THE DIFFUSION EQUATION IN HIGHER DIMENSIONS

1. Derivation. Consider a continuous material medium diffusing in a region Ω in space: a gas, an ink blot, a bacterial population. One way to describe its spatial distribution at any given time is via a *density* or *concentration* $u(x,t)$, $x \in \Omega$, defined so that the total amount (say, mass) $M_D(t)$ of material contained in a sub-region $D \subset \Omega$ at time t is the volume integral of u over D:

$$M_D(t) = \int_D u(x,t)dV.$$

(This will be our notation for multiple integrals, with respect to volume in $\mathbb{R}^3$ or area in $\mathbb{R}^2$.) If there are no 'sources' of the material in Ω (or 'sinks' where material disappears), the only way $M_D(t)$ can change is if there is a net flow of material through the boundary ∂D. This is usually expressed as a *conservation law*: there is a 'current' vector field (representing the direction and intensity of diffusion of the material) $\vec{q}(x,t)$, so that the net flow of material across the boundary per unit time is given by its flux across the boundary:

$$\frac{dM_D}{dt} = -\oint_{\partial D} \vec{q} \cdot \vec{n} dA.$$

(This is our notation for integration on the boundary of a region with respect to boundary area, or arc length for domains in the plane; $\vec{n}$ denotes the outward unit normal of ∂D.) It makes sense that the flow rate across ∂D depends on the dot product with the normal: at points where $\vec{q} \cdot \vec{n} = 0$, the material is moving tangentially to the boundary, and this movement does not contribute to the change in $M_D(t)$. Also, note that $\vec{q}$ measures the *outward* flow rate, which contributes to a *decrease* in M_D; hence the minus sign. Thus, we have:

$$\frac{d}{dt}\int_D u(x,t)dV = -\oint_{\partial D} \vec{q} \cdot \vec{n} dA \text{ (conservation law, integral form).}$$

Recall the *divergence theorem* states the flux of a vector field across the boundary of a region equals the volume integral of its divergence over the region:

$$\oint_{\partial D} \vec{q} \cdot \vec{n} dA = \int_D div(\vec{q}) dV,$$

where, by definition:

$$\vec{q} = (q_1, q_2, q_3) \Rightarrow div(\vec{q}) := (q_1)_x + (q_2)_y + (q_3)_z$$

(in rectangular coordinates in $\mathbb{R}^3$). Combining the last two equations we find:

$$\int_D u_t dV = -\int_D div(\vec{q}) dV,$$

and since this holds for arbitrary regions $D \subset \Omega$, we must have at each point:

$$u_t = -div(\vec{q}). \quad \text{(conservation law, differential form)}$$

The second physical principle entering the derivation relates the intensity and direction of diffusion to local variations in the concentration, which is measured by the gradient of u, a vector field in Ω we denote by ∇u. It is a generally observed fact that the material will diffuse from regions of higher to those of lower concentration, so it is reasonable to assume $\vec{q}$ will follow the direction of fastest *decrease* in concentration, namely $-\nabla u$. The simplest model is to assume these two vector fields are proportional:

$$\vec{q}(x,t) = -k\nabla u(x,t). \text{ ('Fick's law of diffusion')}$$

Combining the conservation law and Fick's law, we arrive at the P.D.E. for $u(x,t)$:

$$u_t = k div(\nabla u).$$

The operator $div(\nabla u)$ is ubiquitous in physical models, and will be denoted by Δu (the Laplacian of u); it is just the sum of second partial derivatives (in rectangular coordinates):

$$\Delta u = div(\nabla u) = u_{xx} + u_{yy} + u_{zz}.$$

Thus we have the final form of the diffusion equation (or 'heat equation', if u represents the temperature of the medium:

$$u_t = k\Delta u,$$

where $k > 0$ is a constant. In fact k can be set to one simply by rescaling the time variable; mathematically it is enough to consider the case $k = 1$, and we'll usually do that.

