THE LAPLACE EQUATION IN THE UPPER HALF-PLANE

and other more-or-less related stuff

1. Deriving the Primitive of the Cauchy/Poisson Kernel from Invariance Properties. The idea is to imitate Strauss's §**2.4**, in which he constructs the solution of the initial-value problem for the diffusion equation in $\{(x,t) : x \in \mathbb{R},\ t > 0\}$ by examining invariance properties of the equation, and through this imitation to construct the solution of the boundary-value problem for the Laplace equation $u_{xx}+u_{yy} = 0$ in the "upper half-plane" $\{(x,y) : x \in \mathbb{R},\ y > 0\}$. As we shall see, there are many similarities and some important differences between the two problems.

The *invariance properties* of the Laplace equation in this region are very similar to those of the diffusion equation:

(a) A *translate* $u(x-a,y)$ of any solution $u(x,y)$ of $\nabla^2 u = 0$ is again a solution, for any fixed $a \in \mathbb{R}$;

(b) Any *derivative* of a solution is again a solution;

(c) A *linear combination* of solutions of $\nabla^2 u = 0$ is again a solution (this is just linearity of ∇^2);

(d) An *integral* of solutions is again a solution. Thus if $P(x,y)$ is a solution of the Laplace equation, then so is $P(x-t,y)$ and so is

$$v(x,y) = \int_{-\infty}^{\infty} P(x-t,y)\, g(t)\, dt\ ,$$

so long as this improper integral converges appropriately. Like Strauss, we'll worry about convergence later.

(e) (This is what makes everything go.) If $u(x,y)$ is a solution of the Laplace equation in $\{(x,y) : x \in \mathbb{R},\ y > 0\}$, and $a > 0$ is a constant, then $u(ax,ay)$ is also a solution. This follows from the chain rule: one has $\partial_x^2 u(ax,ay) = a^2 u_{xx}(ax,ay)$ and $\partial_y^2 u(ax,ay) = a^2 u_{yy}(ax,ay)$ and therefore $\nabla^2 u(ax,ay) = a^2 \nabla^2 u(ax,ay)$, so if the r. h. s. of that equals zero because $u(x,y)$ satisfies the Laplace equation in the upper half-plane, then the l. h. s. is also zero and $(\partial_x^2 + \partial_y^2)u(ax,ay) = 0$ there also.

So again, following Strauss, we observe that if we use boundary values on the x-axis—the line $y = 0$—that have the same invariance property that (e) tells us solutions must have, then the solution of the boundary-value problem with those boundary values will also have the invariance property (e), and—as in Strauss's **Step 2** on pp. 48–49—this will force that solution to be a function really of only one variable, for which the equation $\nabla^2 u = 0$ will turn into an ordinary (one-variable) differential equation. Continuing to emulate Strauss, we use the boundary-value function

$$\phi(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \end{cases}$$

(although we left $u(0,0)$ undefined, we shall see that it really wants to be $1/2$) which shares with the function $Q(x,t)$ that Strauss uses[(1)] the property that it is invariant under dilation: $\phi(ax) = \phi(x)$ for any $a > 0$. Thus the solution of $\nabla^2 u = 0$ in the upper half-plane with this $\phi(x)$ as its boundary values on the x-axis should (assuming uniqueness holds) also be invariant under dilation.

At this point we can see what the solution will have to be—by the simple expedient of passing to polar coördinates. If $u(x,y)$ is a function defined in the upper half-plane and $u(x,y)$ is invariant under dilation, then for any $r > 0$ one has in polar coördinates

$$\begin{aligned} u(x,y) &= u(r\cos\theta, r\sin\theta) && \text{(the usual transformation)} \\ &= u(\cos\theta, \sin\theta) = u(\theta)\ , && \text{(Use } r \text{ as "}a\text{")} \end{aligned}$$

(1) In the box on p. 46.

a function of the polar angle θ only. We know[2] that in polar coördinates the Laplace equation becomes

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0$$

but if u depends on θ only, this is just

$$\frac{d^2 u}{d\theta^2} = 0$$

whose general solution is $u(\theta) = a\theta + b$. If $u(0) = 1$ and $u(\pi) = 0$ then $u(\theta) = \frac{2}{\pi}\left[\frac{\pi}{2} - \theta\right] = 1 - \frac{\theta}{\pi}$. Now let's get the same result by Strauss's method (because the diffusion equation has a different invariance property, Strauss could not glibly pass to polar coördinates to solve it, as we have just done with the Laplace equation).

