TOBY COLDING - TOPICS IN HEAT EQUATIONS

TANG-KAI LEE

ABSTRACT. These are parts of notes from Toby Colding's geometry course taught at MIT in the Spring of 2021. We would like to thank Prof. Colding for an excellent class. Please be aware that it is likely that we have introduced numerous typos and mistakes in our compilation process, and would appreciate it if these are brought to our attention.

Contents

1. **Introduction**

1.1. **Discrete and Continuous Settings.** Imagine that on $\mathbb{Z}^n$, we could measure how many "heat particles" are at any lattice point. Let $p_k(x)$ be the number of heat particles at $x \in \mathbb{Z}^n$ at time $k \in \mathbb{R}_{\geq 0}$. The neighbors of x are the points $y \in \mathbb{Z}^n$ such that $|x - y| = 1$ (so there are $2n$ many neighbors for any x). In this case we will write $x \sim y$. We assume that each heat particle will move to one of its neighbors (randomly chosen) in each unit of time. Then we derive

$$p_{k+1}(x) = \frac{1}{2n} \sum_{y \sim x} p_k(y).$$

As a result,

$$p_{k+1}(x) - p_k(x) = \frac{1}{2n} \sum_{y \sim x} (p_k(y) - p_k(x)).$$

Thus for a function $u \colon \mathbb{Z}^n \to \mathbb{R}$, we define the discrete laplacian of u by

$$\Delta u(x) := \frac{1}{2n} \sum_{y \sim x} (u(y) - u(x)), \tag{1.1}$$

based on which we have

$$p_{k+1(x)} - p_k(x) = \Delta p_k(x).$$

We could formulate it as

$$\partial_t p_k(x) = \Delta p_k(x), \tag{1.2}$$

where ∂_t is the discrete time derivative. (1.2) is called the **discrete heat equation**.

The reason why we define Δ as in (1.1) could be seen from the following. Let $u \colon \mathbb{R}^n \to \mathbb{R}$. For $x \in \mathbb{R}^n$ and $r > 0$, we know

$$\mathrm{Vol}(\partial B_r(x)) = c_n r^{n-1}$$

where c_n is the volume of the unit sphere in $\mathbb{R}^n$. As a result, the average of u over $\partial B_r(x)$ is

$$I(r) := \frac{1}{c_n r^{n-1}} \int_{\partial B_r(x)} u.$$

Then using the polar coordinate,

$$I'(r) = \frac{1}{c_n r^{n-1}} \int_{\partial B_r(x)} \frac{d}{dr} u = \frac{1}{c_n r^{n-1}} \int_{B_r(x)} \Delta u$$

by Stokes' theorem. This provides a heuristic viewpoint to the laplacian (for both discrete and continuous setting).

Now we talk about the continuous analog. Let $u \colon \mathbb{R}^n \times [a, b] \to \mathbb{R}$ where we write $(x, t) \in \mathbb{R}^n \times [a, b]$. Then the heat equation is

$$\partial_t = \Delta u,$$

or

$$(\partial_t - \Delta) u = 0$$

where $\partial_t - \Delta$ is called the heat operator. One way to see the heat equation is the gradient flow for the energy functional. For a smooth function $u\colon \Omega \to \mathbb{R}$, its energy is defined by

$$E(u) = \int_\Omega |\nabla u|^2.$$

We would like to look at variations of u boundary conditions fixed. Thus take $\varphi\colon \Omega \to \mathbb{R}$ with $\varphi|_{\partial\Omega} = 0$, and consider $v_s := u + s\varphi$. Then

$$E(v_s) = \int_\Omega |\nabla(u + s\varphi)|^2 = \int_\Omega (|\nabla u|^2 + 2s \langle \nabla u, \nabla\varphi \rangle + s^2 |\nabla\varphi|^2),$$

so

$$\frac{d}{ds} E(v_s)|_{s=0} = 2\int_\Omega \langle \nabla u, \nabla\varphi \rangle = -2\int_\Omega \varphi\Delta u.$$

Thus if we want $E(v_s)$ to go down as fast as possible, we must let

$$\frac{d}{ds} v_s|_{s=0} = \varphi = \Delta u.$$

This is what it means to be the negative gradient flow for this functional. i.e., if

$$\partial_t u = \Delta u,$$

then $E(u)$ goes down as fast as possible along the flow.

1.2. **Uniqueness.** Suppose u and v satisfy the heat equation on a compact domain $\Omega \times [a, b]$ with $u(\cdot, a) = v(\cdot, a)$ and $u = v$ on $\partial\Omega \times [a, b]$. Then it turns out that $u = v$ everywhere. To see this, let $w := u - v$, which also satisfies the heat equation and vanishes at the initial time and the boundary. Then

$$\frac{d}{dt} \int |\nabla w|^2 = 2\int_\Omega \langle \nabla w_t, \nabla w \rangle = 2\int_\Omega \langle \nabla\Delta, \nabla w \rangle = -2\int_\Omega |\Delta w|^2.$$

That is to say, the energy $E(w(\cdot, t))$ is decreasing. However, the initial energy of w is 0, so it should be 0 at any time and hence $w = 0$, i.e., $u = v$.

If we consider solutions $u, v\colon \mathbb{R}^n \times [a, b] \to \mathbb{R}$ to the heat equation and would like to consider

$$\int_{\mathbb{R}^n} |\nabla u|^2,$$

we need them to **decay** fast near infinity. If it is the case, the argument above still holds. In most cases, one has uniqueness. However, there is a famous example of Tychonoff. That function $u\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ has properties that $u = 0$ for $t \leq 0$ and it grows incredibly fast when $t > 0$. Explicitly, if we let

$$\begin{cases} e^{-\frac{1}{t^2}} \text{ if } t > 0 \\ 0 \text{ if } t \leq 0 \end{cases},$$

then Tychonoff's example is

$$u(x, t) := \sum_{n=0}^{\infty} \varphi^{(n)}(t) \frac{x^{2n}}{2n!}.$$

We could see that at any $t > 0$, $u(\cdot, t)$ grows faster than any exponential functions.

The uniqueness property also gives rise to the heat semigroup property. Let $u(x,t)$ be a solution to the heat equation on $\mathbb{R}^n \times [0,T] \to \mathbb{R}$, with $u(x,0) = u_0(x)$. If we have uniqueness (in either case we discuss above), we could consider an operator P_t by defining

$$P_t u_0(x) := u(x,t).$$

Then it satisfies

$$P_s(P_t u_0) = P_{s+t} u_0$$

by the uniqueness property. This is called the semigroup property of the heat equation.

A crucial property of the heat equation is scaling. Suppose $u\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is a solution to the heat equation. Then for $c > 0$, we define

$$u_c(x,t) := u(cx, c^2 t)$$

for $(x,t) \in \mathbb{R}^n \times \mathbb{R}$. Then

$$\partial_t u_c = c^2 \partial_t u = c^2 \Delta u = \Delta u_c,$$

so u_c is also a solution to the heat equation.

Next we talk about the static solutions to the heat equations, which are harmonic functions, i.e., those $u\colon \mathbb{R}^n \to \mathbb{R}$ with $\Delta u = 0$. We can view u as $u\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ by defining $u(x,t) = u(x)$. Then

$$\partial_t u = 0 = \Delta u$$

if the original u is harmonic.

If $u\colon \mathbb{R}^n \to \mathbb{R}$ satisfies

$$\Delta u + \lambda u = 0$$

for some λ, i.e., u is an eigenfunction with eigenvalue λ, then we consider

$$v(x,t) := e^{-\lambda t} u(x),$$

which satisfies

$$\partial_t v = -\lambda v = \Delta v.$$

A particular example here happens when $u = e^{i\langle y,x\rangle}$, called the plane wave, where $y \in \mathbb{R}^n$ is a constant vector. In this case,

$$\Delta u = \sum_{j=1}^{n} -y_j^2 u = -|y|^2 u.$$

Hence

$$v(x,t) = e^{-|y|^2 t} e^{i\langle y,x\rangle},$$

called the traveling wave, is a solution to the heat equation.

The single most important solution to the heat equation is the fundamental solution

$$u(x,t) := (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}$$

for $(x,t) \in \mathbb{R}^n \times \mathbb{R}_{>0}$. Direct computation shows that it satisfies the heat equation. We will come back to this soon.

1.3. **Other Interesting Topics.** We have no time to discuss some other related interesting topics, such as porous media equations, which is of the form

$$\partial_t u = \Delta u^2 = 2u\Delta u + |2\nabla u|^2$$

for $u > 0$. It also has no finite propagation speed.

2. Dynamics of the Heat Equation

Let $u\colon \Omega \times R_{>0} \to \mathbb{R}$ be a solution to the heat equation, with $u(\cdot,t)|_{\partial\Omega} = 0$. When u comes from a harmonic function, by the maximum principle (assuming Ω is compact), we know $u = 0$, so the only static solution here is 0. In general, we will consider the energy again. We assume the Neumann boundary condition, i.e.,

$$\frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega.$$

We would like to study the dynamics of this heat flow. The first thing we observe is

$$\frac{d}{dt}\int_\Omega u(\cdot,t) = \int_\Omega \Delta u = \int_{\partial\Omega} \frac{\partial u}{\partial n} = 0 \tag{2.1}$$

by Stoke's theorem and our Neumann boundary condition. Thus the average of u is constant (in t). We would like to see the behavior as t becomes large. In sight of the static solution, we may expect it converges to a harmonic function. We already know

$$\frac{d}{dt}\int_\Omega |\nabla u|^2 = 2\int_\Omega \langle \nabla u_t, \nabla u\rangle = -2\int_\Omega (\Delta u)^2. \tag{2.2}$$

To see that u converges to a constant function very fast, we hope to see that

$$\int_\Omega |\nabla u|^2 \to 0$$

very fast. Based on (2.1), we may assume $\int_\Omega u = 0$ by adding a constant. Next, we need the Poincaré inequality, i.e.,

$$\int_\Omega u^2 \le C(\Omega)\int_\Omega |\nabla u|^2. \tag{2.3}$$

The easiest case is when Ω is an interval, in which case the inequality follows from the fundamental theorem of calculus and Cauchy-Schwarz inequality. Since

$$\int_\Omega |\nabla u|^2 = -\int_\Omega u\Delta u \le \left(\int_\Omega u^2\right)^{\frac{1}{2}} \left(\int_\Omega (\Delta u)^2\right)^{\frac{1}{2}},$$

the Poincaré inequality (2.3) implies

$$\int_\Omega |\nabla u|^2 \le \left(C(\Omega)\int_\Omega |\nabla u|^2\right)^{\frac{1}{2}} \left(\int_\Omega (\Delta u)^2\right)^{\frac{1}{2}}$$

so

$$\int_\Omega |\nabla u|^2 \le C(\Omega)\int_\Omega (\Delta u)^2.$$

Hence combining this with (2.2),

$$\frac{d}{dt}\int_\Omega |\nabla u|^2 = -2\int_\Omega (\Delta u)^2 \le -\frac{2}{C}\int_\Omega |\nabla u|^2.$$

Thus if we write

$$E(t) := \int_\Omega |\nabla u|^2,$$

we have

$$E' \le -\widetilde{C}E,$$

which means

$$(e^{\widetilde{C}t}E)' \le 0,$$

implying that E decays exponentially fast to 0 as $t \to \infty$. This is the case when we normalize that $\int u = 0$. In general it will converge to a constant function.

3. **Parabolic Maximum Principle**

Theorem 3.1. Let Ω be a compact domain and assume $u\colon \Omega \times [a,b] \to \mathbb{R}$ satisfies the heat equation. Then

$$\max_{\Omega\times[a,b]} u = \max_{\Omega\times\{0\}\cup\partial\Omega\times[a,b]} u. \tag{3.2}$$

In fact we only need u to be a subsolution to the heat equation, i.e.,

$$(\partial_t - \Delta)u \le 0.$$

(Proof of (3.2)**.)** First we assume a special case that

$$(\partial_t - \Delta)u < 0, \tag{3.3}$$

and say the maximum is achieved at some (x_0, t_0). We may assume this point is not on $\Omega \times \{0\} \cup \partial\Omega \times [a,b]$, so x_0 is in the interior of Ω. Thus

$$\Delta u(x_0, t_0) \le 0, \tag{3.4}$$

and

$$u_t(x_0, t_0) \ge 0. \tag{3.5}$$

To see (3.5), if $u_t < 0$ at that point, the maximality would be violated at some $t_0 - \varepsilon$. Now (3.4) and (3.5) give

$$(\partial_t - \Delta)u(x_0, t_0) \ge 0,$$

contradicting our assumption (3.3).

In general, if $(\partial_t - \Delta)u \le 0$, we consider

$$v_\varepsilon(x,t) := u(x,t) - \varepsilon t$$

for $\varepsilon > 0$. Then

$$(\partial_t - \Delta)v_\varepsilon = (\partial_t - \Delta)u - \varepsilon < 0.$$

Thus by the previous case we know

$$\max_{\Omega\times[a,b]} v_\varepsilon = \max_{\Omega\times\{0\}\cup\partial\Omega\times[a,b]} v_\varepsilon.$$

This holds for any $\varepsilon > 0$, so (3.2) follows. $\square$

4. **Discrete Heat Equation**

4.1. **Background Setting.** We come back to the discrete setting. We could do it more generally. Let Γ be a finite graph, i.e., it consists of a finite set of points, called the vertices, and pairs of some vertices, called edges. We would look at simple graphs, i.e., every vertex is not connected to itself.

For vertices $v_1, v_2 \in \Gamma$, the graph distance between them is the smallest number of edges that connect v_1 and v_2. Then we define the neighbor of a vertex x to be those vertices that of distance 1 to x. We would write $y \sim x$, by which we mean y is a neighbor of x. The number of its neighbor is called the degree of x, denoted by $\deg_x$.

The laplacian on the graph is defined for $u\colon \Gamma \to \mathbb{R}$ by

$$\Delta u\colon \Gamma \to \mathbb{R}, x \mapsto \frac{1}{\deg_x} \sum_{y \sim x} (u(y) - u(x)),$$

which generalizes what we defined before. Obviously Δ is a linear operator. The discrete heat equation for $u\colon \Gamma \times \{0, 1, 2, 3, \cdots\} \to \mathbb{R}$ is

$$\partial_t u = \Delta u, \tag{4.1}$$

where ∂_t is the discrete time derivative, which operates as

$$\partial_t u(x,t) := u(x, t+1) - u(x,t).$$

We could rewrite (4.1) as

$$\begin{aligned} u(x,t+1) - u(x,t) &= \frac{1}{\deg_x} \sum_{y \sim x} (u(y,t) - u(x,t)) \\ &= \frac{1}{\deg_x} \sum_{y \sim x} u(y,t) - u(x,t), \end{aligned}$$

i.e.,

$$u(x,t+1) = \frac{1}{\deg_x} \sum_{y \sim x} u(y,t),$$

which fits our expectation.

4.2. **Adjancecy Map.** There is another viewpoint. For $u\colon \Gamma \to \mathbb{R}$, we could define $Au\colon \Gamma \to \mathbb{R}$ by

$$Au(x) := \frac{1}{\deg_x} \sum_{y \sim x} u(y).$$

This A is called the adjacency map. Thus the heat equation is equivalent to

$$u(x,t+1) = Au(x,t). \tag{4.2}$$

For $u\colon \Gamma \to \mathbb{R}$, the total integral u is defined by

$$\sum_{x\in\Gamma} \deg_x \cdot u(x).$$

Thus the total integral of Δu is

$$\begin{aligned}
\sum_{x\in\Gamma} \deg_x \cdot \Delta u(x) &= \sum_{c\in\Gamma} \deg_x \cdot \left(\frac{1}{\deg_x} \sum_{y\sim x} (u(y) - u(x)) \right) \\
&= \sum_{x} \sum_{y\sim x} (u(y) - u(x)) \\
&= \sum_{x} \sum_{y\sim x} u(y) - \sum_{x} \deg_x \cdot u(x) \\
&= \sum_{x} \deg_y \cdot u(y) - \sum_{x} \deg_x \cdot u(x) \\
&= 0.
\end{aligned}$$

In the continuous setting, we have seen that the Neumann boundary condition guarantees $\frac{d}{dt}\int u = 0$, which is the first law of thermodynamics (conservation of energy). In the discrete setting, since

$$\sum_{x\in\Gamma} \deg_x \cdot \Delta u(x,t) = 0,$$

we get

$$\sum \deg_x \cdot \partial_t u(x,t) = \sum_{x\in\Gamma} \deg_x \cdot \Delta u(x,t) = 0$$

if u satisfies the discrete heat equation. Thus

$$\sum_{x} \deg_x (u(x,t+1) - u(x,t)) = 0,$$

that is to say,

$$\sum_{x} \deg_x \cdot u(x,t+1) = \sum_{x} \deg_x \cdot u(x,t).$$

This is the discrete version of the first law of thermodynamics (conservation of heat particles).