Remark. A more general model than 'Fick's Law' would assume the current vector $\vec{q}(x,t)$ is a linear function of the partial derivatives of u; say, given by a symmetric matrix K with entries depending on (x,t):

$$q_i(x,t) = -\sum_j k_{ij}(x,t) u_{x_j}(x,t), \quad K(x,t) > 0.$$

(Recall a symmetric matrix K is said to be positive if all its eigenvalues are positive.) This would lead to the general diffusion equation:

$$u_t = \sum_{i,j} \partial_{x_i}(k_{ij} u_{x_j}), \quad K(x,t) > 0.$$

(This would correspond to 'directional inhomogeneities' (anisotropy) in the physical processes driving the diffusion.) This model (a 'parabolic equation in $\mathbb{R}^n$ ') is well-understood, and as long as the coefficients are smooth (and the eigenvalues of K are bounded away from zero) is qualitatively similar to the standard diffusion equation.

2. The Heat Kernel and the initial-value problem in $\mathbb{R}^n$.

2.1. Product solutions. The diffusion equation has a remarkable property: products of one-variable solutions are solutions of the equation in $\mathbb{R}^n$! For instance, in $\mathbb{R}^2$ with coordinates (x_1, x_2): consider two solutions $v(x_1,t), w(x_2,t)$ of the standard one-variable heat equation (in different variables):

$$v_t = v_{x_1 x_1}, \quad w_t = w_{x_2 x_2}.$$

Then the Laplacian of the product $u(x_1, x_2, t) = v(x_1,t)w(x_2,t)$ is easily found to be:

$$\Delta u = v_{x_1 x_1} w + v w_{x_2 x_2},$$

and since also $u_t = v_t w + v w_t$, we find that:

$$u_t - \Delta u = (v_t - v_{x_1 x_1})w + v(w_t - w_{x_2 x_2}) = 0,$$

so $u(x,t)$ is a solution of the heat equation in $\mathbb{R}^2$. (Note this depends on the fact that the constants k involved in all three equations are all one.)

Consequence: Earlier we saw that a polynomial function in one variable (at $t = 0$) evolves, under the one-dimensional standard heat equation, to a polynomial function of x (of the same degree) with coefficients depending on t.

Now, polynomial functions of several variables are linear combinations of products of one-variable monomials (for instance, in $\mathbb{R}^2$: linear combinations of terms of the form $x_1^r x_2^s$, where $r, s \geq 0$ are natural numbers.) It follows that if $u_0(x), x \in \mathbb{R}^n$ is a polynomial function of $x \in \mathbb{R}^n$, the solution $u(x,t)$ of the heat equation with initial data u_0 will again be a polynomial function in several variables, with coefficients depending on t. Say, in two variables:

$$u_0(x) = \sum_{r,s \in \mathbb{N}} c_{rs} x_1^r x_2^s.$$

$$v_0(x_1) = x_1^r \Rightarrow v(x_1,t) = \sum_{i=0}^{r} a_i(t) x_1^i,$$

$$w_0(x_2) = x_2^s \Rightarrow w(x_2,t) = \sum_{j=0}^{s} b_j(t) x_2^j.$$

By linearity and uniqueness, it follows the solution $u(x,t)$ with initial data $u_0(x)$ is:

$$u(x,t) = \sum_{r,s\in\mathbb{N}} c_{rs} \sum_{i=0}^{r}\sum_{j=0}^{s} a_i(t) b_j(t) x_1^i x_2^j,$$

where $a_r(0) = 1, a_i(0) = 0$ for $i \neq r$; $b_s(0) = 1, b_j(0) = 0$ for $j \neq s$.

This shows the solution is a polynomial (with time-dependent coefficients). In practice it is easier to find the solution directly; for example, in the two-dimensional case assume the solution takes the form:

$$u(x,t) = \sum_{j=0}^{d} \sum_{r,s\geq 0, r+s=j} c_{rs}^j(t) x_1^r x_2^s.$$

The heat equation implies a system of ordinary differential equations for the coefficients $c_{rs}^j(t)$, with initial conditions given by the initial data.