Strauss's argument in **Step 1**, p. 46, is workable in this situation. Because the solution $u(x, y)$ of the boundary-value problem with the boundary-value function $\phi(x)$ we gave above is invariant under dilation, we can write (letting $a = 1/y$)

$$u(x, y) = u\left(\frac{x}{y}, \frac{y}{y}\right) = u\left(\frac{x}{y}, 1\right)$$

for each $y > 0$. Since the l. h. s. is a function of x/y only, we can try to find a function $g(p)$ for which $u(x, y) = g(x/y)$ satisfies the Laplace equation and the boundary value conditions for the upper half-plane. For $u(x, y)$ of that form, we have

$$\begin{aligned}
\partial_x^2[g(x/y)] &= \left(\frac{1}{y}\right)^2 g''(x/y) \\
\partial_y g(x/y) &= \frac{-x}{y^2}\, g'(x/y) \\
\partial_y^2 g(x/y) &= \frac{2x}{y^3}\, g'(x/y) + \left(\frac{-x}{y^2}\right)^2 g''(x/y)
\end{aligned}$$

and the Laplace equation thus becomes

$$\begin{aligned}
0 = \partial_x^2 g(x/y) + \partial_y^2 g(x/y) &= \left(\frac{1}{y}\right)^2 g''(x/y) + \frac{2x}{y^3}\, g'(x/y) + \left(\frac{-x}{y^2}\right)^2 g''(x/y) \\
0 &= \frac{1}{y^2}\left\{\frac{x^2 + y^2}{y^2}\, g''(x/y) + 2\,\frac{x}{y}\, g'(x/y)\right\} \\
0 &= \frac{1}{y^2}\left\{\left(\frac{x^2}{y^2} + 1\right) g''(x/y) + 2\,\frac{x}{y}\, g'(x/y)\right\}
\end{aligned}$$

from which we see that with $p = \frac{x}{y}$, the function $g(p)$ must satisfy the ordinary differential equation

$$(p^2 + 1)g''(p) + 2pg'(p) = 0\ .$$

That equation is easy to solve by separation of variables: with all the "g" on one side and all the "p" on the other, it becomes

$$\begin{aligned}
\frac{g''(p)}{g'(p)} &= \frac{-2p}{p^2 + 1} \\
\log g'(p) &= -\log(p^2 + 1) + C \\
g'(p) &= C_2 \cdot (p^2 + 1)^{-1} = \frac{C_2}{p^2 + 1} \\
g(p) &= C_1 + C_2 \arctan p \\
u(x, y) = g(x/y) &= C_1 + C_2 \arctan\left(\frac{x}{y}\right)\ .
\end{aligned}$$

---

(2) See Strauss, pp. 150–151, or the notes on this subject.

Paralleling Strauss's **Step 3**, for $x > 0$ and $y \to 0^+$ we want this function to have the limit 1; for $x < 0$ and $y \to 0^+$ we want this function to have the limit 0. This gives the equations

$$C_1 + C_2 \cdot \frac{\pi}{2} = 1 \quad \text{and} \quad C_1 + C_2 \cdot \frac{-\pi}{2} = 0$$

so $C_1 = \dfrac{1}{2}$ and $C_2 = \dfrac{1}{\pi}$. Thus $u(x,y) = \dfrac{1}{2} + \dfrac{1}{\pi} \arctan \dfrac{x}{y}$, which agrees with the result of our polar-coördinate argument since $\arctan \dfrac{x}{y} = \dfrac{\pi}{2} - \theta$.

