distribution as a function, it is infinite at zero and vanishes everywhere else, and has a special integration property.

The way to make this a valid argument is to use some sort of mollifier and limit. This is effectively what we did on page 64 of the script when we chose $v = e^{-4\pi^2|k|^2\epsilon}$ and $\hat{v} = \Phi(x,\epsilon)$. Taking the limit $\epsilon \to 0$ gives $v = 1$ and $\hat{v} = \delta$. This was the essential step in the proof of the inverse Fourier transform. Alternatively we could use the multiplication rule of Fourier transforms

$$\hat{u} = \mathcal{F}[u1] = \hat{u} * \hat{1}.$$

This shows that $\hat{1}$ must be the identity element of convolution, ie δ. But we proved this rule too using the inverse Fourier transform. The moral of the story is that the inverse Fourier transform is an unavoidable nexus in the theory, and that the tricky argument with mollification is essential.

32. One step at a time

Prove the following identity for the fundamental solution in one dimension ($n = 1$):

$$\Phi(x, s+t) = \int_{\mathbb{R}} \Phi(x-y, t)\Phi(y, s)\, dy.$$

(2 points)

Hint. You may use without proof that

$$\int_{\mathbb{R}} \exp(-A + By - Cy^2)\, dy = \sqrt{\frac{\pi}{C}} \exp\left(\frac{B^2}{4C} - A\right).$$

Solution.

$$\begin{aligned}
\int_{\mathbb{R}} \Phi(x-y, t)\Phi(y, s)\, dy &= \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi t}\sqrt{4\pi s}} \exp\left(-\frac{(x-y)^2}{4t} - \frac{y^2}{4s}\right) dy \\
&= \frac{1}{\sqrt{4\pi t}\sqrt{4\pi s}} \int_{\mathbb{R}} \exp\left(-\frac{x^2}{4t} + \frac{x}{2t}y - \left[\frac{1}{4t} + \frac{1}{4s}\right] y^2\right) dy \\
&= \frac{1}{\sqrt{4\pi t}\sqrt{4\pi s}} \sqrt{\frac{\pi}{\frac{s+t}{4st}}} \exp\left(\frac{x^2}{4t^2} \cdot \frac{1}{4} \cdot \frac{4st}{s+t} - \frac{x^2}{4t}\right) \\
&= \frac{1}{\sqrt{4\pi(s+t)}} \exp\left(-\frac{x^2}{4(s+t)}\right) = \Phi(x, s+t).
\end{aligned}$$

We know that convolution of the initial condition with the fundamental solution over the space coordinates gives the solution to the initial value problem. The fact that we have $s + t$ in the time suggests we should do this twice. So if we begin with the initial condition $h(x)$ and then solve up to time t, we have

$$u(x, t) = \int_{\mathbb{R}} \Phi(x-y, t)h(y)\, dy.$$

If we take $x \mapsto u(x, t)$ as the start of a new initial value problem and then calculate what the solution is at time s we get

$$\begin{aligned}
v(x, s) &= \int_{\mathbb{R}} \Phi(x-y, s)u(y, t)\, dy \\
&= \int_{\mathbb{R}} \Phi(x-y, s) \int_{\mathbb{R}} \Phi(y-z, t)h(z)\, dz\, dy \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \Phi(x-y, s)\Phi(y-z, t)\, dy\right) h(z)\, dz \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \Phi(x-z-y, s)\Phi(y, t)\, dy\right) h(z)\, dz.
\end{aligned}$$

(We made a shift substitution to get to the last line, but reused the letter y.) On the other hand, if we began with the initial condition $h(x)$ and computed the solution at time $s + t$ we would find

$$u(x, s+t) = \int_{\mathbb{R}} \Phi(x-z, s+t)h(z)\, dz.$$

The relation we just proved says that these two solutions are the same. This shows that the heat equation does not have 'long term memory', it only depends on the immediately prior state, not on anything that happened before that. This is in contrast to its behaviour in space, where the action at one point can immediately affect points infinitely far away. This property is also called the semigroup property because the solution operators $h \mapsto H_t h := \Phi(\cdot, t) * h$ form a semigroup: $H_s H_t = H_{s+t}$.