Exercise 1. Find the solution $u(p,t)$ of the heat equation $u_t - \Delta u = 0$ in $\mathbb{R}^2$ (with coordinates $p = (x,y)$ with initial condition $u_0(p) = xy^2, p \in R^2$.

Exercise 2. We also saw earlier that, under the one-dimensional heat equation, exponential functions $e^{a_1 x_1}$ evolve as time-dependent exponentials:

$$u_0(x_1) = e^{a_1 x_1} \Rightarrow u(x_1,t) = e^{a_1^2 t + a_1 x_1}.$$

Indeed this is even true for *complex* exponentials! That is:

$$u_0(x_1) = e^{i b_1 x_1} \Rightarrow u(x_1,t) = e^{-b_1^2 t + i b_1 x_1}.$$

Part(i) *Check this.* Recall that $e^{i b_1 x_1} = \cos(b_1 x_1) + i \sin(b_1 x_1)$. We define a complex-valued function to be $u(x,t)$ a solution of the heat equation if its real and imaginary parts are.

What is behind this is the fact that both real and imaginary exponentials are eigenfunctions of the operator 'minus the second derivative', $L[f] = -f_{x_1 x_1}$:

$$f(x_1) = e^{a x_1} \Rightarrow L[f] = -a_1^2 f,$$

$$f(x_1) = e^{ib_1 x_1} \Rightarrow L[f] = b_1^2 f$$

(meaning, in the second case: the real and imaginary parts of f are eigenfunctions, with the same eigenvalue.) The n-dimensional analogue of this is the following: for a function $f(x)$ in $\mathbb{R}^n$, consider the differential operator L given by 'minus the Laplacian':

$$L[f] = -\Delta f.$$

Part (ii). Show that real and complex exponential functions in $\mathbb{R}^n$ (defined by vectors $a, b \in \mathbb{R}^n$ via the dot product) are eigenfunctions of L:

$$f(x) = e^{a \cdot x} \Rightarrow L[f] = -|a|^2 f,$$

$$f(x) = e^{ib \cdot x} \Rightarrow L[f] = |b|^2 f.$$

Part (iii) Show that this implies exponential initial data (real or complex) yield exponential solutions, for the standard heat equation $u_t - \Delta u = 0$ in $\mathbb{R}^n$:

$$u_0(x) = e^{a \cdot x} \Rightarrow u(x,t) = e^{|a|^2 t + a \cdot x},$$

$$u_0(x) = e^{ib \cdot x} \Rightarrow u(x,t) = e^{-|b|^2 t + ib \cdot x},$$

or taking real and imaginary parts:

$$u_0(x) = \cos(b \cdot x) \Rightarrow u(x,t) = e^{-|b|^2 t} \cos(b \cdot x),$$

and similarly for $\sin(b \cdot x)$.

Thus exponential functions in $\mathbb{R}^n$ (with real or pure imaginary exponents) evolve under the standard heat equation by multiplication by an exponential function of t.

2.2 The heat kernel in $\mathbb{R}^n$. Recall the one-dimensional heat kernel in the variable x_1:

$$h(x_1, t) = \frac{1}{\sqrt{4\pi t}} e^{-x_1^2/4t}$$

is a solution of the heat equation $h_t - h_{x_1 x_1} = 0$. From the preceding discussion, it follows that the product of n one-variable copies of the one-dimensional heat kernel (in different variables):

$$p(x,t) = h(x_1,t)h(x_2,t)\dots h(x_n,t) \quad x = (x_1, \dots, x_n), t > 0,$$

will be a solution of the n-dimensional standard heat equation: $p_t - \Delta p = 0$. Recalling that the length squared of a vector $x \in \mathbb{R}^n$ is $|x|^2 = x_1^2 + \ldots + x_n^2$, we may write $h(x,t)$ in the simpler form:

$$p(x,t) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}, \quad x \in \mathbb{R}^n, t > 0.$$