Continuing our emulation of Strauss (in **Step 4**), we should have[(3)] our "$Q(x,y)$" be $\dfrac{1}{2} + \dfrac{1}{\pi} \arctan \dfrac{x}{y}$. However, we shall call our "convolution kernel" by the name $P(x,y)$ rather than $S(x,y)$, because—although it is called the **Cauchy kernel** in certain contexts—it is also the **Poisson kernel for the upper half-plane**. In any event, we'll use

$$P(x,y) = \frac{\partial \left\{ \frac{1}{2} + \frac{1}{\pi} \arctan \frac{x}{y} \right\}}{\partial x} = \frac{1}{\pi} \frac{y}{x^2 + y^2} \; .$$

2. Well, Does it Work? To finish our **Step 4**, we have to form the integral[(4)]

$$u(x,y) = \int_{-\infty}^{\infty} P(x-\xi, y)\phi(\xi)\, d\xi = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{(x-\xi)^2 + y^2} \phi(\xi)\, d\xi$$

and see if it really is the solution of $\nabla^2 u = 0$ in the upper half-plane with boundary values $\phi(x)$ on the x-axis. Some hypotheses on $\phi(x)$ are going to be needed. Following Strauss, we integrate by parts, with the rôle of "u" played by $\phi(x)$ and that of "dv" by the Cauchy kernel, and we assume that $\phi(x) \to 0$ as $|x| \to \infty$:

$$\begin{aligned} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{(x-\xi)^2 + y^2} \phi(\xi)\, d\xi &= - \left[\frac{1}{2} + \frac{1}{\pi} \arctan \frac{x-\xi}{y} \right] \phi(\xi) \Bigg]_{\xi \to -\infty}^{\xi \to \infty} + \int_{-\infty}^{\infty} \left[\frac{1}{2} + \frac{1}{\pi} \arctan \frac{x-\xi}{y} \right] \phi'(\xi)\, d\xi \\ &= \int_{-\infty}^{\infty} \left[\frac{1}{2} + \frac{1}{\pi} \arctan \frac{x-\xi}{y} \right] \phi'(\xi)\, d\xi \; . \end{aligned}$$

We break that integral at x and take the limit as $y \to 0^+$ under the integral sign:

$$\lim_{y \to 0^+} \int_{x}^{\infty} \left[\frac{1}{2} + \frac{1}{\pi} \arctan \frac{x-\xi}{y} \right] \phi'(\xi)\, d\xi = \int_{x}^{\infty} 0 \cdot \phi'(\xi)\, d\xi = 0$$
$$\lim_{y \to 0^+} \int_{-\infty}^{x} \left[\frac{1}{2} + \frac{1}{\pi} \arctan \frac{x-\xi}{y} \right] \phi'(\xi)\, d\xi = \int_{-\infty}^{x} 1 \cdot \phi'(\xi)\, d\xi = \lim_{\xi \to -\infty} [\phi(x) - \phi(\xi)] = \phi(x) \; .$$

So putting the two things together gives us

$$\lim_{y \to 0^+} u(x,y) = \lim_{y \to 0^+} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{(x-\xi)^2 + y^2} \phi(\xi)\, d\xi = \phi(x)$$

which is what we would have hoped for.

---

(3) Actually we could have left off the constant 1/2, since $Q(x,y)$ exists only to be differentiated anyway. This would have made the end of the argument showing that the Cauchy kernel does what we want it to do a little different.

(4) There is a change of notation here, necessitated by the fact that we are thinking of functions of (x,y)—one hopes that it won't be too confusing. Strauss's functions have arguments (x,t), so he has y available as a dummy variable of integration. Here we are using the letter y as the second space coördinate and it is not available for use as a dummy variable. We think of our function as $u(x,y)$, so we use ξ as a dummy variable of integration (this "Greek 'x,' we hope, reminds us that this is integration in one of the "space variables," even though ξ is only a dummy variable of integration).