Now for any function $u\colon \Gamma \to \mathbb{R}$, we have defined

$$Au(x) := \frac{1}{\deg_x} \sum_{y\sim x} u(y).$$

Clearly A is also a linear operator, and by (4.2) we could formulate the heat equation as

$$u_n(x) := u(x,n) = A^n u_0(x). \tag{4.3}$$

In fact, we could view A as a self-adjoint operator with respect to a natural inner product. For $u, v\colon \Gamma \to \mathbb{R}$, we consider

$$\langle u, v\rangle := \sum_{x\in\Gamma} \deg_x \cdot u(x) v(x).$$

Then

$$\begin{aligned}\langle Au, v\rangle &= \sum_x \deg_x \cdot Au(x)v(x) \\ &= \sum_x \deg_x \cdot \left(\frac{1}{\deg_x}\sum_{y\sim x} u(y)\right) v(x) \\ &= \sum_x \sum_{y\sim x} u(y)v(x) \\ &= \sum_y \sum_{x\sim y} u(x)v(y) = \langle u, Av\rangle\,,\end{aligned}$$

which justifies what we just said.

We could identify the space of functions $\Gamma \to \mathbb{R}$ as $\mathbb{R}^\Gamma$ (noting $|\Gamma| < \infty$), and view

$$A\colon \mathbb{R}^\Gamma \to \mathbb{R}^\Gamma.$$

Thus it admits a basis of eigenfunctions. What are the possible eigenvalues? (Our convention here is that λ is an eigenvalue if $Au + \lambda u = 0$.) From now on we assume Γ is connected, i.e., any two different vertices are connected. It turns out that there exists a constraint on the eigenvalues.

Theorem 4.4. If $Au + \lambda u = 0$, then $\lambda \in [-1, 1]$.

(Proof.) Suppose $Au + \lambda u = 0$. i.e.,

$$-\lambda u(x) = \frac{1}{\deg_x}\sum_{y\sim x} u(y).$$

If $|u|$ achieves its maximum at x_M, then

$$|\lambda||u(x_M)| \le \frac{1}{\deg_{x_M}}\sum_{y\sim x_M} |u(y)| \le \frac{1}{\deg_{x_M}}\sum_{y\sim x_M} |u(x_M)| = |u(x_M)|$$

so $|\lambda| \le 1$. $\square$

Note that if $\lambda = -1$, then $Au = u$, which means that u is harmonic. Looking at its value at its maximum, we get that u should be constant (on each connected component). (Otherwise, it could not equal its average.) Equivalently, we prove that harmonic functions are all constant.

When $\lambda = 1$, the graph is very special. It is called a bi-partite graph, in the sense that $\Gamma = \Gamma_1 \dot\cup \Gamma_2$ and any edge of Γ goes from a vertex of Γ_1 to one of Γ_2. To prove this, note that it implies

$$u(x) = -\frac{1}{\deg_x}\sum_{y\sim x} u(y),$$

so

$$u(y) = -u(x_M)$$

for $y \sim x_M$ if $|u|$ achieves its maximum at x_M (as above). Thus we can define $\Gamma_1 = \{x : u(x) = u(x_M)\}$ and $\Gamma_2 = \{x : u(x) = -u(x_M)\}$.

Thus if u is not harmonic and Γ is not bi-partite, then it will be contract under A, i.e.,

$$|u(\cdot,d)|_{L^2} \le \lambda^d |u(\cdot,0)|_{L^2}$$

for some $\lambda \in (0,1)$.

4.3. **On Integral Lattices.** Next, we consider $u\colon \mathbb{Z}^n \times \mathbb{Z}_+ \to \mathbb{R}$, with Δu defined in the same way, so

$$(\Delta u)(x) = \frac{1}{2n}\sum_{y\sim x}(u(y)-u(x)).$$

We may define it a little different, in the sense that

$$(\Delta u)(x) = \frac{1}{2n\delta}\sum_{y\sim x}(u(y)-u(x))$$

for $u\colon \sqrt{\delta}\mathbb{Z}^n \times \delta\mathbb{Z}_+ \to \mathbb{R}$. In this way, the speed is

$$\frac{\sqrt{\delta}}{\delta} = \frac{1}{\sqrt{\delta}} \to \infty$$

as $\delta \to 0$. Why do we consider this kind of scaling? Note that if $(\partial_t - \Delta)u = 0$, then

$$u_a(x,t) := u(ax, a^2 t)$$

also solves the heat equation for $a > 0$. Thus there is a^2 in time and a in space. In fact, as $\delta \to 0$, the solution to the above scaled heat equation will converge to that to the continuous heat equation (with the help of some gradient estimate).

5. Heat Kernel on the Euclidean Space

5.1. **Fundamental Solutions.** Consider $h\colon \mathbb{R}^n \times \mathbb{R}_{>0} \to \mathbb{R}$ defined by

$$h(x,t) := (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}.$$

We claim that h satisfies the heat equation. Before proving that we start with a lemma.

Lemma 5.1. Let $u\colon \mathbb{R}^n \times [a,b] \to \mathbb{R}$ and $f\colon \mathbb{R} \to \mathbb{R}$. Then

$$(\partial_t - \Delta)f(u) = f'(u)(\partial_t - \Delta)u - f''(u)|\nabla u|^2.$$

(Proof of lemma 5.1.) Note

$$\nabla f(u) = f'(u)\nabla u.$$

Thus

$$\Delta f(u) = \operatorname{div}(\nabla f(u)) = f''|\nabla u|^2 + f'\Delta u.$$

On the other hand,

$$\partial_t f(u) = f'\partial_t u.$$

Combining these gives the result. $\square$

Now we apply lemma 5.1 to

$$f(s) := e^s \text{ and } u := -\frac{n}{2}\log t - \frac{|x|^2}{4t},$$

which gives

$$f(u) = t^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}} = (4\pi)^{\frac{n}{2}} h.$$

Hence it is equivalent to prove that $f(u)$ is a solution. To see that $f(u)$ satisfies the heat equation, note

$$\Delta u = -\frac{\Delta |x|^2}{4t} = -\frac{2n}{4t}$$

and

$$\nabla u = -\frac{x}{2t},$$

which tells us that

$$|\nabla u|^2 = \frac{|x|^2}{4t^2}.$$

On the other hand,

$$u_t = -\frac{n}{2t} + \frac{|x|^2}{4t^2}.$$

Thus

$$(\partial_t - \Delta)u = \frac{|x|^2}{4t^2} = |\nabla u|^2.$$

Hence lemma 5.1 implies

$$(\partial_t - \Delta) f(u) = e^u (\partial_t - \Delta) u - e^u |\nabla u|^2 = 0.$$

In conclusion, we derive that $h(x,t)$ is a solution to the heat equation on the euclidean space, and we call h the fundamental solution (the reason of which could be seen later).

5.2. **Heat Kernels.** Now we assume $u(x,t)$ is a solution to the heat equation. Then for any $y \in \mathbb{R}^n$,

$$v(x,t) := u(x-y,t)$$

also solves the heat equation. This means that

$$H(x,y,t) := h(x-y,t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x-y|^2}{4t}}$$

is a solution to the heat equation (in (x,t)). Notice that H is symmetric in x and y, i.e., $H(x,y,t) = H(y,x,t)$. Then the fact that

$$(\partial_t - \Delta_x) H = 0$$

implies

$$(\partial_t - \Delta_y) H = 0$$

also.

Another nice property, which explains why 4π comes in, is that the integral of h is

$$\int_{\mathbb{R}^n} h(x,t)dx = \int_{\mathbb{R}^n} (4\pi)^{-\frac{n}{2}} e^{-\frac{|z|^2}{4}} dz = 1$$

for any fixed t, by a change of variable $z = \frac{x}{\sqrt{t}}$.

Observing the pattern of h (i.e., concentration to the origin as $t \to 0$) along with the fact that it has integral 1, for any continuous and bounded u, we have the limit

$$\lim_{t\to 0^+} \int_{\mathbb{R}^n} h(x,t)u(x)dx = u(0).$$

Thus

$$\lim_{t\to 0^+} \int_{\mathbb{R}^n} H(x,y,t)u(y)dy = u(x).$$

To be more precise, for a continuous and bounded $v\colon \mathbb{R}^n \to \mathbb{R}$, consider

$$u(x,t) := \int_{\mathbb{R}^n} H(x,y,t)v(y)dy.$$

Then u satisfies, as mentioned above,

$$\lim_{t\to 0^+} u(x,t) = v(x)$$

and furthermore,

$$(\partial_t - \Delta_x)u = \int_{\mathbb{R}^n} \left((\partial_t - \Delta_x)H\right)(x,y,t)v(y)dy = 0.$$

That is, u is a solution to the heat equation with initial value $u(x,0) = v(x)$. This explains the terminology of the fundamental solution. (Beware that we need some mild condition on the initial data v. Here of course boundedness suffices.)

We summarize the properties of $H(x,y,t)$.

1. $H(x,y,t) = H(y,x,t)$.
2. $(\partial_t - \Delta_x)H = (\partial_t - \Delta_y)H = 0$.
3. Reproducing property: $u(x,t) := \int_{\mathbb{R}^n} H(x,y,t)v(y)dy$. solves the heat equation with initial data v with mild growth.

The next question is if this applies when we change $\mathbb{R}^n$ to other spaces. For example, consider the circle S^1. Functions on S^1 could be viewed as periodic functions on $\mathbb{R}$. We need to recall some Fourier analysis. Let

$$c_k := \frac{\sin kx}{\sqrt{\pi}} \text{ and } s_k := \frac{\sin kx}{\sqrt{\pi}}.$$

Then

$$c_k^2 + s_k^2 = \frac{1}{\pi}(\cos^2 kx + \sin^2 kx) = \frac{1}{\pi}.$$

Thus

$$\int_0^{2\pi} c_k^2 + \int_0^{2\pi} s_k^2 = 2,$$

in which we have

$$\int_0^{2\pi} c_k^2 = \int_0^{2\pi} s_k^2 = 1.$$

In fact, c_k's and s_k's ($k \in \mathbb{N}$) along with 1 form an orthonormal basis for all L^2 functions on S^1.

Note that
$$\Delta c_k = c_k'' = -k^2 c_k$$
and likewise
$$\Delta s_k = -k^2 s_k.$$
These lead to considering
$$H(x,y,t) := \sum_k \left(c_k(x)c_k(y)e^{-k^2t} + s_k(x)s_k(y)e^{-k^2t} \right) = H(y,x,t),$$
which is well-defined for any $t > 0$ (by the boundedness of c_k and s_k and the fast decay of the exponential function). The reason to add the exponential function is the following observation
$$(\partial_t - \Delta)(c_k e^{-k^2t}) = -k^2 c_k e^{-k^2t} + k^2 e^{-k^2t} = 0.$$
Similarly we have $(\partial_t - \Delta)(s_k e^{-k^2t}) = 0$. Thus we conclude
$$(\partial_t - \Delta)H = 0.$$
Next we look at the reproducing property. For a continuous function v, we can write it as
$$v(y) = \sum_k (a_k c_k(y) + b_k s_k(y))$$
where a_k and b_k are its Fourier coefficients. Then by the orthogonality,
$$\begin{aligned}\int_{S^1} H(x,y,t)v(y)dy &= \sum_{k,l} \int_{S^1} \left(c_k(x)c_k(y)e^{-k^2t}a_l c_l(y) + s_k(x)s_k(y)e^{-k^2t}b_l s_l(y) \right) dy \\ &= \sum_k \left(c_k(x)a_k + s_k(x)b_k\right) e^{-k^2t} \to v(x)\end{aligned}$$
as $t \to 0^+$ for any $x \in S^1$. Thus we are fine.

The third example is in the discrete setting. Recall what we did in the section 4. Let Γ be a finite simple graph which is not bi-partite. If u is a solution (to the discrete heat equation) orthogonal to constant, i.e., $\frac{1}{\deg_x}\sum_{y\sim x} u(y) = 0$, we have $\|Au\| \leq \mu\|u\|$ for some $\mu < 1$. Thus by (4.3),
$$\|u(\cdot,n)\| \leq \mu^k \|u(\cdot,0)\|.$$
Hence $u(\cdot,n) \to 0$ exponentially fast. In general, since $A1 = 1$, we can find c such that $u(\cdot,0) - c \perp 1$, so $u(\cdot,n) \to c$.

Note that we can write $\Delta = A - 1$ (where 1 means the identity map). Thus by theorem 4.4, its eigenvalues lie in $[0,2]$. Also, the same arguments as that for A shows that we have an ONB of eigenfunctions, say φ_i's with eigenvalues λ_i's ($i = 1,2,\cdots,|\Gamma|$). Then we consider (a finite sum)
$$H(x,y,t) := \sum_i \varphi_i(x)\varphi_i(y)(1-\lambda_i)^t.$$

We check that this is the heat kernel. It is clear that it is symmetric in x and y. To see that it is a solution, note

$$\begin{aligned} H(x,y,t+1) - H(x,y,t) &= \sum_i \varphi_i(x)\varphi_i(y)(1-\lambda_i)^{t+1} - \sum_i \varphi_i(x)\varphi_i(y)(1-\lambda_i)^t \\ &= \sum_i \varphi_i(x)\varphi_i(y)(1-\lambda_i)^t \cdot (-\lambda_i). \end{aligned}$$

On the other hand,

$$\Delta_x H = \sum_i -\lambda_i \varphi_i(x)\varphi_i(y)(1-\lambda_i)^t,$$

so $\partial_t H = \Delta_x H$.

The next thing to prove is the reproducing property. For a function v on Γ, consider

$$u(x,t) := \sum_{y\in\Gamma} \deg_y H(s,y,t)v(y) = \sum_y \sum_i \deg_y \varphi_i(x)\varphi_i(y)(1-\lambda_i)^t v(y)$$

Clearly u satisfies the heat equation. Furthermore,

$$\begin{aligned} u(x,0) &= \sum_y \sum_i \deg_y \varphi_i(x)\varphi_i(y)v(y) \\ &= \sum_i \varphi_i(x) \sum_y \deg_y \varphi_i(y)v(y) \\ &= \sum_i \varphi_i(x) \langle v, \varphi_i \rangle = v(x) \end{aligned}$$

since φ_i's form an ONB.

5.3. **Green Functions.** Recall the heat kernel

$$H(x,y,t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x-y|^2}{4t}}$$

on $\mathbb{R}^n$. Here we assume $n \geq 3$ (to make the function integrable). Then we define

$$G(x,y) = \int_0^\infty H(x,y,t)dt.$$

Since

$$\lim_{t\to 0^+} H(x,y,t) = 0$$

for $x \neq y$ (by the concentration nature of H), we have

$$\Delta_x G(x,y,t) = \int_0^\infty \Delta_x H(x,y,t)dt = \int_0^\infty H_t(x,y,t)dt = 0$$

by the fundamental theorem of calculus.

We can use the function G, called the Green function, to solve the Poisson equation. To be precise, for a given function f, Poisson problem asks for the solution to $\Delta u = f$. Indeed, if we assume f has compact support, and consider

$$u(x) := -\int_{\mathbb{R}^n} G(x,y)f(y)dy,$$

then

$$\begin{aligned}
\Delta_x u &= -\Delta_x \int_{\mathbb{R}^n} \left(\int_0^\infty H(x,y,t)dt \right) f(y)dy \\
&= -\int_{\mathbb{R}^n} \left(\int_0^\infty \Delta_x H(x,y,t)dt \right) f(y)dy \\
&= -\int_{\mathbb{R}^n} \left(\int_0^\infty \frac{d}{dt} H(x,y,t)dt \right) f(y)dy \\
&= -\int_0^\infty \left(\frac{d}{dt} \int_{\mathbb{R}^n} H(x,y,t)f(y)dy \right) dt = -(0 - f(x)) = f(x).
\end{aligned}$$

5.4. **Parabolic Mean Value Inequalities.** Let $u \colon \mathbb{R}^n \times [-T, 0] \to \mathbb{R}$ with $(\partial_t - \Delta)u \leq 0$. We want to estimate $u(0,0)$. First, consider

$$H_b(x,t) = (-4\pi t)^{-\frac{n}{2}} e^{\frac{|x|^2}{4t}} = h(x, -t)$$

where h is the fundamental solution. Then

$$\partial_t H_b = -\partial_t h = -\Delta h = -\Delta H_b.$$

That is to say, H_b satisfies the backward heat equation $(\partial_t + \Delta)H_b$.

Next, if we let

$$I(t) := (-4\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} u(x,t) e^{\frac{|x|^2}{4t}} dx = \int_{\mathbb{R}^n} u(x,t) H_b(x,t) dx,$$

then

$$I' = \int u_t H_b + \int u (H_b)_t \leq \int \Delta u \cdot H_b - \int u \Delta H_b = 0 \tag{5.2}$$

after integration by parts. (Here of course we need to assume u does not grow too fast.) Thus we know that I is decreasing.