In addition to being a solution, the one-dimensional heat kernel has the normalization property:

$$\int_{-\infty}^{\infty} h(x_1,t)dx_1 = 1, \forall t > 0.$$

Exercise 3. Show that this implies the normalization property for the heat kernel in $\mathbb{R}^n$:

$$\int_{\mathbb{R}^n} p(x,t)dV = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1,\ldots,x_n,t)dx_1 \ldots dx_n = 1.$$

In addition to being a positive solution of the heat equation on $\mathbb{R}$, with finite (normalized) integral over the real line, the one-dimensional heat kernel $h(x_1,t)$ has the crucial property: its 'mass' (or integral, or area under the curve) is increasingly concentrated near the origin, as we let t approach zero. Equivalently, for any $R > 0$ the 'area under the curve' for the 'tail part' $|x| > R$ tends to zero as $t \to 0_+$. The n-dimensional heat kernel also has this property: for any $R > 0$, the total integral outside of a ball of side lradius R (centered at the origin) tends to zero:

$$\forall R > 0 \lim_{t \to 0_+} \int_{|x|>R} p(x,t)dV = 0.$$

To see this, compute in polar doordinates (r,ω) in $\mathbb{R}^n$ (where $r = |x| > 0$ is distance to the origin, and $\omega = \frac{x}{|x|}$ is a point on the unit sphere S^{n-1} ($|\omega| = 1$); so $x = r\omega$). Recall:

$$dV = r^{n-1} d\omega dr,$$

where $d\omega$ is the 'element of area' on the unit sphere S^{n-1}. In more familiar notation: if $n = 2$, $d\omega$ is just the 'element of angle' (in radians) on the unit circle S^1 (often denoted by $d\theta$):

$$\int_{S^1} d\omega = 2\pi, \quad dV = r d\omega$$

In standard spherical coordinates (r, ϕ, θ) in $\mathbb{R}^3$, $\omega = (\phi, \theta)$, where $\phi \in [0, \pi], \theta \in [0, 2\pi]$:

$$d\omega = \sin^2 \phi d\phi d\theta, \int_{S^2} d\omega = \int_0^{2\pi} \int_0^{\pi} \sin^2 \phi d\phi d\theta = 4\pi, dV = r^2 d\omega.$$

Thus the integral in question is:

$$\int_{|x|>R} p(x,t)dV = \frac{1}{(4\pi t)^{n/2}} \int_R^{\infty} \int_{S^{n-1}} e^{-r^2/4t} r^{n-1} d\omega dr$$

$$= \frac{A(S^{n-1})}{(4\pi t)^{n/2}} \int_R^{\infty} e^{-r^2/4t} r^{n-1} dr,$$

where $A(S^{n-1}) = \int_{S^{n-1}} d\omega$, the 'area of the unit sphere in $\mathbb{R}^n$', depends only on n, for instance: $A(S^1) = 2\pi, A(S^2) = 4\pi$. (And $A(S^0) = 2$!)

In this one-dimensional integral, make the change of variable $s = \frac{r}{\sqrt{4t}}$ to find:

$$\int_{|x|>R} p(x,t)dV = \frac{A(S^{n-1})}{\pi^{n/2}} \int_{\frac{R}{\sqrt{4t}}}^{\infty} e^{-s^2} s^{n-1} ds.$$

One way to express this is the following. Define the 'n-dimensional error function' $erf_n(\rho)$ of $\rho > 0$ as the normalized integral of $e^{-|y|^2}$ over the ball $\{y \in \mathbb{R}^n; |y| < \rho\}$. That is, we define:

$$erf_n(\rho) = \frac{1}{\pi^{n/2}} \int_{\{|y|<\rho\}} e^{-|y|^2} dV = \frac{A(S^{n-1})}{\pi^{n/2}} \int_0^{\rho} e^{-r^2} r^{n-1} dr.$$

The normalizing constant can be explained by considering the integral over $\mathbb{R}^n$ of $p(\cdot, 1/4)$:

$$1 = \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} e^{-|y|^2} dV = \frac{A(S^{n-1})}{\pi^{n/2}} \int_0^{\infty} e^{-r^2} r^{n-1} dr,$$

so $\lim_{\rho\to\infty} erf_n(\rho) = 1$. Using this definition, we have:

$$\int_{|x|>R} p(x,t)dV = 1 - erf_n(\frac{R}{\sqrt{4t}}),$$

which shows that, for each fixed $R > 0$, $\lim_{t\to 0+} \int_{|x|>R} p(x,t)dV = 0$. More precisely, the heat kernel has the following property.