3. Sigh—More Hypotheses. Strauss's treatment of the diffusion equation is that of an initial-value problem. Although he chooses to assume that his initial-value function $\phi(x)$ equals zero for $|x|$ large, that is really an overkill hypothesis: it suffices that $\phi(x)$ fall off to zero fast enough as $|x| \to \infty$ that the energy proof of uniqueness on pp. 42–43 makes sense, so that one needs to know that the integrals $\int_{-\infty}^{\infty} \phi(x)^2\,dx < \infty$ and $\int_{-\infty}^{\infty} \phi'(x)^2\,dx < \infty$. This (and a few other considerations that Lebesgue integration theory disposes of efficiently) makes the energy proof just as valid on the "infinite interval" $(-\infty, \infty)$ as on the finite interval $[0, \ell]$. We have been calling the corresponding problem for the Laplace equation a boundary-value problem, because there are many points of view from which that is what it is. Unfortunately, the simple example of the function $u(x, y) = y$—which certainly satisfies $\nabla^2 y = 0$—shows that there are solutions of the Laplace equation in the upper half-plane that are zero on the boundary (= the x-axis) without being identically zero. To get uniqueness for the Laplace equation, we must thus add some kind of "growth condition as $\sqrt{x^2 + y^2} \to \infty$." For what we're doing now, let us just add the requirement that $u(x, y) \to 0$ as $\sqrt{x^2 + y^2} \to \infty$. That excludes our counterexample $u(x, y) = y$, and it is easy to prove a uniqueness theorem for these functions by the maximum principle: if $u(x, y)$ is continuous on $\{(x, y) : x \in \mathbb{R},\ y \ge 0\}$ and $\nabla^2 u = 0$ in $\{(x, y) : x \in \mathbb{R},\ y > 0\}$, and if also $u(x, y) \to 0$ as $\sqrt{x^2 + y^2} \to \infty$, then the condition $u(x, 0) \equiv 0$ forces $u(x, y) \equiv 0$ throughout the upper half-plane. The reason is that given any $\epsilon > 0$ we can find an $R > 0$ such that $\sqrt{x^2 + y^2} \ge R$ implies $|u(x, y)| < \epsilon$. If $a \ge R$, then $|u(x, y)| < \epsilon$ on the boundary of the "half-disc"$\{(x, y) : x \in \mathbb{R},\ y \ge 0,\ \sqrt{x^2 + y^2} \le a\}$, because it is zero on the x-axis and smaller than ϵ on the semicircle $x^2 + y^2 = a^2,\ y \ge 0$. But then, by the maximum principle, $|u(x, y)| < \epsilon$ throughout the "half-disc"$\{(x, y) : x \in \mathbb{R},\ y \ge 0,\ \sqrt{x^2 + y^2} \le a\}$. Because $a > R$ could be arbitrarily large, $|u(x, y)| < \epsilon$ throughout the upper half-plane $\{(x, y) : x \in \mathbb{R},\ y \ge 0\}$. Because $\epsilon > 0$ could be arbitrarily small, $u(x, y) = 0$ must hold at every point of the upper half-plane. That shows "zero boundary values on the x-axis" implies that the function is identically zero; so, in the presence of the limit-zero-at-infinity condition, we have uniqueness of solutions (consider the difference of two competing solutions, as usual).

4. Convolution and Convolution Kernels. These general solutions for both the diffusion and the Laplace equations have an extremely important form. If $K(x)$ is a function defined on the real line for which $\displaystyle\int_{-\infty}^{\infty} |K(x)|\,dx < \infty$, then for a large variety of other functions $\phi(x)$ the integral[5]

$$(K * \phi)(x) \int_{-\infty}^{\infty} K(x - \xi)\phi(\xi)\,d\xi$$

will make sense, and will usually have all the nice properties (continuity, differentiability, etc.) that either of $K(x)$ and $\phi(x)$ possesses. This function is routinely called $K * \phi$ and called the **convolution of K and ϕ**. The order is not important: a simple change of variable[6] shows that $K * \phi = \phi * K$. Usually one will want to think of K as fixed and ϕ as being open to many possibilities: in these situations one calls K the **convolution kernel** and $\phi \mapsto K * \phi$ gives us the **convolution transform** of ϕ (by K, if there might be some ambiguity).