Note that

$$\lim_{t \to 0^-} I(t) = u(0,0)$$

by the property of the heat kernel, and since I is decreasing,

$$(4\pi t)^{-\frac{n}{2}} \int u(x,t) e^{-\frac{|x|^2}{4T}} dx = I(-T) \geq I(0^-) = u(0,0).$$

This is the simplest form of the parabolic mean value inequality.

6. **Central Limit Theorem**

Assume $u\colon \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfies $(\partial_t - \Delta)u = 0$ and $\int u_0 dx = 1$. Then with some mild growth assumption, we know

$$u(x,t) = \int H(x,y,t)u_0(y)dy.$$

If we consider

$$v(x,t) := t^{\frac{n}{2}} u(\sqrt{t}x, t),$$

it turns out that

$$v(x,t) \to (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}} \tag{6.1}$$

as $t \to \infty$. To see (6.1), note

$$\begin{aligned} v(x,t) &= (4\pi)^{-\frac{n}{2}} \int u_0(y) e^{-\frac{|\sqrt{t}x - y|^2}{4t}} dy \\ &= (4\pi)^{-\frac{n}{2}} \int u_0(y) e^{-\frac{|x|^2}{4}} \cdot e^{-\frac{|y|^2}{4t} + \frac{1}{2\sqrt{t}}\langle x, y\rangle} dy. \end{aligned}$$

By dominated convergence theorem,

$$\begin{aligned} v(x,t) &\to (4\pi)^{-\frac{n}{2}} \int u_0(y) e^{-\frac{|x|^2}{4}} dy \\ &= (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}} \int u_0(y) dy \end{aligned}$$

and by our assumption of unit integral, (6.1) follows. We will use (6.1), together with the monotonicity (5.2), to show inequalities not directly related to the heat equation.

We will use the theorem to prove some functional inequalities.

6.1. **Application 1. Hölder Inequality.**

Theorem 6.2. Let h_1 and h_2 be C^1 functions with compact support. Then for $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\int |h_1 h_2| \leq \|h_1\|_{L^p} \|h_2\|_{L^q} \tag{6.3}$$

and the equality holds if and only if $f = cg$ for some $c \in \mathbb{R}$.

Before proving this inequality, we first observe that if f and g are positive subsolutions to the heat equation, then by letting

$$u := \log\left(f^{\frac{1}{p}} g^{\frac{1}{q}}\right) = \frac{1}{p}\log f + \frac{1}{q}\log g,$$

we have

$$\nabla u = \frac{1}{p}\frac{\nabla f}{f} + \frac{1}{q}\frac{\nabla g}{g} \tag{6.4}$$

and

$$\Delta u = \frac{1}{p}\left(\frac{\Delta f}{f} - \frac{|\nabla f|^2}{f^2}\right) + \frac{1}{q}\left(\frac{\Delta g}{g} - \frac{|\nabla g|^2}{g^2}\right).$$

Thus

$$\begin{aligned}(\partial_t - \Delta)u &= \frac{1}{p}\left(\frac{(\partial_t - \Delta)f}{f} + \frac{|\nabla f|^2}{f^2}\right) + \frac{1}{q}\left(\frac{(\partial_t - \Delta)g}{g} + \frac{|\nabla g|^2}{g^2}\right)\\ &\geq \frac{1}{p}\frac{|\nabla f|^2}{f^2} + \frac{1}{q}\frac{|\nabla g|^2}{g^2},\end{aligned}$$

so by (6.4),

$$\begin{aligned}(\partial_t - \Delta)e^u &= e^u\left((\partial_t - \Delta)u - |\nabla u|^2\right)\\ &\geq e^u \frac{1}{p}\left(\frac{(\partial_t - \Delta)f}{f} + \frac{|\nabla f|^2}{f^2}\right) + \frac{1}{q}\left(\frac{(\partial_t - \Delta)g}{g} + \frac{|\nabla g|^2}{g^2}\right)\\ &\geq \frac{1}{p}\frac{|\nabla f|^2}{f^2} + \frac{1}{q}\frac{|\nabla g|^2}{g^2} - \left|\frac{1}{p}\frac{\nabla f}{f} + \frac{1}{q}\frac{\nabla g}{g}\right|^2.\end{aligned}$$

After simplification, we obtain

$$\begin{aligned}e^{-u}(\partial_t - \Delta)e^u &= \frac{1}{pq}\left(\frac{|\nabla f|^2}{f^2} + \frac{|\nabla g|^2}{g^2} - \frac{2\langle \nabla f, \nabla g\rangle}{fg}\right)\\ &\geq \frac{1}{pq}\left|\frac{\nabla f}{f} - \frac{\nabla g}{g}\right|^2\\ &= \frac{1}{pq}\left|\nabla \log \frac{f}{g}\right|^2.\end{aligned}$$

Thus

$$(\partial_t - \Delta)(f^{\frac{1}{p}} g^{\frac{1}{q}}) \geq \frac{f^{\frac{1}{p}} g^{\frac{1}{q}}}{qp}\left|\nabla \log \frac{f}{g}\right|^2 \geq 0.$$

If we consider

$$I(t) := \int_{\mathbb{R}^n \times \{t\}} f^{\frac{1}{p}} g^{\frac{1}{q}},$$

then by integration by parts (and mild growth assumption of f and g)

$$I'(t) = \int_{\mathbb{R}^n \times \{t\}} \partial_t\left(f^{\frac{1}{p}} g^{\frac{1}{q}}\right) = \int_{\mathbb{R}^n \times \{t\}} (\partial_t - \Delta)\left(f^{\frac{1}{p}} g^{\frac{1}{q}}\right) \geq 0$$

and if $I'(t) = 0$, we get $f = cg$.

(Proof of (6.3)**.)** We may assume h_1 and h_2 are non-negative, by seeing their absolute values. In fact, we let

$$f := |h_1|^p \text{ and } g := |h_2|^q.$$

What we really want to do is the following. We also write f and g be the solutions to the heat equation with initial data

$$f(\cdot, 0) = |h_1|^p \text{ and } g(\cdot, 0) := |h_2|^q.$$

Then the parabolic maximum principle implies f and g are non-negative, with

$$(\partial_\Delta)f = (\partial_\Delta)g = 0.$$

By the observation above, the quantity

$$\int f^{\frac{1}{p}} g^{\frac{1}{q}}$$

is increasing. Now we consider

$$u(x,t) := t^{\frac{n}{2}} f(\sqrt{t}x, t)$$

and

$$v(x,t) := t^{\frac{n}{2}} g(\sqrt{t}x, t),$$

we have

$$u(x,t) \to c_f (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}$$

and

$$v(x,t) \to c_g (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}$$

as $t \to \infty$ by the central limit theorem, with the constants $c_f = \int f$ and $c_g = \int g$. Note that

$$\int f^{\frac{1}{p}} g^{\frac{1}{q}} = \int u^{\frac{1}{p}} v^{\frac{1}{q}} \tag{6.5}$$

since

$$\begin{aligned} u^{\frac{1}{p}} v^{\frac{1}{q}} &= \left(t^{\frac{n}{2}} f(\sqrt{t}x, t)\right)^{\frac{1}{p}} \left(t^{\frac{n}{2}} g(\sqrt{t}x, t)\right)^{\frac{1}{q}} \\ &= t^{\frac{n}{2}} f^{\frac{1}{p}}(\sqrt{t}x, t) g^{\frac{1}{q}}(\sqrt{t}x, t) \end{aligned}$$

by the exponent assumption. By a change of variable, (6.5) follows (at a specific time). Since (6.5) is increasing and we have the convergence

$$\begin{aligned} \int u^{\frac{1}{p}} v^{\frac{1}{q}} \to& \int \left(c_f (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}\right)^{\frac{1}{p}} \left(c_g (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}\right)^{\frac{1}{q}} \\ &= c_f^{\frac{1}{p}} c_g^{\frac{1}{q}} \int (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}} \\ &= c_f^{\frac{1}{p}} c_g^{\frac{1}{q}} \\ &= \left(\int f\right)^{\frac{1}{p}} \left(\int g\right)^{\frac{1}{q}} \\ &= \left(\int |h_1|^p\right)^{\frac{1}{p}} \left(\int |h_2|^q\right)^{\frac{1}{q}}. \end{aligned}$$

In particular, by the monotonicity,

$$\left(\int |h_1|^p\right)^{\frac{1}{p}} \left(\int |h_2|^q\right)^{\frac{1}{q}} \geq \int f^{\frac{1}{p}}(x,0) g^{\frac{1}{q}}(x,0) = \int |h_1 h_2|$$

and the inequality follows. When we have equality, the monotonicity implies the integral $I(t)$ is constant in time, in which case we know $f = cg$ since $I'(t) = 0$. $\square$

The spirit is that we use the monotonicity of a quantity to reduce the proof to a limiting case, in which some quantities are equal, or easy to compare.

6.2. **Generalizations.** Let $F\colon \mathbb{R}^2 \to \mathbb{R}$, and $u, v\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$. (In the case of the Höler inequality, $F(x,y) = x^{\frac{1}{p}} y^{\frac{1}{q}}$.) Then we have the following generalized result.

Proposition 6.6. Suppose $F = F(x,y)$ satisfies $F_x \geq 0$, $F_y \geq 0$ and that F is concave, which is equivalent to say that Hess_F is negative semi-definite. If $(\partial_t - \Delta)u \geq 0$ and $(\partial_t - \Delta)v \geq 0$, then $(\partial - \Delta)F(u,v) \geq 0$.

(Proof.) Let $w := F(u,v)$. Then

$$\nabla w = F_x \nabla u + F_y \nabla v, w_t = F_x u_t + F_y v_t,$$

and

$$\Delta w = F_x \Delta u + F_y \Delta v + F_{xx} |\nabla u|^2 + F_{yy} |\nabla v|^2 + 2F_{xy} \langle \nabla u, \nabla v \rangle .$$

Therefore, by assumptions we have

$$\begin{aligned}(\partial_t - \Delta)w &= F_x(\partial_t - \Delta)u + F_y(\partial_t - \Delta)v - F_{xx}|\nabla u|^2 - F_{yy}|\nabla v|^2 - 2F_{xy} \langle \nabla u, \nabla v \rangle \\ &\geq -F_{xx}|\nabla u|^2 - F_{yy}|\nabla v|^2 - 2F_{xy} \langle \nabla u, \nabla v \rangle .\end{aligned}$$

We would like to see that this sum is also non-negative. In fact, it is at least,

$$\begin{pmatrix} |\nabla u| & |\nabla v| \end{pmatrix} \begin{pmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{pmatrix} \begin{pmatrix} |\nabla u| \\ |\nabla v| \end{pmatrix},$$

so by the concavity assumption, we are done. □

If we write $I(t) = \int F(u,v)$, with some mild growth assumptions, again we can use integration by parts to derive

$$I'(t) = \int (\partial_t - \Delta)F(u,v) \geq 0$$

with the proposition 6.6

In the Hölder inequality case, we can compute F_x, F_y, F_{xx} and F_{yy} directly and see that $F_x, F_y \geq 0$, $F_{xx} + F_{yy} \leq 0$ and $F_{xx}F_{yy} - (F_{xy})^2 \geq 0$. Thus the observation in the section 6.1 is covered.

We ould also obtain another interesting inequality, as follows. If we consider

$$F(x,y) := \frac{xy}{x+y}$$

for $x, y > 0$, then we can check that F satisfies the assumption of the proposition 6.6. Thus if u and v solve the heat equation, then $(\partial_t - \Delta)F(u,v) \geq 0$, and hence again we get

$$\frac{d}{dt} \int F(u,v) \geq 0$$

with some mild growth assumptions on u and v. Using the central limit theorem and this monotonicity, we can derive

$$\int \frac{uv}{u+v} \leq \frac{\int u \int v}{\int u + \int v}.$$

7. **Drift Laplacian**

Let $u, v\colon \mathbb{R}^n \to \mathbb{R}$ be smooth with compact support. Then we could consider

$$\langle u, v \rangle = \int uv$$

and

$$\langle\langle \nabla u, \nabla v \rangle\rangle = \int \langle \nabla u, \nabla v \rangle = -\int u \Delta v.$$

Thus Δ is self-adjoint with respect to this inner product.

We could also consider weighted inner product. For example, for $\varphi\colon \mathbb{R}^n \to \mathbb{R}$, we could consider

$$\langle u, v \rangle_\varphi = \int uve^{-\varphi}$$

and

$$\langle\langle \nabla u, \nabla v \rangle\rangle = \int \langle \nabla u, \nabla v \rangle e^{-\varphi}.$$

With respect to this weighted inner product, Δ is not self-adjoint any more. However, note that

$$\operatorname{div}(ue^{-\varphi}\nabla v) = \langle \nabla u, \nabla v \rangle e^{-\varphi} - u \langle \nabla \varphi, \nabla v \rangle e^{-\varphi} + ue^{-\varphi}\Delta v,$$

which, by the Stokes' theorem, implies

$$\begin{aligned} \langle\langle \nabla u, \nabla v \rangle\rangle_\varphi &= \int \langle \nabla u, \nabla v \rangle e^{-\varphi} \\ &= -\int (u\Delta v - u \langle \nabla \varphi, \nabla v \rangle) e^{-\varphi} \end{aligned}$$

Thus if we define

$$L_\varphi v := \Delta v - \langle \nabla \varphi, \nabla v \rangle,$$

we obtain

$$\int \langle \nabla u, \nabla v \rangle e^{-\varphi} = -\int (uL_\varphi v) e^{-\varphi}.$$

L_φ is called the **drift Laplacian**, and the inner product term is called the drift term.

Example 1. When φ is constant, $L_\varphi = \Delta$.

Example 2. When $\varphi = \frac{|x|^2}{4}$, we have

$$L_\varphi v = \Delta v - \frac{1}{2} \langle x, \nabla v \rangle, \tag{7.1}$$

called **the Ornstein-Uhlenbeck operator**, which plays a crucial role when we try to understanding the scaling of the heat equation.

Example 3. When $\varphi = -\frac{|x|^2}{4}$, we have

$$L_\varphi v = \Delta v + \frac{1}{2}\langle x, \nabla v\rangle,$$

closely related to the Mehler operator.

7.1. **Reverse Poincaré Inequality.** In general, if u is a drift harmonic function, in the sense that $L_\varphi u = 0$, then we have the following reverse Poincaré inequality.

Theorem 7.2. If $L_\varphi u = 0$, then

$$\int_{B_r} |\nabla u|^2 e^{-\varphi} \leq \frac{4}{r^2}\int_{B_{2r}\setminus B_r} u^2 e^{-\varphi}.$$

(Proof.) Let $\eta \geq 0$ be a smooth function with compact support. Then

$$\begin{aligned} 0 &= \int \eta^2 u L_\varphi u e^{-\varphi} \\ &= -\int \left\langle \nabla(\eta^2 u), \nabla u\right\rangle e^{-\varphi} \\ &= -2\int \eta u \langle \nabla\eta, \nabla u\rangle e^{-\varphi} - \int \eta^2|\nabla u|^2 e^{-\varphi}. \end{aligned}$$

Thus by the Cauchy-Schwarz inequality and the AM-GM inequality,

$$\begin{aligned} \int \eta^2|\nabla u|^2 e^{-\varphi} &= -2\int \eta u \langle \nabla\eta, \nabla u\rangle e^{-\varphi} \\ &\leq \frac{1}{2}\int \eta^2|\nabla u|^2 e^{-\varphi} + 2\int u^2|\nabla\eta|^2 e^{-\varphi}, \end{aligned}$$

so

$$\int \eta^2|\nabla u|^2 e^{-\varphi} \leq 4\int u^2|\nabla\eta|^2 e^{-\varphi}.$$

Now by considering η to be 1 on B_r and 0 outside B_{2r} with $|\nabla\eta| \leq \frac{1}{r}$, the conclusion follows. □

Corollary 7.3. If $u \in L^2_\varphi$ (in the sense that $\int u^2 e^{-\varphi} < \infty$) satisfies $L_\varphi u = 0$, then u must be constant.

(Proof.) We have

$$\int_{B_r} |\nabla u|^2 e^{-\varphi} \leq \frac{4}{r^2}\int_{B_{2r}\setminus B_r} u^2 e^{-\varphi} \to 0$$

as $r \to \infty$. □

We would like to look at some slight generalization.