Uniform concentration property. For any $R > 0$ and every $\epsilon > 0$, there exists $\tau > 0$ (depending only on R and ϵ) so that for all $x \in \mathbb{R}^n$:

$$0 < t < \tau \Rightarrow \int_{\{|y-x|>R\}} p(x-y,t)dV(y) < \epsilon.$$

From the preceding discussion, it suffices to choose τ so that:

$$1 - erf_n(\frac{R}{\sqrt{4\tau}}) < \epsilon.$$

Remark 1: erf_1 is related to the function ERF defined earlier by:

$$ERF(\rho) = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^\rho e^{-x^2} dx = \frac{1}{2} + \frac{1}{2} erf_1(\rho), \rho > 0.$$

(Note $S^0 = \{-1, 1\}$, so $A(S^0) = 2$.)

Remark 2: Gamma function. The Gamma function is defined as the convergent improper integral:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, x > 0.$$

One shows easily that (i)$\Gamma(1) = 1$, (ii)$\Gamma(\frac{1}{2}) = \sqrt{\pi}$, (iii)$\Gamma(x+1) = x\Gamma(x)$, (iv) $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$, and by change of variable:

$$\int_0^\infty e^{-r^2} r^{n-1} dr = \frac{1}{2}\Gamma(\frac{n}{2}).$$

Combining with the observation above we find a formula for the area of the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$:

$$A(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}, n \geq 1.$$

It is easier to show that the integral of $p(x,t)$ over the exterior of any cube in $\mathbb{R}^n$ tends to zero as $t \to 0_+$ (or equivalently, that the integral over any cube tends to one.) That is:

$$\lim_{t \to 0_+} \int_{Q_R} p(x,t)dV = 1, \quad Q_R = \{x \in \mathbb{R}^n; |x_i| \leq R \text{ for } i = 1, \ldots, n\}.$$

Exercise 4. Prove this, using the fact that $p(x,t)$ is the product of one-dimensional heat kernels. In fact, prove the 'cubes version' of the uniform

concentration property: given $R > 0$, for any $\epsilon > 0$ we may find $\tau > 0$ so that, for each $x \in \mathbb{R}^n$:

$$0 < t < \tau \Rightarrow \int_{\mathbb{R}^n \setminus Q_R(x)} p(x-y,t)dV(y) < \epsilon$$

where:

$$Q_R(x) = \{y \in \mathbb{R}^n; |y_i - x_i| \leq R \text{ for } i = 1, \ldots, n\}.$$

2.3 Solution of the initial-value problem.

With the properties of the heat kernel in $\mathbb{R}^n$, the proof of the solution formula for the initial-value problem in $\mathbb{R}^n$ follows the same lines as the one-dimensional proof. As Exercise 4 suggests, both the limit/continuity definitions and the proofs may be phrased in terms of either 'balls' or 'cubes' (since every cube contains a ball, and vice-versa). We choose the latter.

A function $u : \mathbb{R}^n \to \mathbb{R}$ is continuous at a given $x \in \mathbb{R}^n$ if:

$$(\forall \epsilon > 0)(\exists \delta > 0) y \in B_\delta(x) \Rightarrow |u(x) - u(y)| < \epsilon.$$

In general δ will depend both on ϵ and on x. We say u is *uniformly continuous* on a set $A \subset \mathbb{R}^n$ if it is continuous at every point of A and, given ϵ, the same δ works for all points x of A.