Both the solution formulas we have found for the diffusion equation

$$u(x, t) = \frac{1}{\sqrt{4\pi k t}} \int_{-\infty}^{\infty} e^{-(x-\xi)^2/4kt}\,\phi(\xi)\,d\xi$$

and the Laplace equation

$$u(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{(x - \xi)^2 + y^2}\,\phi(\xi)\,d\xi$$

are obviously in the form of convolutions. Much of what makes these formulas so easy to work with is due to that fact.

---

(5) It is possible—in some cases even necessary and desirable—to finesse the requirement that convolution kernels have a finite integral. The problem is that then sometimes one only has a "symbolic integral" and considerable special pleading is necessary to make the formalism meaningful. We shall try to avoid considerations of that type in this course.

(6) Which the reader should undertake to find . . .

5. The Semi-Group Property. This is important for a number of reasons, but Strauss does not talk about it—let us look. In both cases we know about—the diffusion equation and the Laplace equation—the convolution kernels that solve the equations have the property

$$K(x,t_1) * K(x,t_2) = K(x,t_1+t_2) \ \text{ for diffusion; for Laplace } \quad K(x,y_1) * K(x,y_2) = K(x,y_1+y_2)\ .$$

This "additivity" property, which is called the **semi-group property**, is hard to prove directly[(7)] but flows right out of the differential equations. Consider first the Cauchy kernel

$$P(x,y) = \frac{1}{\pi}\,\frac{y}{x^2+y^2}\ .$$

For a fixed value of $y = y_2 > 0$, this is a perfectly good function to use as the boundary-value function for the Laplace equation, *i.e.*, we can take

$$\phi(x) = \frac{1}{\pi}\,\frac{y_2}{x^2+y_2^2}\ .$$

Now we have two ways to solve the boundary-value problem for the Laplace equation with these boundary values: we can convolve with the Poisson kernel and get

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{(x-\xi)^2+y^2}\,\frac{1}{\pi}\,\frac{y_2}{\xi^2+y_2^2}\,d\xi$$

or we can just look at the function itself:

$$u(x,y) = \frac{1}{\pi}\,\frac{y+y_2}{x^2+(y+y_2)^2}$$

is a solution of the Laplace equation in the upper half-plane and takes the right boundary values on the x-axis (when $y=0$, we get $\frac{1}{\pi}\frac{y_2}{x^2+y_2^2}$ back). By uniqueness, which we will have because these functions clearly have limit zero as $\sqrt{x^2+y^2}\to\infty$,

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{(x-\xi)^2+y^2}\,\frac{1}{\pi}\,\frac{y_2}{\xi^2+y_2^2}\,d\xi = \frac{1}{\pi}\,\frac{y+y_2}{x^2+(y+y_2)^2} \qquad \text{for all } y>0.$$

Let $y = y_1$ in this relation and you get a relation valid for all $y_1,\ y_2 > 0$:

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y_1}{(x-\xi)^2+y_1^2}\,\frac{1}{\pi}\,\frac{y_2}{\xi^2+y_2^2}\,d\xi = \frac{1}{\pi}\,\frac{y_1+y_2}{x^2+(y_1+y_2)^2}$$

$$P(x,y_1) * P(x,y_2) = P(x,y_1+y_2)\ .$$

It is possible to establish this relation directly, but who would want to? Similarly, consider the diffusion kernel

$$S(x,t) = \frac{1}{\sqrt{4\pi kt}}\,e^{-x^2/4kt}\ .$$

For a fixed value of $t = t_2 > 0$, this is a perfectly good function to take as the initial-value function for the diffusion equation; *i.e.*, we can take

$$\phi(x) = \frac{1}{\sqrt{4\pi kt_2}}\,e^{-x^2/4kt_2}$$

as giving the values of a solution of the diffusion equation at time $t=0$. Now we have two ways to solve the initial-value problem for the diffusion equation with these boundary values: we can convolve with the diffusion kernel and get