Theorem 7.4. Suppose V is a bounded function. If $L_\varphi u + Vu = 0$, then

$$\int_{B_r} |\nabla u|^2 e^{-\varphi} \le \frac{4}{r^2} \int_{B_{2r}\setminus B_r} u^2 e^{-\varphi} + 2\int_{B_{2r}} V u^2 e^{-\varphi}.$$

(Proof.) For any cut-off function η,

$$\begin{aligned} 0 &= \int \eta^2 u (L_\varphi u + Vu) e^{-\varphi} \\ &= -2\int \eta u \langle \nabla \eta, \nabla u \rangle e^{-\varphi} - \int \eta^2 |\nabla u|^2 e^{-\varphi} + \int \eta^2 V u^2 \end{aligned}$$

so

$$\begin{aligned} \int \eta^2 |\nabla u|^2 e^{-\varphi} &= -2\int \eta u \langle \nabla \eta, \nabla u \rangle e^{-\varphi} + \int \eta^2 V u^2 \\ &\le \frac{1}{2}\int \eta^2 |\nabla u|^2 e^{-\varphi} + 2\int u^2 |\nabla \eta|^2 e^{-\varphi} + \int V \eta^2 u^2 \end{aligned}$$

which implies

$$\int \eta^2 |\nabla u|^2 e^{-\varphi} \le 4\int u^2 |\nabla \eta|^2 e^{-\varphi} + 2\int V \eta^2 u^2.$$

By taking the same cut-off function as in the previous theorem, we get the result. □

Corollary 7.5. If V is bounded, $L_\varphi u + Vu = 0$, and $u \in L^2_\varphi$ in the sense that $\int u^2 e^{-\varphi} < \infty$, then $|\nabla u| \in L^2_\varphi$.

(Proof.) By letting $r \to \infty$ in theorem 7.4, the bound follows. □

7.2. **Eigenfunctions of the Drift Laplacian.** Suppose

$$L_\varphi u + \lambda u = 0.$$

We know if $u \in L^2_\varphi$, then $|\nabla u| \in L^2_\varphi$ by the corollary 7.5. Thus if v is another eigenfunction satisfying

$$L_\varphi v + \mu v = 0$$

with $\mu \neq \lambda$, then

$$-\lambda \int u v e^{-\varphi} = \int L_\varphi u \cdot v e^{-\varphi} = \int L_\varphi v \cdot u e^{-\varphi} = -\mu \int u v e^{-\varphi}$$

so

$$\int u v e^{-\varphi} = 0.$$

7.3. **Maximum Principle.** Another important propety is the maximum principle for the drift Laplacian.

Theorem 7.6. Let Ω be a compact domain in $\mathbb{R}^n$. If $L_\varphi u > 0$ on Ω, then

$$\max_{\Omega} u = \max_{\partial\Omega} u.$$

(Proof.) Suppose $x \in \mathrm{int}\Omega$ satisfied $u(x) = \max u$. Then by calculus we know

$$\nabla u(x) = 0 \text{ and } \Delta u(x) \leq 0,$$

but this implies

$$L_\varphi u(x) = \Delta u(x) - \langle \nabla\varphi, \nabla u \rangle (x) \leq 0,$$

contradicting to the assumption. □

In genera, we have a stronger result.

Theorem 7.7. If $L_\varphi u \geq 0$ on a compact domain Ω, then

$$\max_{\Omega} u = \max_{\partial\Omega} u.$$

(Proof.) Consider $v = e^{\alpha x_1} > 0$. Then $\nabla v = \alpha v \partial_1$ and $\Delta v = \alpha^2 v$, so

$$L_\varphi v = \alpha v \left(\alpha - \frac{\partial\varphi}{\partial x_1} \right).$$

Thus we can take α so large that $L_\varphi v > 0$ since Ω is compact. Then we have

$$L_\varphi(u + \varepsilon v) > 0$$

for all $\varepsilon > 0$. Thus

$$\max_{\Omega}(u + \varepsilon v) = \max_{\partial\Omega}(u + \varepsilon v)$$

for all $\varepsilon > 0$. By taking $\varepsilon \to 0$, the maximum principle follows. □

7.4. **Weighted Energy.** We go back to the equation

$$L_\varphi u + V u = 0$$

with V bounded. It turns out that there is a natural energy functional associated to this equation, which is sort of a weighted version of the Dirichlet energy. Consider

$$E_{\varphi,V}(u) = \int \left(|\nabla u|^2 - V u^2 \right) e^{-\varphi}.$$

This makes sense for $u \in L^2_\varphi$ based on the corollary (7.5).

We write

$$L_V := L_\varphi + V u.$$

An example is when $\varphi = -\frac{|x|^2}{4}$ and $V = \frac{n}{2}$. In this case, the operator is called the Mehler operator, denoted by

$$L_M := \Delta v + \frac{1}{2}\langle x, \nabla v\rangle + \frac{n}{2}u.$$

In this case, we have the Mehler energy

$$E_M(u) := \int \left(|\nabla u|^2 - \frac{n}{2}u^2\right) e^{\frac{|x|^2}{4}}.$$

Lemma 7.8. If there exists $g > 0$ such that $L_V g = L_\varphi g + Vg \leq 0$, then $E_{\varphi,V}(u) \geq 0$ for all $u \in L^2_\varphi$ with $|\nabla u| \in L^2_\varphi$. Moreover, if $E_{\varphi,V}(u) = 0$, then either $u = 0$ or $u = cg$ with $c \neq 0$, in which case $L_V g = 0$.

(Proof.) Since $g > 0$, we consider

$$\begin{aligned} L_\varphi \log g &= \frac{L_\varphi g}{g} - \frac{|\nabla g|^2}{g^2} \\ &\leq \frac{-Vg}{g} - \frac{|\nabla g|^2}{g^2} \\ &= -V - \frac{|\nabla g|^2}{g^2}. \end{aligned}$$

Thus, for any cut-off function η,

$$\int \eta^2 u^2 L_\varphi \log g e^{-\varphi} \leq -\int \eta^2 u^2 V e^{-\varphi} - \int \eta^2 u^2 \frac{|\nabla g|^2}{g^2} e^{-\varphi}.$$

On the other hand, integration by parts give

$$\int \eta^2 u^2 L_\varphi \log g e^{-\varphi} = -2\int \eta u^2 \frac{\langle \nabla \eta, \nabla g\rangle}{g} e^{-\varphi} - 2\int u\eta^2 \frac{\langle \nabla u, \nabla g\rangle}{g} e^{-\varphi}.$$

Combining these we get

$$-2\int \eta u^2 \frac{\langle \nabla \eta, \nabla g\rangle}{g} e^{-\varphi} - 2\int u\eta^2 \frac{\langle \nabla u, \nabla g\rangle}{g} e^{-\varphi} \leq -\int \eta^2 u^2 V e^{-\varphi} - \int \eta^2 u^2 \frac{|\nabla g|^2}{g^2} e^{-\varphi}.$$

This is equivalent to

$$\begin{aligned} -2\int \eta u^2 \frac{\langle \nabla \eta, \nabla g\rangle}{g} e^{-\varphi} &\leq 2\int u\eta^2 \frac{\langle \nabla u, \nabla g\rangle}{g} e^{-\varphi} - \int \eta^2 u^2 V e^{-\varphi} - \int \eta^2 u^2 \frac{|\nabla g|^2}{g^2} e^{-\varphi} \\ &\leq \left(\int \eta^2 |\nabla u|^2 e^{-\varphi} + \int \eta^2 u^2 \frac{|\nabla g|^2}{g^2} e^{-\varphi}\right) - \int \eta^2 u^2 V e^{-\varphi} - \int \eta^2 u^2 \frac{|\nabla g|^2}{g^2} e^{-\varphi} \\ &= \int \eta^2 (|\nabla u|^2 + Vu^2) e^{-\varphi} \end{aligned}$$

where we use the AM-GM inequality.

Intuitively, we would like to choose $\eta = 1$, which seems to imply $0 \leq E_{\varphi,V}(u)$. But in this case the integral may converge in general.

First by putting $\eta = 1$, we prove that $E_{\varphi,V}(u) \geq 0$ if u has compact support, and the energy vanishes if and only if $\nabla \log \frac{u}{g} = 0$. In general, we may take η approximating the identity (i in B_r and linearly decreasing on $B_{2r} \setminus B_r$). In this case, if $E_{\varphi,V}(u) = 0$, we can prove by a limiting argument that

$$\left| u\frac{\nabla g}{g} - \nabla u \right| = 0.$$

If $u = cg$ for some $c \neq 0$, then

$$0 = E_{\varphi,V}(u) = \int (|\nabla u|^2 - Vu^2)e^{-\varphi} = -\int \left(uL_\varphi u + Vu^2 \right) e^{-\varphi}$$

since $u, \nabla u \in L^2_\varphi$. □

Note that no integrability condition on g is assumed in this lemma.

The case we are most interested in is the Mehler operator, that is,

$$L_M u = e^{-\frac{|x|^2}{4}} \operatorname{div}(e^{\frac{|x|^2}{4}} \nabla u) + \frac{n}{2} u = \Delta u + \frac{1}{2} \langle x, \nabla u \rangle + \frac{n}{2} u.$$

If we take $g = e^{-\frac{|x|^2}{4}} > 0$, then

$$L_M g = e^{-\frac{|x|^2}{4}} \operatorname{div}\left(-\frac{x}{2} \right) + \frac{n}{2} e^{-\frac{|x|^2}{4}} = 0.$$

Thus by the lemma,

$$E_M(u) = \int \left(|\nabla u|^2 - \frac{n}{2} u^2 \right) e^{\frac{|x|^2}{4}} \geq 0$$

if u and ∇u are in $L^2_{-\frac{|x|^2}{4}}$, and $E_M(u) = 0$ if and only if $u = ce^{-\frac{|x|^2}{4}}$.

7.5. **Relations to the Heat Equation and the Mehler Flow.** Now we go back to the heat equation. Recall the scaling property of the heat equation. That is, if u is a solution to the heat equation, then

$$v_c(x,t) := u(cx, c^2 t)$$

is also a solution. Also, we established the central limit theorem (6.1). That is, if u is a solution with mild growth and $u(\cdot, 0) = u_0$, then

$$v(x,t) = t^{\frac{n}{2}} u(\sqrt{t} x, t) \to Ce^{-\frac{|x|^2}{4}}.$$

as $t \to \infty$. If we consider

$$w(x,s) := v(x, e^s) = e^{\frac{n}{2}s} u(e^{\frac{s}{2}} x, e^s),$$

it satisfies the Mehler flow equation

$$\partial_s w = L_M w = \Delta w + \frac{1}{2} \langle x, \nabla w \rangle + \frac{n}{2} w. \tag{7.9}$$

To see this, note

$$\partial_s w = \frac{n}{2} w + \frac{1}{2} e^{\frac{n}{2}s} \left\langle e^{\frac{s}{2}} x, \nabla u \right\rangle + e^{\frac{n}{2}s} u_t \cdot e^s.$$

On the other hand,

$$\nabla w = e^{\frac{n}{2}s} e^{\frac{s}{2}} \nabla u$$

and

$$\Delta w = e^{\frac{n}{2}s} e^{s} \Delta u,$$

which implies

$$\begin{aligned}
\partial_s w &= \frac{n}{2} w + \frac{1}{2} e^{\frac{n}{2}s} \left\langle e^{\frac{s}{2}} x, \nabla u \right\rangle + e^{\frac{n}{2}s} \Delta u \cdot e^{s} \\
&= \frac{n}{2} w + \frac{1}{2} \langle x, \nabla w \rangle + \Delta w = L_M w.
\end{aligned}$$

In summary, if u_0 has compact, then we could construct a solution to the heat equation with its initial data u_0, and construct such a solution w to the Mehler flow equation (7.9). In fact, as $s \to \infty$, $w(s,t)$ will converge to a solution g to the Mehler equation $L_M g = 0$.

7.6. **Ornstein-Uhlenbeck Operator from the heat equation.** Suppose $(\partial_t - \Delta u) = 0$. We may want to start the time from any t_0, but we may assume $t_0 = 0$ by translation. Then take the following rescaling and set

$$v(x,t) = u(\sqrt{|t|}x, t)$$

and

$$w(x,s) = u\left(\sqrt{e^{-s}} x, -e^{-s}\right).$$

We put the minus sign to make the time to approach 0. Then we see that

$$\partial_s w = -\frac{1}{2} e^{-\frac{s}{2}} \langle x, \nabla u \rangle + e^{-s} u_t,$$

$$\nabla w = e^{-\frac{s}{2}} \nabla u,$$

$$\Delta w = e^{-s} \Delta u,$$

and hence

$$\partial_s w = \Delta w - \frac{1}{2} \langle x, \nabla w \rangle = L_{OU} w,$$

where L_{OU} is the Ornstein-Uhlenbeck operator (7.1).

In general, for $u \colon \mathbb{R}^n \times [a,b] \to \mathbb{R}$. If we consider a general energy

$$E(u) = \int \left(|\nabla u|^2 - V u^2 \right) e^{-\varphi},$$

then we get

$$\begin{aligned}\frac{d}{dt}E(u) &= \frac{d}{dt}\int \left(|\nabla u|^2 - Vu^2\right) e^{-\varphi} \\ &= 2\int (\langle \nabla u, \nabla u_t\rangle - Vuu_t)e^{-\varphi} \\ &= -2\int (L_\varphi u \cdot u_t - Vuu_t)e^{-\varphi} \\ &= -2\int u_t Lu \cdot e^{-\varphi}\end{aligned}$$

if we define $L = L_\varphi + V$, where we need some assumption on decays to integrate them by parts without boundary terms. Thus if u satisfies

$$\partial_t = Lu, \tag{7.10}$$

then

$$\frac{d}{dt}E(u) = -2\int |Lu|^2 e^{-\varphi}.$$

That is to say, up to a constant, (7.10) is the negative gradient flow of E.

8. **Central Limit Theorem from the Variational Viewpoint**

The variational point of view is opposed to the representation formula (though in this case it does not prove anything new here).

First we note that the Mehler flow sits inside a larger class of examples. We define

$$Lu := \Delta u - \operatorname{div}(u\nabla\varphi) = \Delta u - \langle \nabla u, \nabla \varphi\rangle - u\Delta\varphi = L_\varphi u - u\Delta\varphi.$$

That is, we choose $V = -\Delta\varphi$. In particular, if we take $\varphi = -\frac{|x|^2}{4}$, then

$$Lu = L_{\frac{|x|^2}{4}} u + \frac{n}{2}u$$

is the Mehler operator L_M. In general,

$$\int Lu = 0$$

if there is no boundary term. Thus if u satisfies $\partial_t u = Lu$, then

$$\frac{d}{dt}\int u = \int Lu = 0.$$

In the special case that $L = L_M$, we know $\int u$ is conserved. Suppose $u\colon \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfies $\partial_t u = L_M u$. Assume that

$$\int u^2 e^{\frac{|x|^2}{4}} \text{ and } \int |\nabla u|^2 e^{\frac{|x|^2}{4}} < \infty \tag{8.1}$$

for each time (not in a uniform sense). Then

$$\begin{aligned}
\frac{d}{dt}\int u^2 e^{\frac{|x|^2}{4}} &= 2\int u_t u e^{\frac{|x|^2}{4}} \\
&= 2\int u L_M u e^{\frac{|x|^2}{4}} \\
&= 2\int u\left(L_{-\frac{|x|^2}{4}} u + \frac{n}{2}u\right) e^{\frac{|x|^2}{4}} \\
&= -2\int \left(|\nabla u|^2 - \frac{n}{2}u^2\right) e^{\frac{|x|^2}{4}} = -2E_M(u).
\end{aligned}$$

We already know that $E_M(u) \geq 0$, so we know that the L^2 weighted norm $\frac{d}{dt}\int u^2 e^{\frac{|x|^2}{4}}$ is decreasing. Besides, we already know that

$$\frac{d}{dt}E_M(u) = -2\int |L_M u|^2 e^{\frac{|x|^2}{4}}$$

so $E_M(u)$ is also decreasing. Hence we indeed have a uniform bound on $\int |\nabla u|^2 e^{\frac{|x|^2}{4}}$.

Proposition 8.2. As $t \to \infty$, $E_M(u(\cdot,t)) \to 0$.

(Proof.) We know

$$\frac{d}{dt}\int u^2 e^{\frac{|x|^2}{4}} = -2E_M(u).$$

Hence we could take a sequence $t_i \to \infty$ such that $E(u(\cdot,t_i)) \to 0$. Since the energy in monotone, we get the conclusion. □

What we aiming at is to show that $u(\cdot,t)$ converges to $ce^{-\frac{|x|^2}{4}}$. In practice, we would like to prove that $u(\cdot,t) \to g$ with $E_M(g) = 0$. However, we are afraid that the convergence is weak.