A one parameter family of functions $u(\cdot, t)$ defined in $\mathbb{R}^n$ (where $t > 0$ is the parameter) *converges uniformly in* R^n to a function u_0 as $t \to 0_+$ if:

$$(\forall \epsilon > 0)(\exists \tau > 0) 0 < t < \tau \Rightarrow (\forall x \in \mathbb{R}^n)|u(x,t) - u_0(x)| < \epsilon.$$

(The point is that τ depends on ϵ, but not on x.)

Theorem. Let $u_0 : \mathbb{R}^n \to \mathbb{R}$ be bounded and piecewise continuous. Then the function:

$$u(x,t) = \int_{\mathbb{R}^n} p(x-y,t)u_0(y)dy$$

is a solution of the heat equation $u_t - \Delta u = 0$ in $\mathbb{R}^n$ for all $t > 0$.

(i) If u_0 is continuous at x, we have: $\lim_{t \to 0_+} u(x,t) = u_0(x)$;

(ii) If u is uniformly continuous in $\mathbb{R}^n$, $u(\cdot, t) \to u_0$ as $t \to 0_+$, uniformly in $\mathbb{R}^n$;

(iii) $u(\cdot, t)$ is a smooth function in $\mathbb{R}^n$, for each $t > 0$.

Proof (outline). This is very similar to the one-dimensional case, so we only outline the proof of (ii). For a given $R > 0$, split the integral:

$$|u(x,t) - u_0(x)| = |\int_{\mathbb{R}^n} p(x-y,t)(u_0(x) - u_0(y))dV(y)|$$

$$\leq \int_{\{|y|\leq R\}} p(x-y,t)|u_0(x)-u_0(y)|dV(y) + \int_{\{|y|>R\}} p(x-y,t)|u_0(x)-u_0(y)|dV(y).$$

Let $M = \sup_{\mathbb{R}^n} |u_0|$. Given $\epsilon > 0$, by uniform continuity we may choose R so that $|u_0(x) - u_0(y)| < \epsilon/2$ for any $x, y \in \mathbb{R}^n$ with $|x - y| \leq R$; this implies the first integral is smaller than $\epsilon/2$. Then choose (using the uniform concentration property) $\tau > 0$ so that we have: $\int_{\{|y|>R\}} p(x-y,t)dV(y) < \epsilon/4M$ for $0 < t < \tau$. (Note that τ depends only on R and ϵ, hence only on u_0 and ϵ.) Then the second integral will also be smaller than $\epsilon/2$, so we have:

$$|u(x,t) - u_0(x)| \leq \epsilon$$

for $t < \tau$. This concludes the proof of uniform convergence (part (ii)).

Exercise 5. (Characteristic functions as initial data). The *characteristic function* χ_A of a set $A \subset \mathbb{R}^n$ is defined by: $\chi_A(x) = 1, x \in A; \chi_A(x) = 0, x \in \mathbb{R}^n \setminus A$.

(i) Find the solution of the one-dimensional heat equation with initial data $u_0 = \chi_{B_R(x_0)}$, where $B_R(x_0)$ is the interval $[x_0 - R, x_0 + R]$ $(x_0 \in \mathbb{R}, R > 0)$. Express your answer in terms of the function erf_1 (which is defined only for *positive* values of its argument.)

(ii) Find the solution of the heat equation in $\mathbb{R}^n$, with initial data $u_0 = \chi_{Q_R(x_0)}$, where $Q_R(x_0)$ is the cube:

$$Q_R(x_0) = \{x \in \mathbb{R}^n; |x^i - x_0^i| \leq R, i = 1, \ldots, n\}, (x_0 \in \mathbb{R}^n, R > 0).$$

(Note that $\chi_{Q_R(x_0)}$ is a product of one-dimensional characteristic functions.)

(iii) Find the value at the origin of solution of the heat equation in $\mathbb{R}^n$ with initial data $u_0 = \chi_{B_R}$, where $B_R = \{x \in \mathbb{R}^n; |x| \leq R\}$ is the ball with center $0 \in \mathbb{R}^n$, radius $R > 0$. Express your answer $(u(0,t))$ in terms of the function erf_n.