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi kt}}\,e^{-(x-\xi)^2/4kt}\,\frac{1}{\sqrt{4\pi kt_2}}\,e^{-\xi^2/4kt_2}\,d\xi$$

(7) Try establishing it directly if you enjoy really unpleasant integration problems.

or we can just look at the function itself:

$$u(x,t) = \frac{1}{\sqrt{4\pi k(t+t_2)}}\, e^{-x^2/[4k(t+t_2)]}$$

is a solution of the diffusion equation in the upper half-plane and takes the right boundary values on the x-axis (when $t=0$, we get $\dfrac{1}{\sqrt{4\pi k t_2}}\, e^{-x^2/4kt_2}$ back). By uniqueness, which we will have because these functions clearly have square-integrable values and square-integrable derivatives (so that energy proofs of uniqueness are valid),

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi kt}}\, e^{-(x-\xi)^2/4kt}\, \frac{1}{\sqrt{4\pi kt_2}}\, e^{-\xi^2/4kt_2}\, d\xi = \frac{1}{\sqrt{4\pi k(t+t_2)}}\, e^{-x^2/[4k(t+t_2)]}\ .$$

Let $t = t_1$ in this relation and you get a relation valid for all t_1, $t_2 > 0$:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi kt_1}}\, e^{-(x-\xi)^2/4kt_1}\, \frac{1}{\sqrt{4\pi kt_2}}\, e^{-\xi^2/4kt_2}\, d\xi = \frac{1}{\sqrt{4\pi k(t_1+t_2)}}\, e^{-x^2/[4k(t+t_2)]}$$

$$S(x,t_1) * S(x,t_2) = S(x,t_1+t_2)\ .$$

6. Connections with Probability Theory. This § is not required material for the course, because there is no probability prerequisite for the course; however, some of these considerations may be helpful for some persons. At the time these notes are written the textbook for the basic probability course at Rutgers–NB is Sheldon Ross, *A First Course in Probability*, 5th ed., Prentice-Hall (1997); references to it are given below for the reader's convenience.

A basic theorem about the probability distributions of random variables is the fact that if X and Y are independent random variables and Y is continuously distributed with probability density function f_Y, then the cumulative distribution function of $X+Y$—and, if X is also continuously distributed, the probability density function of $X+Y$—are given by convolution integrals:(8)

$$F_{X+Y}(a) = \int_{-\infty}^{\infty} F_X(a-y)\, f_Y(y)\, dy \quad \text{and}$$

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)\, f_Y(y)\, dy \quad \text{respectively.}$$

For each $t>0$ and $y>0$ respectively, each of the functions

$$x \mapsto S(x,t) = \frac{1}{\sqrt{4\pi kt}}\, e^{-x^2/4kt}$$

$$x \mapsto P(x,y) = \frac{1}{\pi}\, \frac{y}{x^2+y^2}$$

is a nonnegative function with integral 1, so each can be interpreted as a probability density function. The function $S(x,t)$ is particularly familiar: it is the p. d. f. of a normal random variable(9) with mean zero and standard deviation $\sigma = \sqrt{2kt}$. Similarly, the function $P(x,y)$ is well known in the context of statistics, although it may be less familiar: it is the p. d. f. of a random variable with a Cauchy distribution,(10) which explains one of its names ("Poisson kernel" is not chosen to have a connection with the [discrete] Poisson distributions of probability theory, however). In the case in which the initial-value function $\phi(x)$ (for the diffusion equation) or boundary-value function $\phi(x)$ (for the Laplace equation) is a probability density

(8) Ross, pp. 264 ff.