To deal with this issue, note that $E_M(u)$ is bounded from below, in the sense that

$$\int |\nabla u - u\log g|^2 e^{\frac{|x|^2}{4}} \leq E(u)$$

where $g(x) = e^{-\frac{|x|^2}{4}}$, $\log g = -\frac{|x|^2}{4}$ and $\nabla \log g = -\frac{x}{2}$. Thus

$$E_M(u) \geq \int \left|\nabla u + \frac{x}{2}u\right|^2 e^{\frac{|x|^2}{4}}.$$

Thus we have a uniform bound (in t) for

$$\int \left|\nabla u + \frac{x}{2}u\right|^2 e^{\frac{|x|^2}{4}}.$$

Since

$$\frac{|x|^2}{4}u^2 \leq \frac{1}{2}|\nabla u|^2 + \frac{1}{2}\left|\nabla u + \frac{x}{2}u\right|^2$$

by the inequality $(a+b)^2 \le 2a^2 + 2b^2$, we thus obtain a unifrom bound for

$$\int |x|^2 u^2 e^{\frac{|x|^2}{4}}.$$

Then we can estimate that $\int_{B_R} u^2 e^{\frac{|x|^2}{4}}$ must be almost all $\int u^2 e^{\frac{|x|^2}{4}}$. In fact, if we say

$$\int |x|^2 u^2 e^{\frac{|x|^2}{4}} \le C,$$

then

$$R^2 \int u^2 e^{\frac{|x|^2}{4}} \le \int_{bbR^n \setminus B_R} |x|^2 u^2 e^{\frac{|x|^2}{4}} \le C,$$

so

$$\int_{\mathbb{R}^n \setminus B_R} u^2 e^{\frac{|x|^2}{4}} \le \frac{C}{R^2}.$$

Then just take R sufficiently large. In conclusion, we see that $\int u^2 e^{\frac{|x|^2}{4}}$ in concentrated on a large ball. This tells us that it will not go to infinity.

Recall our assumption (8.1), based on which we get uniform bounds for them and $\int |x|^2 u^2 e^{\frac{|x|^2}{4}}$. Then with the proposition 8.2, we could get a uniform bound for the derivatives of u (in space). Then the Arzela-Ascoli theorem gives us a convergenct subsequence, say $u(\cdot, t_i) \to v$ with $t_i \to \infty$ as $i \to \infty$. By the convergence of the energy, we know $E_M(v) = 0$, which means that v is Gaussian. Since $\int u$ is constant in time, we know that $\int v = \int u_0 =: c$, which is independent of the subsequence derived from the Arzela-Ascoli theorem.

We remark that this method is more robust, in the sense that it could be applied to fairly large cases of different PDE.

Recall what we did in the section 6.1. We could give a different viewpoint now also. For non-negative g and g with compact support (in $\mathbb{R}^n$), consider $f(x,t)$ and $g(x,t)$ satisfying the Mehler flow equation. If we take $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, then it follows from direct computation that

$$(\partial_t - L_M) f^{\frac{1}{p}} g^{\frac{1}{q}} = \frac{1}{pq} \left| \nabla \log \frac{f}{g} \right|^2 f^{\frac{1}{p}} g^{\frac{1}{q}}.$$

Thus

$$\partial_t \int f^{\frac{1}{p}} g^{\frac{1}{q}} = \int (\partial_t - L_M) f^{\frac{1}{p}} g^{\frac{1}{q}} + \int L_M (f^{\frac{1}{p}} g^{\frac{1}{q}}) = \frac{1}{pq} \int \left| \nabla \log \frac{f}{g} \right|^2 f^{\frac{1}{p}} g^{\frac{1}{q}} \ge 0.$$

Since f and g converge to $c_f (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}$ and $c_g (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}$ with $c_f = \int f$ and $c_g = \int g$, the monotonicity proves the Hölder inequality.

9. **Shannon Entropy and Fisher Information**

Let u be a positive function. Then we define

$$S(u) := -\int u \log u$$

if it is finite, called the Shannon entropy. If u satisfies the heat equation, we could consider

$$S(t) := S(u(\cdot, t)) = -\int u(\cdot, t) \log u(\cdot, t),$$

and we get

$$\begin{aligned} \frac{d}{dt} S(t) &= -\int u_t \log u - \int u \frac{u_t}{u} \\ &= -\int \Delta u \log u - \int \Delta u \\ &= \int \frac{|\nabla u|^2}{u} - 0 \\ &= \int |\nabla \log u|^2 u, \end{aligned}$$

called the Fisher information. There are also discrete versions of these notions. Here we recall the classic Bochner formula, which could be derived by direct computation.

Lemma 9.1. For a C^2 function v, we have

$$\frac{1}{2} \Delta |\nabla v|^2 = \langle \nabla v, \nabla \Delta v \rangle + |\mathrm{Hess}_v|^2.$$

Also, we can derive

$$\frac{1}{2} \partial_t |\nabla v|^2 = \sum_i \frac{\partial^2 v}{\partial x_i \partial t} \frac{\partial v}{\partial x_i} = \langle \nabla v_t, \nabla v \rangle .$$

Combining these, if v solves the heat equation, we then have

$$\frac{1}{2} (\partial_t - \Delta) |\nabla v|^2 = -|\mathrm{Hess}_v|^2.$$

Going bact to u, since

$$(\partial_t - \Delta) |\nabla \log u|^2 = -2 |\mathrm{Hess}_{\log u}|^2 + 2 \langle \nabla (\partial_t - \Delta) \log u, \nabla \log u \rangle ,$$

we know

$$\begin{aligned} (\partial_t - \Delta)(|\nabla \log u|^2 u) &= (\partial_t - \Delta) |\nabla \log u|^2 u + |\nabla \log u|^2 (\partial_t - \Delta) u - 2 \left\langle \nabla |\nabla \log u|^2, \nabla u \right\rangle \\ &= -2 |\mathrm{Hess}_{\log u}|^2 u + 2 \langle \nabla (\partial_t - \Delta) \log u, \nabla \log u \rangle u - 2 \left\langle \nabla |\nabla \log u|^2, \nabla u \right\rangle . \end{aligned}$$

since u solves the heat equation, which also implies

$$(\partial_t - \Delta) \log u = |\nabla \log u|^2. \tag{9.2}$$

In conclusion, we get

$$
\begin{aligned}
(\partial_t - \Delta)(|\nabla \log u|^2 u) &= -2|\mathrm{Hess}_{\log u}|^2 u + 2\left\langle \nabla |\nabla \log u|^2, \nabla \log u \right\rangle u - 2\left\langle \nabla |\nabla \log u|^2, \nabla u \right\rangle \\
&= -2|\mathrm{Hess}_{\log u}|^2 u.
\end{aligned}
$$

In particular,

$$\int \partial_t(|\nabla \log u|^2 u) = -2\int |\mathrm{Hess}_{\log u}|^2 u.$$

Thus if we define

$$F(t) = F(u(\cdot, t)) = \int |\nabla \log u|^2 u,$$

then we get

$$F'(t) = -2\int |\mathrm{Hess}_{\log u}|^2 u.$$

These quantities paly important roles in many areas, especially in the information theory.

9.1. W**-functional.** It is reasonable to look at a combination $f + tf'$ (for example, thinking of f as a polynomial). Thus we consider

$$W(t) := S(t) + tS'(t) - \frac{n}{2}\log t = S(t) + tF(t) - \frac{n}{2}\log t \tag{9.3}$$

where the last term is for normalization. This functional W is discovered by Perelman, called Perelman's W-functional.

Example. Suppose

$$u(x,t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}.$$

Then

$$\log u = -\frac{n}{2}\log(4\pi t) - \frac{|x|^2}{4t},$$

and

$$S(u) = \frac{n}{2}\log(4\pi t)\int u + \int \frac{|x|^2}{4t} u = \frac{n}{2}\log(4\pi t) + \int \frac{|x|^2}{4t} u.$$

We need to find out the second term. Consider a change of variable $y = \frac{x}{\sqrt{t}}$, so if we define

$$v(x,t) = \frac{|x|^2}{4t} u(x,t),$$

then

$$v(y,t) = \frac{|y|^2}{4} e^{-\frac{|y|^2}{4}}$$

so

$$\int \frac{|x|^2}{4t} u = \int \frac{|y|^2}{4} (4\pi)^{-\frac{n}{2}} e^{-\frac{|y|^2}{4}}.$$

We know if we let $\varphi = -\frac{|x|^2}{4}$, then

$$L_\varphi = \Delta - \frac{1}{2}\left\langle x, \nabla(\cdot)\right\rangle,$$

so

$$L_\varphi |x|^2 = 2n - |x|^2.$$

Therefore

$$0 = \int L_\varphi |x|^2 e^{-\frac{|x|^2}{4}} = \int (2n - |x|^2) e^{-\frac{|x|^2}{4}}$$

and hence

$$\int \frac{|x|^2}{4} (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}} = \frac{2n}{4} \int (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}} = \frac{n}{2}.$$

Putting this back to $S(u)$, we get

$$S(u) = \frac{n}{2} \log(4\pi t) + \frac{n}{2}.$$

This suggests the normalization in (9.3).

We can compute

$$\begin{aligned} W' &= \Big(S + tF - \frac{n}{2}\log t\Big)' \\ &= 2F + tF' - \frac{n}{2t} \\ &= 2\int |\nabla \log u|^2 u - 2t \int |\mathrm{Hess}_{\log u}|^2 u - \frac{n}{2t}. \end{aligned}$$

Notice if we let

$$A = \mathrm{Hess}_{\log u} \text{ and } B = \frac{1}{2t}\delta_{ij},$$

then

$$\left|\mathrm{Hess}_{\log u} + \frac{1}{2t}\delta_{ij}\right|^2 = \mathrm{tr}(A+B)^2 = |\mathrm{Hess}_{\log u}|^2 + \frac{1}{t}\Delta \log u + \frac{n}{4t^2}.$$

Therefore, after assuming $\int u = 1$ (by normalization since it is constant in time),

$$\begin{aligned} W' &= -2t \int \left(|\mathrm{Hess}_{\log u}|^2 + \frac{n}{4t^2} - \frac{|\nabla \log u|^2}{t} \right) u \\ &= -2t \int \left|\mathrm{Hess}_{\log u} + \frac{1}{2t}\delta_{ij}\right|^2 u + 2t \int \left(\frac{\Delta \log u}{t} + \frac{|\nabla \log u|^2}{t} \right) u. \end{aligned}$$

By (9.2), we have

$$\int |\nabla \log u|^2 u = \int ((\partial_t - \Delta)\log u) u = \int u_t - \int (\Delta \log u) u = -\int (\Delta \log u) u$$

so the last term in W' vanishes, and finally we derive

$$W' = -2t \int \left|\mathrm{Hess}_{\log u} + \frac{1}{2t}\delta_{ij}\right|^2 u. \tag{9.4}$$

In particular, tF is constant for the fundamental solution since in this case, $S' = \frac{n}{2t}$.

9.2. **Log Sobolev Inequality.** We recall the we use the methods related to the heat equation to derive some functional inequalities. We use 1. the central limit theorem, and 2. the monotonicity. We will use the monotonicity of the W-functional to prove the following log Sobolev inequality.

Theorem 9.5. Suppose $w > 0$ satisfies $(2\pi)^{-\frac{n}{2}} \int w^2 e^{-\frac{|x|^2}{2}} = 1$. Then
$$\int w^2 \log w \cdot e^{-\frac{|x|^2}{2}} \leq \int |\nabla w|^2 e^{-\frac{|x|^2}{2}}.$$

This was first proven by L. Gross, but it turns out that there is a nice proof using the W-functional.

Note that in the usual Sobolev inequality, the power p on the left hand side should be larger than 2. A downside is that it depends on the dimension of the space, which is not uniform. It is not a case in the log Sobolev inequality. At the same time, we get a log term.

(Proof.) The first thing to do is to reformulate the inequality. If we consider
$$u := (2\pi)^{-\frac{n}{2}} w^2 e^{-\frac{-|x|^2}{2}},$$
then we get $u > 0$, $\int u = 1$, and
$$w = (2\pi)^{\frac{n}{4}} \sqrt{u} e^{\frac{|x|^2}{4}}.$$
Thus
$$\nabla w = \frac{\nabla u}{2\sqrt{u}} (2\pi)^{\frac{n}{4}} e^{\frac{|x|^2}{4}} + \frac{x}{2} w$$
and
$$|\nabla w|^2 = \frac{|\nabla u|^2}{4u} (2\pi)^{\frac{n}{2}} e^{\frac{|x|^2}{2}} + \frac{|x|^2}{4} w^2 + \frac{1}{2} \cdot \frac{w}{\sqrt{u}} (2\pi)^{\frac{n}{2}} e^{\frac{|x|^2}{2}} \langle x, \nabla u \rangle .$$
Therefore,
$$|\nabla w|^2 e^{\frac{-|x|^2}{2}} = \frac{|\nabla u|^2}{4u} (2\pi)^{\frac{n}{2}} + \frac{|x|^2}{4} w^2 e^{\frac{-|x|^2}{2}} + \frac{1}{2} \cdot \frac{w}{\sqrt{u}} (2\pi)^{\frac{n}{2}} \langle x, \nabla u \rangle .$$
On the other hand, by noting that
$$\int \langle \nabla u, x \rangle = \int \left\langle \nabla u, \nabla \frac{|x|^2}{2} \right\rangle = -\frac{1}{2} \int u \Delta |x|^2 = -n \int u = -n,$$
we could translate the original inequality we want to show to
$$n + \frac{n \log(2\pi)}{2} \leq - \int u \log u + \frac{1}{2} \int \frac{|\nabla u|^2}{u}. \tag{9.6}$$
If u is a solution to the heat equation on $\mathbb{R}^n \times [\frac{1}{2}, \infty)$, then
$$W\left(\frac{1}{2}\right) = - \int u \left(\cdot, \frac{1}{2}\right) \log u \left(\cdot, \frac{1}{2}\right) + \frac{1}{2} \int \frac{|\nabla u|^2}{u} + \frac{n}{2} \log \frac{1}{2}.$$
Thus (9.6) is equivalent to
$$n + \frac{n \log(2\pi)}{2} \leq W\left(\frac{1}{2}\right) + \frac{n}{2} \log \frac{1}{2}.$$

We know that W is decreasing by (9.4). By the central limit theorem, we know

$$W(t) \to \text{ the } W\text{-functional of } (4\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4}}.$$

Hence we get the desired inequality. $\square$

9.3. **Renyi Entropy.** We could define another notion of entropy by

$$R_p := \frac{1}{1-p} \log \int u^p$$

for a solution $u > 0$ to the heat equation and $p \neq 1$. Then we have

$$\begin{aligned} \partial_t R_p &= \frac{p}{1-p} \frac{\partial_t \int u^p}{\int u^p} \\ &= \frac{p}{1-p} \frac{\int u^{p-1} \Delta u}{\int u^p} \\ &= \frac{p(p-1)}{1-p} \frac{\int |\nabla u|^2 u^{p-2}}{\int u^p}, \end{aligned}$$

which proves that R_p is also monotone. If for a fixed function u, we think of R_p as a function of p, say $f(p) = \log \int u^p$, note

$$(u^p)' = \log u \cdot u^p$$

so

$$f'(p) = \frac{\int \log u \cdot u^p}{\int u^p}.$$

If $\int u = 1$ after normalization, we have

$$f'(1) = \int u \log u,$$

which is exactly the Shannon entropy (up to a sign). In fact, we could say

$$\lim_{p \to 1} R_p = -\lim_{p \to 1} \frac{\log \int u^p - \log \int u}{p-1} = -f'(1) = -\int u \log u = S.$$

10. Differential Harnack Inequality

10.1. **Li-Yau estimate.** The first inequality in this section was established by Li and Yau [LY86].

Theorem 10.1 (Li-Yau)**.** Suppose a positive C^2 function $u \colon \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ solves the heat equation. Then

$$\frac{|\nabla u|^2}{u^2} - \frac{u_t}{u} \leq \frac{n}{2t}.$$

Example. Let

$$u(x,t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}.$$

Then

$$\log u = -\frac{n}{2}\log(4\pi t) - \frac{|x|^2}{4t},$$
$$\nabla \log u = -\frac{x}{2t},$$

and

$$(\log u)_t = -\frac{n}{2t} + \frac{|x|^2}{4t}.$$

Then theorem 10.1, which is equivalent to

$$|\log u|^2 - (\log u)_t \le \frac{n}{2t},$$

becomes

$$\frac{|x|^2}{4t^2} + \frac{n}{2t} - \frac{|x|^2}{4t^2} = \frac{n}{2t}.$$

Thus the equality of the Li-Yau inequality is achieved when u is the fundamental solution, so it is sharp.

(Proof of theorem 10.1.) Write $v = \log u$. Then

$$(\partial_t - \Delta)v = \frac{u_t}{u} - \frac{\Delta u}{u} + \frac{|\nabla u|^2}{u^2} = \frac{(\partial_t - \Delta)u}{u} + |\nabla v|^2 = |\nabla v|^2.$$

If we define

$$F(x,t) = t\left(\frac{|\nabla u|^2}{u^2} - \frac{u_t}{u}\right),$$

then we can derive

$$(\partial_t - \Delta)F \le \frac{F}{t} + 2\langle \nabla F, \nabla v\rangle - \frac{2}{nt}F^2. \tag{10.2}$$

To see (10.2), note that

$$\begin{aligned}
(\partial_t - \Delta)F &= (\partial_t - \Delta)\big(t(|\nabla v|^2 - v_t)\big)\\
&= |\nabla v|^2 - v_t + t\partial_t|\nabla v|^2 - t\Delta|\nabla v|^2 - t(\partial_t - \Delta)v_t\\
&= \frac{F}{t} + t(\partial_t - \Delta)|\nabla v|^2 - t((\partial_t - \Delta)v)_t\\
&= \frac{F}{t} + t(\partial_t - \Delta)|\nabla v|^2 - t(|\nabla v|^2)_t\\
&= \frac{F}{t} + t(\partial_t - \Delta)|\nabla v|^2 - 2t\langle \nabla v_t, \nabla v\rangle\\
&= \frac{F}{t} + t\big(2\langle \nabla v_t, \nabla v\rangle - 2|\mathrm{Hess}_v|^2 - 2\langle \nabla\Delta v, \nabla v\rangle\big) - 2t\langle \nabla v_t, \nabla v\rangle\\
&= \frac{F}{t} + t\big(2\langle \nabla(\partial_t - \Delta)v, \nabla v\rangle - 2|\mathrm{Hess}_v|^2\big) - 2t\langle \nabla v_t, \nabla v\rangle\\
&= \frac{F}{t} + t\big(2\langle \nabla|\nabla v|^2, \nabla v\rangle - 2|\mathrm{Hess}_v|^2\big) - 2t\langle \nabla v_t, \nabla v\rangle\\
&= \frac{F}{t} + 2\langle \nabla F, \nabla v\rangle - 2t|\mathrm{Hess}_v|^2
\end{aligned}$$

where we use the Bochner formula

$$\Delta|\nabla w|^2 = 2|\mathrm{Hess}_w|^2 + 2\langle \nabla\Delta w, \nabla w\rangle .$$

Then (10.2) follows since by the Cauchy-Schwarz inequality, we have

$$|\mathrm{Hess}_v|^2 = \mathrm{tr}(\mathrm{Hess}_v \cdot \mathrm{Hess}_v) \geq \frac{1}{n}\mathrm{tr}\mathrm{Hess}_v = \frac{1}{n}(\Delta v)^2$$

and we know

$$-\Delta v = |\nabla v|^2 - v_t = \frac{F}{t}.$$

Remember that we want to show $F \leq \frac{n}{2}$. Note that $F|_{t=0} = 0$. Now suppose F achieves its maximum on $\mathbb{R}^n \times [0,T]$ at some (x_0,t_0). Then at (x_0,t_0), we have

$$\nabla F = 0, \Delta F \leq 0, \text{ and } \partial_t F \geq 0.$$

Thus $(\partial_t - \Delta)F \geq 0$ at (x_0,t_0). Then (10.2) implies

$$0 \leq (\partial_t - \Delta)F \leq \frac{F}{t} + 2\langle \nabla F, \nabla v\rangle - \frac{2}{nt}F^2 = \frac{F}{t} - \frac{2}{nt}F^2 = \frac{F}{t}\left(1 - \frac{2}{n}F\right)$$

at (x_0,t_0). Thus either $\frac{F}{t} \leq 0$ or $1 - \frac{2}{n}F \geq 0$. Both of them imply $F \leq \frac{n}{2}$.

In general, when the maximum is not achieved, some cut-off functions need to be introduced, causing some lower-order error terms, which could be addressed in a straightforward manner in view of (10.2). □

Corollary 10.3 (Harnack inequality)**.** If u is a positive solution to the heat equation, then for $t > s > 0$ and $x, y \in \mathbb{R}^n$,

$$u(x,t) \geq \left(\frac{s}{t}\right)^{\frac{n}{2}} e^{-\frac{|x-y|^2}{4(t-s)}} u(y,s).$$

Intuitively, the heat could not completely disappear after a finite time if it initially exists, and a reference point like this corollary could give us a rough lower bound. However, this does not say anything if we want to compare u at the same time.

(Proof.) Again set $v = \log u$. For $r \in [0, t-s]$, consider

$$w(r) := v\left(x + \frac{r}{t-s}(y-x), t-r\right),$$

which is $v(x,t)$ if $r = 0$ and $v(y,s)$ if $r = t-s$. Note

$$w' = \left\langle \nabla v, \frac{y-x}{t-s}\right\rangle - v_t \leq |\nabla v|^2 + \frac{|y-x|^2}{4(t-s)^2} - v_t,$$

which implies

$$w(t-s) - w(0) \leq \int_0^{t-s}\left(\frac{n}{2(t-r)} + \frac{|y-x|^2}{4(t-s)^2}\right)dr = \frac{n}{2}\log\frac{t}{s} + \frac{|y-x|^2}{4(t-s)}$$

by theorem 10.1, and the conclusion follows. □

10.2. **Hamilton's Harnack inequality.** After Li and Yau's result in 1986, Hamilton observed in 1993 that theorem 10.1 is a special case of a much more general inequality. The reference is [?Ham93].

Theorem 10.4 (Hamilton's matrix differential Harnack inequality)**.** If u is a positive solution to the heat equation, then

$$\mathrm{Hess}_{\log u} \geq -\frac{\delta_{ij}}{2t}$$

in the matrix sense. i.e., $\mathrm{Hess}_{\log u} + \frac{\delta_{ij}}{2t}$ is a semi-positive definite matrix.

Note that after taking the trace of the inequality in theorem 10.4, we get theorem 10.1. Also, since $\mathrm{Hess}_{|x|^2} = 2\delta_{ij}$, it also implies

$$\mathrm{Hess}_{\log u} \geq -\frac{1}{4t}\mathrm{Hess}_{|x|^2},$$

i.e.

$$\mathrm{Hess}_{\log\left(ue^{\frac{|x|^2}{4t}}\right)} \geq 0$$

that is to say, $f(x,t) := \log\left(ue^{\frac{|x|^2}{4t}}\right)$ is convex. Thus, we can derive

$$f((1-s)x + s(y-x), t) \leq (1-s)f(x,t) + sf(y,t)$$

for $s \in [0,1]$, which means

$$\log\left(u((1-s)x+sy,t)\cdot e^{\frac{|(1-s)x+sy|^2}{4t}}\right) \leq (1-s)\log\left(u(x,t)\cdot e^{\frac{|x|^2}{4t}}\right) + s\log\left(u(y,t)\cdot e^{\frac{|y|^2}{4t}}\right).$$

Thus we get

$$u((1-s)x+sy)\cdot e^{\frac{|(1-s)x+sy|^2}{4t}} \leq \left(u(x,t)e^{\frac{|x|^2}{4t}}\right)^{1-s}\left(u(y,t)e^{\frac{|y|^2}{4t}}\right)^{s}.$$

This assists us to compare the information of u at the same time (which is, infact, sharp), which is not accessible from Li and Yau's theorem 10.1.

(Proof of theorem 10.4.) For convenience as we have done, define $v = \log u$. Consider

$$M := \mathrm{Hess}_v + \frac{1}{2t}\delta_{ij},$$

so our goal is to prove M is semi-positive definite. Note that as $t \to 0^+$ (or at least for t very small), M is positive (if we assume Hess_v is bounded in space), so we want to see this positivity is in some sense preserved. Therefore we will use a parabolic matrix maximum principle (c.f. theorem 10.8).

Componentwise, we have $M_{ij} = v_{ij} + \frac{1}{2t}\delta_{ij}$. It turns out that

$$(\partial_t - \Delta)(uM_{ij}) = 2uM_{ik}M_{kj} - \frac{2u}{t}M_{ij}. \tag{10.5}$$

To see (10.5), first note that

$$\begin{aligned}(\partial_t - \Delta)(uv_{ij}) &= v_{ij}(\partial_t - \Delta)u + u(\partial_t - \Delta)v_{ij} - 2u_k v_{ijk} \\ &= u(\partial_t - \Delta)v_{ij} - 2u_k v_{ijk} \\ &= u\left((\partial_t - \Delta)v\right)_{ij} - 2u_k v_{ijk} \\ &= u\left(|\nabla v|^2\right)_{ij} - 2u_k v_{ijk} \\ &= u \cdot (2v_{ik}v_{jk} + 2v_k v_{ijk}) - 2u_k v_{ijk} = 2uv_{ik}v_{jk}\end{aligned}$$

since $(\partial_t - \Delta)v = (\partial_t - \Delta)\log u = \frac{(\partial_t - \Delta)u}{u} + |\nabla v|^2$ and $uv_k = u_k$. On the other hand,

$$(\partial_t - \Delta)\left(\frac{u}{2t}\delta_{ij}\right) = \frac{(\partial_t - \Delta)u}{2t}\delta_{ij} - \frac{u}{2t^2}\delta_{ij} = -\frac{u}{2t^2}\delta_{ij}.$$

Combining these, we get

$$(\partial_t - \Delta)(uM_{ij}) = 2uv_{ik}v_{jk} - \frac{u}{2t^2}\delta_{ij}.$$

To identify this in terms of M, note that

$$\begin{aligned}2uM_{ik}M_{kj} - \frac{2u}{t}M_{ij} &= 2u\left(v_{ik} + \frac{1}{2t}\delta_{ik}\right)\left(v_{jk} + \frac{1}{2t}\delta_{jk}\right) - \frac{2u}{t}\left(v_{ij} + \frac{1}{2t}\delta_{ij}\right) \\ &= \left(2uv_{ik}v_{jk} + \frac{2u}{4t^2}\delta_{ij} + \frac{2u}{2t}v_{ij}\right) + \left(\frac{2u}{2t}v_{ij} - \frac{2u}{t}v_{ij} - \frac{u}{t^2}\delta_{ij}\right) \\ &= 2uv_{ik}v_{jk} - \frac{u}{2t^2}\delta_{ij}.\end{aligned}$$

Thus (10.5) follows. In matrix forms, it could be written as

$$(\partial_t - \Delta)(uM) = 2uM^2 - \frac{2u}{t}M. \tag{10.6}$$

Hence the conclusion follows from the matrix maximum principle (theorem 10.8 in the next section). More details are duscussed after the proof of theorem 10.8. □

10.3. **Matrix Maximum Principle.** We will see a toy case of the matrix maximum principle first.

Proposition 10.7. Let $Q\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be a differentiable function with a property that $Q(0, x, t) \geq 0$ for all $(x, t) \in \mathbb{R}^n \times \mathbb{R}$. If a C^2 function $u\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ satisfies

$$(\partial_t - \Delta)u(x, t) > Q(u(x, t), x, t)$$

with $u(\cdot, 0) > 0$, then $u > 0$.

(Proof.) Suppose not. i.e., we can take the first time t_0 such that $u(x_0, t_0) = 0$ for some $x_0 \in \mathbb{R}^n$. Thus the infimum of u is achieved at x_0, so

$$\Delta u(x_0, t_0) \geq 0.$$

Therefore,

$$\partial_t u(x_0, t_0) > Q(u(x_0, t_0), x_0, t_0) + \Delta u(x_0, t_0) > Q(0, x_0, t_0) \geq 0,$$

which is impossible since t_0 is the first time such that u vanishes somewhere. □

Next, we consider the space $\mathrm{Sym}_{n\times n}(\mathbb{R})$ of all symmetric $n \times n$ real matrices.

Theorem 10.8. Let $Q\colon \mathrm{Sym}_{n\times n}(\mathbb{R}) \times \mathbb{R}^n \times [0,\infty) \to \mathrm{Sym}_{n\times n}(\mathbb{R})$ be a differentiable function with the null vector condition, in the sense that $\langle Q(B,x,t)v, v\rangle \geq 0$ for all B and v satisfying $Bv = 0$. If a C^2 function $A\colon \mathbb{R}^n \times \mathbb{R} \to \mathrm{Sym}_{n\times n}(\mathbb{R})$ satisfies

$$(\partial_t - \Delta)A(x,t) \geq Q(A(x,t),x,t)$$

with $A(\cdot, 0) \geq 0$ (semi-positive definite), then $A \geq 0$.

(Proof sketch.) First we assume $(\partial_t - \Delta)A > Q(A,x,t)$. Using the same idea, take the first time t_0 such that $A(x_0,t_0)$ admits a non-trivial kernel. i.e., there exists $v \neq 0$ such that

$$\langle A(x_0,t_0)v, v\rangle = 0.$$

Consider

$$w := \langle A(x,t)v, v\rangle .$$

Then we have at (x_0,t_0),

$$\begin{aligned}(\partial_t - \Delta)w &= \langle ((\partial_t - \Delta)A)v, v\rangle \\ &> \langle Q(A,x,t)v, v\rangle \geq 0\end{aligned}$$

by assumption and the null vector condition, so at (x_0,t_0),

$$\partial_t w > \Delta w \geq 0,$$

contradicting to the assumption that t_0 is the first time such that A admits a non-trivial kernel.

For the general case that $(\partial_t - \Delta)A \geq Q(A,x,t)$, consider

$$B = A + \varepsilon(t+\kappa)\delta_{ij}$$

for some $\kappa > 0$. Then

$$(\partial_t - \Delta)B = (\partial_t - \Delta)A + \varepsilon\delta_{ij} \geq Q(A,x,t) + \varepsilon\delta_{ij}.$$

Suppose $B(x_0,t_0)v_0 = 0$ at some (x_0,t_0) with t_0 smallest. Then $Av = -\varepsilon(t+\kappa)v$. By the null vector condition, we can take the (positive) infimums of $\langle Q(A,x,t)v,v\rangle$ and $|Av|$ over $v \in (\ker A)^\perp$ with unit length. Thus for any unit vector v,

$$\langle Q(A,x,t)v, v\rangle \geq -C|Av|$$

for some C depending only on the eigenvalues of A. As a result, for any unit vector v,

$$\begin{aligned}\langle (\partial_t - \Delta)Bv, v\rangle &\geq \langle Q(A,x,t)v, v\rangle + \varepsilon \\ &\geq -C|Av| + \varepsilon \\ &= -C\varepsilon(t+\kappa) + \varepsilon = \varepsilon\left(-C(t+\kappa) + 1\right).\end{aligned}$$

This holds for any $\kappa > 0$, so after choosing κ small, we have

$$\langle (\partial_t - \Delta)Bv, v\rangle \geq \frac{\varepsilon}{2}.$$

However, since $\langle B(x_0,t_0)v_0, v_0\rangle = 0$ and $\langle B(x,t_0)v_0, v_0\rangle \geq 0$, we have

$$\Delta \langle Bv_0, v_0\rangle (x_0) \geq 0$$

so $\partial_t \langle Bv_0, v_0\rangle < 0$, contradicting to the assumption that t_0 is the first time. Hence we conclude that B is positive for a fixed time forward (independent of ε). $\square$

In the last step of the proof of theorem 10.4, we apply theorem 10.8 with $A = uM$, so (10.6) becomes

$$(\partial_t - \Delta)A = 2\frac{A^2}{u} - 2\frac{A}{t},$$

and hence

$$Q(A, x, t) = 2\frac{A^2}{u} - 2\frac{A}{t} = 2\left(\frac{A}{u} - \frac{1}{t}\delta_{ij}\right)A.$$

Therefor, Q satisfies the null vector condition since if $Av = 0$, then $Qv = 2\left(\frac{A}{u} - \frac{1}{t}\delta_{ij}\right)Av = 0$.

11. **Other Maximum Principles**

11.1. **Elliptic Strong Maximum Principle.** First we recall the elliptic case. Suppose $u\colon B_1 \to \mathbb{R}$ is a harmonic function. The strong maximum principle asserts that if u achieves an extremum in an interior point, then u must be constant.

To see this, consider

$$h_\alpha(x) := e^{-\alpha|x|^2} - e^{-\alpha}.$$

Then

$$\begin{aligned}\Delta h_\alpha &= -\alpha e^{-\alpha|x|^2}\cdot 2n + \alpha^2 e^{-\alpha|x|^2}\cdot 4|x|^2\\ &= 2\alpha\left(2\alpha|x|^2 - 2n\right)e^{-\alpha|x|^2}\end{aligned}$$

by applying with $f = e^{-\alpha s} - e^{-\alpha}$ and $g = |x|^2$ (so $h_\alpha = f\circ g$). In particular, if we take $\delta \in (0,1)$ and a sufficiently large α, we have

$$\Delta h_\alpha \ge 2\alpha\left(2\alpha\delta^2 - 2n\right)e^{-\alpha|x|^2} > 0. \tag{11.1}$$

Based on this, we can establish Hopf's lemma.