(9) Ross, p. 204 ff.

(10) Ross, p. 304, #33: this is the distribution of a quotient of two independent normal random variables of mean zero; "y" depends on the ratio of their variances.

function—which is simply the requirement that $\phi(x) \geq 0$ and $\displaystyle\int_{-\infty}^{\infty} \phi(x)\,dx = 1$—so that one can think of it[11] as the p. d. f. of a random variable Φ, the solution of the corresponding equation at time t or "height" y, *i.e.*,

$$u(x,t) = \int_{-\infty}^{\infty} S(x-\xi, y)\phi(\xi)\,d\xi = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{-(x-\xi)^2/4kt}\,\phi(\xi)\,d\xi \quad \text{or}$$

$$u(x,y) = \int_{-\infty}^{\infty} P(x-\xi, y)\phi(\xi)\,d\xi = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y}{(x-\xi)^2 + y^2}\,\phi(\xi)\,d\xi \quad \text{respectively,}$$

can thus be interpreted as the probability density function of a random variable $\Phi + N_t$ or $\Phi + Y_y$, where N_t is independent of Φ and normally distributed with mean zero and $\sigma = \sqrt{2kt}$, or Y_y is independent of Φ and Cauchy-distributed with parameter $y > 0$. What one sees as t or y becomes large, then—and this is particularly striking with the diffusion kernel—is that the original distribution of Φ becomes increasingly "corrupted" or "spread out" by having an independent, very highly spread out normally- or Cauchy-distributed random variable added to it. On the other hand, as $t \to 0^+$ or $y \to 0^+$ the "spread" of N_t or Y_y shrinks down to zero, so that the initial distribution ϕ of Φ is attained.[12] Indeed, if we refer back to the derivations of the diffusion and Poisson kernels respectively, we see that their antiderivatives, *i.e.*, the cumulative distribution functions of the corresponding random variables—which are respectively

$$Q(x,t) = \frac{1}{2} + \frac{1}{2}\,\mathcal{E}rf\left(\frac{x}{\sqrt{4kt}}\right) \quad \text{and}$$

$$Q(x,y) = \frac{1}{2} + \frac{1}{\pi}\,\arctan\frac{x}{y}$$

—each converge to the c. d. f. of a "nonrandom variable with value 0," *i.e.*, the function that is 0 for $x < 0$ and 1 for $x > 0$, as $t \to 0^+$ or $y \to 0^+$ respectively. Thus each of the kernels can be viewed as giving the probability distribution of the location of some "randomly placed object" that is placed at the origin at $t = 0$ (or $y = 0$) but that at time $t > 0$ (or when $y > 0$) is to be found somewhere on the x-axis, where "somewhere" has the interpretation of "randomly located with probability distribution $S(x,t)$ (or $P(x,y)$)."

This approach may help to answer objections of the following kind to the diffusion equation: "How can a dye which at $t = 0$ is contained in the interval $[-1/2, 1/2]$ in an 'infinitely long pipe' suddenly find itself everywhere in the pipe at time $t > 0$, no matter how small t is?" Diffusion processes are fundamentally probabilistic: there are simply too many molecules of dye running around to treat each molecule individually. Thus when one considers the diffusion process with initial-value function $\phi(x) = \begin{cases} 1 & \text{if } -1/2 \leq x \leq 1/2 \\ 0 & \text{otherwise} \end{cases}$, one is really saying: if the probability of finding dye molecules at time $t = 0$ is uniformly distributed on $[-1/2, 1/2]$—meaning that the probability of finding molecules is proportional to the length of the interval of $[-1/2, 1/2]$ in which one looks for them, and is zero outside that interval—then how is the probability of finding them distributed on the x-axis when $t > 0$? The answer is given by Strauss's exercise §**2.4, #1** with $\ell = 1/2$, namely by

$$u(x,t) = \mathcal{E}rf\left(\frac{x+1/2}{\sqrt{4kt}}\right) - \mathcal{E}rf\left(\frac{x-1/2}{\sqrt{4kt}}\right)\,.$$

This function is nonzero everywhere on the x-axis for all $t > 0$. That does not mean that individual molecules of dye will have instantaneously migrated out to all points of the x-axis, no matter how small t is. It means, rather,[13] that if one were able to perform repeated, independent experiments with dye diffusion and able

(11) I am just using the letter Φ here so as to match the name of the function ϕ; it should not be construed as referring to the c. d. f. of a unit normal random variable!