Theorem 11.2 (Hopf's lemma)**.** Let $u \in C^2(\overline{B}_1)$ be a harmonic function satisfying

$$u(x_0) = \max_{B_1} u > u(0)$$

for some $x_0 \in \partial B_1$. Then

$$\frac{\partial u}{\partial r}(x_0) > 0.$$

(Proof.) By continuity, we can take small $\delta > 0$ such that

$$u(x_0) = \max_{B_\delta} u.$$

Then we can take α so large that (11.1) holds, so

$$\Delta(u + \varepsilon h_\alpha) > 0$$

for any $\varepsilon > 0$. By assumption, we can take ε so small that

$$u + \varepsilon h_\alpha < u(x_0)$$

on ∂B_δ. At the same time, we have

$$u + \varepsilon h_\alpha = u \le u(x_0)$$

on ∂B_1 by the definition of h_α. The maximum principle then implies that

$$\max_{B_1 \setminus B_\delta} (u + \varepsilon h_\alpha) = u(x_0) = (u + \varepsilon h_\alpha)(x_0).$$

As a result,

$$\frac{\partial(u + \varepsilon h_\alpha)}{\partial r}(x_0) \ge 0.$$

Therefore

$$\frac{\partial u}{\partial r}(x_0) \ge -\varepsilon \frac{\partial h_\alpha}{\partial r}(x_0) = 2\alpha\varepsilon e^{-\alpha},$$

which is positive. □

Based on this, we have the strong maximum principle.

Theorem 11.3 (elliptic strong maximum principle)**.** Let $u \in C^2(\overline{\Omega})$ be a harmonic function on a connected domain Ω. If u achieves its maximum at an interior point, then u is constant.

(Proof.) Suppose u achieve its maximum at an interior point x_0, and take $B_\delta(x_0)$ contained in Ω. Hopf's lemma implies u is constant on $B_\delta(x_0)$. This implies that the subset $\{u = u(x_0)\}$ is open. Then the conclusion follows by the connectivity of Ω. □

11.2. **Parabolic Strong Maximum Principle.** As expected, we have a parabolic version as follows.

Theorem 11.4. Let Ω be a connected domain and $u \in C^2(\Omega \times [0,T])$ satisfies $(\partial_t - \Delta)u = 0$. If $x_0 \in \mathrm{int}\Omega$ satisfies

$$\max_{\Omega \times [0,T]} u = u(x_0, T),$$

then u is constant.

We remark that the discrete heat equation fails the strong maximum principle. Also, this is related to infinite propagation speed, which says that if u_0 non-negative, and positive at some point, then $u(\cdot, t) > 0$ for $t > 0$, which could be seen from the representation formula for the solution to the heat equation. (Note that the parabolic Harnack inequality provides a more robust statament.) To see this, let $u_0 := u(\cdot, 0)$, and we let

$$v(x,t) := (4\pi(t+\varepsilon))^{-\frac{n}{2}} e^{-\frac{|x|^2}{4(t+\varepsilon)}}$$

be the fundamental solution (with time shifted) such that v is below u (i.e., $v \le u$). Then the property follows from the parabolic strong maximum principle.

11.3. **More General Setting.** Let M be a Riemannian manifold with a time-varying metric $g(t)$. For a symmetric 2-tensor A, we can define

$$\Delta A := \mathrm{tr}(\mathrm{Hess}_A),$$

where $\mathrm{Hess}_A(X, Y) = \nabla_X \nabla_Y A - \nabla_{\nabla_X Y} A$ for tangent vectors X and Y. We could also define

$$\partial_t A := \nabla_{\partial_t} A.$$

Then we can look at the inequality

$$(\partial_t - \Delta)A \geq 0,$$

in the sense that it is semi-positive definite. The point is that what we have seen could be gneralized to this setting. That is to say, if $A \geq 0$ on the parabolic boundary, then it remains so all the time (weak version). The idea is to prove that the smallest eigenvalue λ remains non-negative.

As an example, let $\Sigma_t^n \subseteq \mathbb{R}^{n+1}$ and let A be its second fundamental form. Then it holds

$$(\partial_t - \Delta)A = |A|^2 A$$

if the hypersurfaces are flowed by the MCF. Thus the maximum principle implies that if the initial hypersurface is convex ($A \geq 0$), then it remains so for later time.

12. Growth of Solutions to Some PDEs

12.1. **Laplace Equation.** Fred Almgren was interested in the regularities of minimal surfaecs (starting in 1960s). To deal with this, he did some graphical approximations by harmonic functions, letting him understand better their grwoth.

The objects we are going to see are functions on $\mathbb{R}^n$. Let $u \colon \mathbb{R}^n \to \mathbb{R}$ be a harmonic function. Then we consider

$$I(r) := r^{1-n} \int_{\partial B_r} u^2,$$

the scale-invariant energy

$$D(r) := r^{2-n} \int_{B_r} |\nabla u|^2,$$

and Almgren's frequency

$$U(r) := \frac{D(r)}{I(r)}.$$

(We could replace B_r with $B_r(x)$ for any other $x \in \mathbb{R}^n$.) The important things are their derivatives. As we have seen, by a change of variable, we have

$$I'(r) = r^{1-n} \int_{\partial B_r} (u^2)_r = 2r^{1-n} \int_{\partial B_r} u u_r = r^{1-n} \int_{B_r} \Delta u^2 = 2r^{1-n} \int_{B_r} |\nabla u|^2 = \frac{2D(r)}{r} \tag{12.1}$$

by Stokes' theorem and the fact that u is harmonic. Next,

$$D'(r) = \left(r^{2-n} \int_0^r \int_{\partial B_s} |\nabla u|^2 \right)' = \frac{2-n}{r} D(r) + r^{2-n} \int_{\partial B_r} |\nabla u|^2. \tag{12.2}$$

We introduce the Pohozaev identity, which holds for any function, but particularly useful when the function is harmonic.

Theorem 12.3 (the Pohozaev identity)**.** For a C^2 function u,

$$(2-n)\int_{B_r} |\nabla u|^2 + \int_{B_r} \left\langle \nabla|x|^2, \nabla u \right\rangle \Delta u = 2r\int_{\partial B_r} u_r^2 - r\int_{\partial r} |\nabla u|^2.$$

(Proof.) We want to use the divergence theorem on the vector field

$$X := \left\langle \nabla|x|^2, \nabla u \right\rangle \nabla u - \frac{1}{2}|\nabla u|^2 \nabla |x|^2.$$

Its divergence is

$$\begin{aligned}\text{div}X &= \left\langle \nabla|x|^2, \nabla u \right\rangle \Delta u + \left\langle \nabla \left\langle \nabla|x|^2, \nabla u \right\rangle, \nabla u \right\rangle - \frac{1}{2}|\nabla u|^2 \Delta|x|^2 - \frac{1}{2}\left\langle \nabla|\nabla u|^2, \nabla|x|^2 \right\rangle \\ &= \left\langle \nabla|x|^2, \nabla u \right\rangle \Delta u - n|\nabla u|^2 + \left\langle \nabla \left\langle \nabla|x|^2, \nabla u \right\rangle, \nabla u \right\rangle - \frac{1}{2}\left\langle \nabla|\nabla u|^2, \nabla|x|^2 \right\rangle.\end{aligned}$$

Note

$$\left\langle \nabla|x|^2, \nabla u \right\rangle = \langle 2x_i\partial_i, \nabla u \rangle = 2x_i u_i,$$

so

$$\nabla \left\langle \nabla|x|^2, \nabla u \right\rangle = (2\delta_{ij}u_i + 2x_i u_{ij})\partial_j$$

which implies

$$\left\langle \nabla \left\langle \nabla|x|^2, \nabla u \right\rangle, \nabla u \right\rangle = 2u_j^2 + 2x_i u_{ij} u_j = 2|\nabla u|^2 + 2x_i u_{ij} u_j$$

The second thing we want to calculate is

$$\left\langle \nabla|\nabla u|^2, \nabla|x|^2 \right\rangle = 2x_i\partial_i(u_j^2) = 4x_i u_{ij} u_j.$$

As a result, the terms involving $x_i u_{ij} u_j$ cancel. That is,

$$\text{div}X = \left\langle \nabla|x|^2, \nabla u \right\rangle \Delta u + (2-n)|\nabla u|^2.$$

The divergence theorem then implies

$$\begin{aligned}\int_{B_r} \left\langle \nabla|x|^2, \nabla u \right\rangle \Delta u + (2-n)\int_{B_r} |\nabla u|^2 &= \int_{B_r} \text{div}X \\ &= \int_{\partial B_r} \left\langle X, \frac{\partial}{\partial r} \right\rangle \\ &= \int_{\partial B_r} \left(2ru_r^2 - \frac{1}{2}|\nabla u|^2(2r) \right),\end{aligned}$$

which proves the identity. □

Corollary 12.4 (Almgren's version of Hadamard's three circles theorem)**.** If u is a harmonic function on $\mathbb{R}^n$, then $U'(r) \geq 0$.

(Proof.) We know

$$U' = \left(\frac{D}{I}\right)' = \frac{D'I - I'D}{I^2}.$$

Thus it suffices to show

$$D'I - I'D \geq 0.$$

Observe that (12.1) implies $D = \frac{rI'}{2}$, and that

$$I'D = \frac{rI'}{2} \cdot I' = 2r^{3-2n} \left(\int_{\partial B_r} u u_r \right)^2 .$$

On the other hand, (12.2) and the Pohozaev identity give (noting u is harmonic)

$$D' = \frac{2-n}{r} D(r) + r^{1-n} \left(2r \int_{\partial B_r} u_r^2 - (2-n) \int_{B_r} |\nabla u|^2 \right) = 2r^{2-n} \int_{\partial B_r} u_r^2$$

by the definition of D. (In particulet, we know that $D' \geq 0$.) As a result,

$$D'I = 2r^{2-n} \int_{\partial B_r} u_r^2 \cdot r^{1-n} \int_{\partial B_r} u^2 = 2r^{3-2n} \left(\int_{\partial B_r} u_r^2 \right) \left(\int_{\partial B_r} u^2 \right).$$

Thus

$$D'I - I'D = 2r^{3-2n} \left(\left(\int_{\partial B_r} u_r^2 \right) \left(\int_{\partial B_r} u^2 \right) - \left(\int_{\partial B_r} u u_r \right)^2 \right),$$

so it remains to show

$$\left(\int_{\partial B_r} u_r^2 \right) \left(\int_{\partial B_r} u^2 \right) \geq \left(\int_{\partial B_r} u u_r \right)^2,$$

which is true by the Cauchy-Schwarz inequality. □

From the proof, the Cauchy-Schwarz inequality tells us that $U'(r) = 0$ if and only if

(12.5) $$u(r) = c(r) u$$

for some $c(r)$. In fact, we could get more.

Corollary 12.6. If $U' = 0$, then $u = r^U \cdot g(\theta)$ where g only depends on the spherical direction.

(Proof.) Note that

$$I' = \frac{2D}{r} = 2r^{1-n} \int_{\partial B_r} u u_r = 2cr^{1-n} \int_{\partial B_r} u^2 = 2cI$$

by (12.5). Thus

$$U = \frac{D}{I} = \frac{rI'}{2I} = \frac{2rcI}{2I} = rc.$$

As a result

$$c = \frac{U}{r}.$$

Thus

$$u_r = cu = \frac{U}{r} u$$

and we get

$$\frac{\partial}{\partial r} \left(r^{-U} u \right) = 0.$$

As a result, $r^{-U} u = g(\theta)$ for some g not depending on r. □

Note that a function u of the form $r^U \cdot g(\theta)$ is harmonic if and only if g is an eigenfunction of the laplacian on the sphere.

Another observation is

$$(\log I)' = \frac{I'}{I} = \frac{2D}{rI} = \frac{2}{r}U. \tag{12.7}$$

Thus if $I(r) \approx r^{2d}$, then $(\log I)' \approx \frac{2d}{r}$ so heuristically the exponent d indicates the frequency.

12.2. **Unique Continuation Property.** Another consequence of the monotonicity of the frequency is the unique continuation property. Suppose u is a harmonic funtion on a connectd domain $\Omega \subseteq \mathbb{R}^n$. The unique continuation of a PDE problem asserts that if a solution vanishes on a non-empty open subset, then it vanishes anywhere. The strong unique continuation says that if u and all its derivatives vanish at a point, then it vanishes everywhere.

Theorem 12.8. Suppose $\Omega \subseteq \mathbb{R}^n$ is an open connected subset and u is harmonic on Ω with $\nabla^k u(p) = 0$ for all k at some $p \in \Omega$, then $u = 0$ on Ω. That is to say, the strong unique continuation property holds for the laplacian equation.

(Proof.) We may assume $p = 0$. It suffices to show that $\nabla^k u = 0$ for all k in a neighborhood of 0 (since Ω is connected).

By (12.7), we have

$$2\int_r^s \frac{U(\tau)}{\tau} d\tau = \int_r^s (\log I)'(\tau) d\tau = \log\left(\frac{I(s)}{I(r)}\right)$$

if $r \le s$ and $B_s \subseteq \Omega$. Since U is increasing by the corollary 12.4, we know $U(\tau) \le U(s) =: c$ so

$$\log\left(\frac{I(s)}{I(r)}\right) \le 2c \int_r^s \frac{1}{\tau} d\tau = \left(\log \tau^{2c}\right)|_r^s$$

and hence

$$s^{-2c} I(s) \le r^{-2c} I(r),$$

which tends to 0 as $r \to 0$ (by seeing the Taylor expansion). That is to say, $I(s) = 0$ (for all s). In particular, the conclusion follows. □

Recall the Ornstein-Uhlenbeck operator

$$Lu = e^{\frac{|x|^2}{4}} \operatorname{div}\left(e^{-\frac{|x|^2}{4}} \nabla u\right),$$

which is one of many possible drift operators. It could be rewrite as

$$Lu = \Delta u - \frac{1}{2}\langle x, \nabla u\rangle.$$

If $u\colon \mathbb{R}^n \to \mathbb{R}$ is an eigenfunction of Δ. i.e.,

$$Lu + \lambda u = 0$$

for some $\lambda \geq 0$, we can still define the function $I(r)$, and get

$$I'(r) = r^{1-n} \int_{\partial B_r} (u^2)_r = r^{1-n} e^{\frac{r^2}{4}} \int_{B_r} \operatorname{div} \left(e^{-\frac{|x|^2}{4}} \nabla u^2 \right)$$

by Stokes' theorem. Thus we get

$$I'(r) = r^{1-n} e^{\frac{r^2}{4}} \int_{B_r} L u^2 e^{-\frac{|x|^2}{4}}.$$

Since

$$Lu^2 = -2\lambda u^2 + 2|\nabla u|^2,$$

we get

$$I'(r) = r^{1-n} e^{\frac{r^2}{4}} \int_{B_r} \left(-2\lambda u^2 + 2|\nabla u|^2 \right) e^{-\frac{|x|^2}{4}}.$$

Thus if we define

$$D := r^{2-n} e^{\frac{r^2}{4}} \int_{B_r} \left(|\nabla u|^2 - \lambda u^2 \right) e^{-\frac{|x|^2}{4}}, \tag{12.9}$$

we still get

$$I' = \frac{2D}{r}, \tag{12.10}$$

and hence

$$(\log I)' = \frac{2U}{r}$$

where $U := \frac{D}{I}$ again.

Based on (12.9), we have

$$D' = \frac{2-n}{r} D + \frac{r}{2} D + r^{2-n} \int_{\partial B_r} \left(|\nabla u|^2 - \lambda u^2 \right). \tag{12.11}$$

It turns out that in this case, U is no longer monotone. However, we could still derive some useful information as follows.

$$\frac{U'}{U} \geq \frac{2-n}{r} + \frac{r}{2} - \frac{\lambda r}{U} - \frac{U}{r}. \tag{12.12}$$

To see (12.12), note that by (12.10),

$$D^2 = \left(\frac{rI'}{2} \right)^2 = \left(r^{2-n} \int_{\partial B_r} u u_r \right)^2 \leq \left(r^{1-n} \int_{\partial B_r} u^2 \right) \left(r^{3-n} \int_{\partial B_r} u_r^2 \right) = I \cdot \left(r^{3-n} \int_{\partial B_r} u_r^2 \right)$$

by the Cauchy-Schwarz inequality. Thus we get

$$D^2 \leq I \cdot \left(r^{3-n} \int_{\partial B_r} |\nabla u|^2 \right).$$

Therefore,

$$\frac{D}{I} \leq \frac{1}{D} \left(r^{3-n} \int_{\partial B_r} |\nabla u|^2 \right),$$

so

$$\frac{U}{r} \leq \frac{1}{D}\left(r^{2-n}\int_{\partial B_r}|\nabla u|^2\right). \tag{12.13}$$

As a consequence, since by (12.11) we have

$$D' = \frac{2-n}{r}D + \frac{r}{2}D + r^{2-n}\int_{\partial B_r}|\nabla u|^2 - \lambda I r,$$

we derive

$$\frac{D'}{D} = \frac{2-n}{r} + \frac{r}{2} + \frac{1}{D}r^{2-n}\int_{\partial B_r}|\nabla u|^2 - \frac{\lambda r}{U},$$

which along with (12.10) gives

$$\begin{aligned}
\frac{U'}{U} &= \frac{D'}{D} - \frac{I'}{I} \\
&= \left(\frac{2-n}{r} + \frac{r}{2} + \frac{1}{D}r^{2-n}\int_{\partial B_r}|\nabla u|^2 - \frac{\lambda r}{U}\right) - \frac{2D}{rI} \\
&= \frac{2-n}{r} + \frac{r}{2} - \frac{\lambda r}{U} - \frac{2U}{r} + \frac{1}{D}r^{2-n}\int_{\partial B_r}|\nabla u|^2 \\
&\geq \frac{2-n}{r} + \frac{r}{2} - \frac{\lambda r}{U} - \frac{2U}{r} + \frac{U}{r}
\end{aligned}$$

by (12.13), which proves (12.12).