(12) This statement is quite precise in the case of the diffusion kernel, which has finite moments of all orders. It is less so in the case of the Cauchy distribution: the random variable $|P|$ does not even have finite expectation (try computing it!), much less finite variance, so a quantitative description of how it spreads out is much harder to give.

(13) This discussion unabashedly takes the "relative frequency" viewpoint of probability, which is to some extent vindicated by the strong law of large numbers.

to measure the amount of dye found in a small interval $[x, x + \Delta x]$ of the x-axis at time $t > 0$, then on the average one would find that the experimental probability of finding dye in that interval was about $u(x,t)\cdot\Delta x$. In other words, the diffusion equation gives statistical information, not molecule-by-molecule deterministic information.(14)

Note also, by the way, how beautifully the semigroup property of convolution with the diffusion and Poisson kernels fits in with the probabilistic interpretation. Let's look at diffusion. If Φ is an initial probability distribution with p. d. f. $\phi(x)$, then at time $t_1 > 0$ the solution $u(x,t)$ of the diffusion equation with $u(x,0) = \phi(x)$ and $t = t_1$ is the p. d. f. of $\Phi + N_{t_1}$, where N_{t_1} is an independent normal random variable with mean zero and variance $\sigma = \sqrt{2kt_1}$. If we allow time t_2 to elapse from that time, then the semigroup property of the diffusion kernel says that we have our choice of looking at the random variable telling us the probability distribution at time $t_1 + t_2$ in either of the forms

$$\Phi + N_{t_1} + N_{t_2} \quad \text{or} \quad \Phi + N_{(t_1+t_2)} \ ,$$

where—perhaps surprisingly—N_{t_1} and N_{t_2} are independent of each other and of Φ, and of course $N_{(t_1+t_2)}$ is independent of Φ. In other words, to find the probability distribution at time $t_1 + t_2$ one can simply take the probability distribution in the state to which it has evolved by the end of the time interval $[0, t_1]$—without worrying about how it got there—and just allow that to evolve in the time interval $[0, t_2]$. The new random variable—whether we look at it as $\Phi + N_{t_1} + N_{t_2}$ or as $\Phi + N_{t_1+t_2}$—will have the same probability distribution.(15) This ability to "start anew" tells us that diffusion is really being governed by the probabilistic construct called a **Markov process**, a discussion of which is totally outside the scope of this course.(16)

Similar considerations are available for the Poisson kernel, although the fact that it does not have finite absolute expectation introduces some difficulty in conceptualization.

---

(14) Points of view like this one should be familiar to anyone who has dealt with the Schrödinger equation. But please notice that we can't get too relentlessly physical about what this equation means. One may argue that if c denotes the speed of light, then at time t one cannot find diffused molecules out farther than $1/2 + ct$ units from the origin. Ok, but what do you mean by "t" at that point on the axis? You can't get news of time zero out to that point at speeds faster than c either. We're not doing physics here; we're trying to develop intuition for solutions of the diffusion equation. The mathematics of the diffusion equation stands by itself, irrespective of (quasi-)physical interpretations that might be placed upon it.

(15) Of course, if you think of the random process as physical, then "individual molecules" might wind up in different places. But if you look at it that way, then—as the probabilists say—you are looking at a particular sample path of a Markov process, rather than at the probability distribution that you would have at that time.

(16) Rudiments of the theory of continuous-time Markov processes are discussed in the second-semester probability course at Rutgers–NB, Mathematics 478.