Using (12.12), we get a dichotomy.

Theorem 12.14. Given $\delta > 0$ and $\lambda > 0$, there exists $R_0 > 0$ suc htaht the following holds. If $Lu + \lambda u = 0$ where L is the Ornstein-Uhlenbeck operator, then $U(r) > (2+\delta)\lambda$ for $r > R_0$, i.e.,

$$\lim_{r\to\infty}\frac{U(r)}{r^2} \geq \frac{1}{2}.$$

Roughly it says that for an eigenfunction u, either $U \leq 2\lambda$ or U grows like $r^2/2$.

(Proof.) If $U(r) \in [(2+\delta)\lambda, 2r]$, then by (12.12),

$$\begin{aligned}
(\log U)' &\geq \frac{2-n}{r} + \frac{r}{2} - \frac{\lambda r}{U} - \frac{U}{r} \\
&\geq \frac{2-n}{r} + \frac{r}{2} - 2 - \frac{r\lambda}{(2+\delta)\lambda} \\
&= \frac{2-n}{r} + r\left(\frac{1}{2} - \frac{1}{2+\delta}\right) - 2.
\end{aligned}$$

On the other hand, if $U(r) \geq 2r$,

$$\begin{aligned}
(\log U)' &\geq \frac{2-n}{r} + \frac{r}{2} - \frac{\lambda r}{U} - \frac{U}{r} \\
&\geq \frac{2-n}{r} + \frac{r}{2} - \frac{U}{r} - \frac{\lambda}{2}.
\end{aligned}$$

Thus either $U \geq (1-\varepsilon)r^2/2$, or $U < (1-\varepsilon)r^2/2$, in which case we have

$$(\log U)' \geq \frac{2-n}{r} + \frac{r}{2} - \frac{r}{2} - \frac{\lambda}{2} = \frac{1-\varepsilon}{2} r,$$

which gives exponential growth, leading to a contradiction. Thus the conclusion follows. □

12.3. **Heat Equation.** Recall that our favorite solution to the heat equation is

$$u(x,t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}.$$

We also denote $H_b(x,t) := u(x,-t)$, which is a solution to the backward solution since

$$(\partial_t + \Delta)H_b = -u_t + \Delta u = 0.$$

When discussing the heat equation, we could not focus on local spaces, since solutions to the heat equation have infinite propogation speed. i.e., even if the initial data has compact support, it will not after any positive time. However, we could localize the time.

Now we let u be any solution to the heat equation on the time interval $(-\infty, 0]$. Consider a natural analog as what we did for harmonic functions, that is,

$$I(t) := \int u^2 H_b dx = \int u^2(x,t) H_b(x,t) dx$$

with

$$H_b(x,t) = (-4\pi t)^{-\frac{n}{2}} e^{\frac{|x|^2}{4t}}.$$

defined above, which has integral 1 at any time. Thus

$$\begin{aligned} I'(t) &= \int 2uu_t H_b + \int u^2 (H_b)_t \\ &= \int 2u\Delta u H_b - \int u^2 \Delta H_b. \end{aligned}$$

Since $\Delta u^2 = 2|\nabla u|^2 + 2u\Delta u$ (for any C^2 functions), we have

$$I' = \int \Delta(u^2) H_b - 2\int |\nabla u|^2 H_b - \int \Delta(u^2) H_b = -2\int |\nabla u|^2 H_b. \tag{12.15}$$

after integration by parts since H_b decays incredibly fast at infinity at any time and we assume u does too. We would like to have a similar formula as (12.1). Thus we define

$$D(t) := -t \int |\nabla u|^2 H_b, \tag{12.16}$$

which implies $I' = \frac{2D}{t}$. Lastly, we again define the frequency

$$U := \frac{D}{I}.$$

Thus of course

$$(\log I)' = \frac{I'}{I} = \frac{2D}{tI} = \frac{2}{t} U.$$

We want to revise D. In fact, using $\Delta u^2 = 2|\nabla u|^2 + 2u\Delta u$, we get

$$D = -\frac{t}{2}\int \Delta(u^2)H_b + t\int u\Delta u H_b \tag{12.17}$$
$$= t\int u\langle \nabla u, \nabla H_b\rangle + t\int u\Delta u H_b$$
$$= t\int u\left(\langle \nabla u, \nabla \log H_b\rangle + \Delta u\right)H_b.$$

after integration by parts.

Next, we will see that D is also monotone.

Lemma 12.18. We have

$$D' = 2t\int \left(\Delta u + \langle \nabla u, \nabla \log H_b\rangle\right)^2 H_b.$$

In particular, $D' \le 0$.

Corollary 12.19. The frequency is decreasing. That is, $U' \le 0$.

(Proof of corollary 12.19.) We have

$$U' = \frac{D'I - I'D}{I^2}.$$

Thus it suffices to show $D'I - I'D \le 0$. By lemma (12.18), (12.15) and (12.17),

$$D'I - I'D = 2t\int \left(\Delta u + \langle \nabla u, \nabla \log H_b\rangle\right)^2 H_b \cdot \int u^2 H_b - \frac{2}{t}D^2$$
$$= 2t\int \left(\Delta u + \langle \nabla u, \nabla \log H_b\rangle\right)^2 H_b \cdot \int u^2 H_b - 2t\left(\int u\left(\langle \nabla u, \nabla \log H_b\rangle + \Delta u\right)H_b\right)^2.$$

Viewing $A = u$ and $B = \langle \nabla u, \nabla \log H_b\rangle + \Delta u$, the corollary follows from the Cauchy-Schwarz inequality (using the weighted integral as the inner product). □

(Proof of lemma 12.18.) By the original definition (12.16), we have

$$D' = -\int |\nabla u|^2 H_b - 2t\int \langle \nabla u_t, \nabla u\rangle H_b - t\int |\nabla u|^2 (H_b)_t$$
$$= -\int |\nabla u|^2 H_b - 2t\int \langle \nabla \Delta u, \nabla u\rangle H_b + t\int |\nabla u|^2 \Delta H_b$$
$$= -\int |\nabla u|^2 H_b + 2t\int \left((\Delta u)^2 H_b + \langle \nabla u, \nabla \log H_b\rangle H_b \Delta u\right) - t\int \left\langle \nabla |\nabla u|^2, \nabla \log H_b\right\rangle H_b$$

using integration by parts. Note that we can write the last term in local coordinate. i.e.,

$$\begin{aligned}
&\int \left\langle \nabla |\nabla u|^2, \nabla \log H_b \right\rangle H_b \\
&= 2 \int u_{ij} u_i (\log H_b)_j H_b \\
&= -2 \int u_j u_{ii} (\log H_b)_j H_b - 2 \int u_j u_i (\log H_b)_{ij} H_b - 2 \int u_j u_i (\log H_b)_j (\log H_b)_i H_b \\
&= -2 \int \Delta u \left\langle \nabla u, \nabla \log H_b \right\rangle H_b - 2 \int \mathrm{Hess}_{\log H_b} (\nabla u, \nabla u) H_b - 2 \int \left\langle \nabla u, \nabla \log H_b \right\rangle^2 H_b.
\end{aligned}$$

Since

$$H_b(x,t) = (-4\pi t)^{-\frac{n}{2}} e^{\frac{|x|^2}{4t}},$$

we have

$$\nabla \log H_b = \nabla \left(\frac{|x|^2}{4t} \right) = \frac{x}{2t}$$

and

$$\mathrm{Hess}_{\log H_b} = \frac{1}{2t}.$$

As a result,

$$\begin{aligned}
&\int \left\langle \nabla |\nabla u|^2, \nabla \log H_b \right\rangle H_b \\
&= -2 \int \Delta u \left\langle \nabla u, \nabla \log H_b \right\rangle H_b - \frac{1}{t} \int |\nabla u|^2 H_b - 2 \int \left\langle \nabla u, \nabla \log H_b \right\rangle^2 H_b.
\end{aligned}$$

Putting this back to D', we get

$$\begin{aligned}
D' &= -\int |\nabla u|^2 H_b + 2t \int \big((\Delta u)^2 H_b + \left\langle \nabla u, \nabla \log H_b \right\rangle H_b \Delta u \big) \\
&\quad + 2t \int \Delta u \left\langle \nabla u, \nabla \log H_b \right\rangle H_b + \int |\nabla u|^2 H_b + 2t \int \left\langle \nabla u, \nabla \log H_b \right\rangle^2 H_b \\
&= 2t \int H_b \left((\Delta u)^2 + 2\Delta u \left\langle \nabla u, \nabla \log H_b \right\rangle + \left\langle \nabla u, \nabla \log H_b \right\rangle^2 \right),
\end{aligned}$$

which is what we want. □

Note that by the argument in the proof of corollary 12.19, $U' = 0$ if and only if

$$\left\langle \nabla u, \nabla \log H_b \right\rangle + \Delta u = c(t) u.$$

In this case,

$$D = t \int u \left(\left\langle \nabla u, \nabla \log H_b \right\rangle + \Delta u \right) H_b = c(t) t \int u^2 H_b = c(t) t I,$$

so

$$U = \frac{D}{I} = c(t) t,$$

which implies $c(t) = \frac{U}{t}$, and hence

$$\langle \nabla u, \nabla \log H_b \rangle + \Delta u = \frac{U}{t} u,$$

which is equivalent to

$$Uu = tu_t + \frac{1}{2} \langle \nabla u, x \rangle . \tag{12.20}$$

If we consider

$$v(t) := t^{-U} u(\sqrt{t}x, -t),$$

then by (12.20),

$$\begin{aligned} v'(t) &= -Ut^{-U-1}u + t^{-U} \left\langle \nabla u, \frac{x}{2\sqrt{t}} \right\rangle - t^{-U} u_t \\ &= t^{-U} \left(\frac{-Uu}{t} + \left\langle \nabla u, \frac{x}{2\sqrt{t}} \right\rangle - u_t \right) \\ &= t^{-U} \left(\frac{-Uu}{t} + \frac{1}{2t} \left\langle \nabla u, \sqrt{t}x \right\rangle - u_t \right) = 0 \end{aligned}$$

since the spacial variable is $\sqrt{t}x$ and the time variable is $-t$ now (so (12.20) becomes

$$\frac{-Uu}{t} = u_t - \frac{1}{2t} \left\langle \nabla u, \sqrt{t}x \right\rangle$$

in this case). As a result,

$$t^{-U} u(\sqrt{t}x, -t) = v(t) = v(1) = u(x, -1),$$

so

$$u(\sqrt{t}x, -t) = t^U u(x, -1),$$

which is equivalent to

$$u(x,t) = (-t)^U u \left(\frac{x}{\sqrt{-t}}, -1 \right),$$

in which the variables are separated, as what we have seen for the case of harmonic functions in corollary 12.6.

12.4. **Backward Uniqueness.** We claim that solutions to the heat equation also enjoy some uniqueness continuation properties.

Theorem 12.21. If $u \colon \mathbb{R}^n \times [-a, 0] \to \mathbb{R}$ is a solution to the heat equation and

$$\lim_{t \to 0} \frac{I(t)}{(-t)^d} = 0$$

for all $d > 0$, then $u = 0$.

The condition means that u vanishes to infinite order at the origin. We remark that this is not true for discrete heat equation.

(Proof.) Since $(\log I)' = \frac{2U}{t}$, for $t_2 < t_1 < 0$, we have

$$\log\left(\frac{I(t_1)}{I(t_2)}\right) = 2\int_{t_2}^{t_1} \frac{U}{t} dt.$$

By corollary 12.19, $U(t_1) \leq U(t_2)$. Thus

$$\log\left(\frac{I(t_1)}{I(t_2)}\right) \geq 2\int_{t_2}^{t_1} \frac{U(t_2)}{t} dt = 2U(t_2)\log\left(\frac{t_1}{t_2}\right).$$

Thus

$$\frac{I(t_1)}{I(t_2)} \geq \left(\frac{t_1}{t_2}\right)^{2U(t_2)}.$$

That is to say,

$$I(t_2)(-t_2)^{-2U(t_2)} \leq I(t_1)(-t_1)^{-2U(t_2)}.$$

Let $t_1 \to 0$, we get $I(t_2) = 0$. □

12.5. **More General Setting.** The content we are going to see is based on [CM20]. For a fixed smooth function φ, consider

$$L_\varphi u := e^{-\varphi}\text{div}\left(e^{-\varphi}\nabla u\right).$$

If u satisfies $(\partial_t - L_\varphi)u = 0$, we can define again

$$I(t) = \int u^2 e^{-\varphi},$$

$$D(t) = -\int |\nabla u|^2 e^{-\varphi} = \int u L_\varphi u e^{-\varphi},$$

and

$$U(t) := \frac{D(t)}{I(t)}.$$

Then

$$I'(t) = 2\int u u_t e^{-\varphi} = 2\int u L_\varphi u e^{-\varphi} = -2\int |\nabla u|^2 e^{-\varphi} = 2D,$$

and

$$D'(t) = -2\int \langle \nabla u_t, \nabla u\rangle e^{-\varphi} = 2\int u_t L_\varphi u e^{-\varphi} = 2\int (L_\varphi u)^2 e^{-\varphi}.$$

We assume we can do integration by parts again (with some assumptions on boundary data). Thus

$$(\log U)' = \frac{D'}{D} - \frac{I'}{I} = \frac{D'I - I'D}{DI},$$

where

$$D'I = 2\int (L_\varphi u)^2 e^{-\varphi} \cdot \int u^2 e^{-\varphi}$$

and

$$I'D = 2\left(\int uL_\varphi ue^{-\varphi}\right)^2,$$

so $D'I - I'D \geq 0$ by the Cauchy-Schwarz inequality, i.e., $(\log U)' \geq 0$. Also, we have

$$(\log I)' = \frac{I'}{I} = \frac{2D}{I} = 2U.$$

Note that the Cauchy-Schwarz inequality tells us that U is constant if and only if

$$L_\varphi u = c(t)u.$$

If it is the case, since

$$D = \int uL_\varphi ue^{-\varphi} = c(t)\int u^2 e^{-\varphi} = c(t)I,$$

we get $U = c(t)$. Thus $U' = 0$ if and only if $u_t = L_\varphi u = Uu$, which means

$$(e^{-Ut}u)' = 0,$$

that is,

$$u(x,t) = e^{Ut}u(x,0).$$

In fact, this more general monotonicity implies Poon's monotonicity. To see this, if $u_t = \Delta u$ for $t < 0$, consider

$$w(x,s) = u(\sqrt{-t}x, t) = u\left(e^{-\frac{s}{2}}x, -e^{-s}\right)$$

where $t = -e^{-s}$. Then as we have seen in previous sections,

$$w_s = -\frac{1}{2}e^{-\frac{s}{2}}\langle \nabla u, x\rangle + e^{-s}u_t,$$

$$\nabla w = e^{-\frac{s}{2}}\nabla u,$$

and

$$\Delta w = e^{-s}\Delta u.$$

Thus

$$w_s - L_{\frac{|x|^2}{4}} = w_s - \Delta w + \frac{1}{2}\langle x, \nabla w\rangle = 0$$

since $u_t = \Delta u$. Note that w is defined for all time (in the variable s). Thus for w, we have $I_w(s)$, $D_w(s)$ and $U_w(s)$ defined above, and for u, we have

$$I_u(t) = (-4\pi t)^{-\frac{n}{2}}\int u^2 e^{-\frac{|x|^2}{4t}}.$$

Note that by a change of variables $t = e^{-s}$, $I_w(e^{-s}) = (-t)^{-\frac{n}{2}}I_u(t)$. Thus the monotonicity for w implies that for u.

12.6. **Backward Uniqueness in the General Setting.**

Theorem 12.22. Suppose $u_t = L_\varphi u$ on the time interval $[a,b]$. If $u = 0$ at time b, then $u \equiv 0$.

(Proof.) Since $(\log I)' = 2U$, for $a \leq t \leq b$,

$$\log I(b) - \log I(t) = \int_t^b 2U \geq 2U(t)(b-t)$$

since U is increasing. Taking exponential gives

$$\frac{I(b)}{I(t)} \geq e^{2U(t)(b-t)},$$

so if $I(b) = 0$, then $I(t) = 0$ for all t. □

It is remarkable that all the arguments become simpler in this general setting.

13. **Ancient Solutions to the Heat Equation**

The content of this section is based on [CM21]. A function u is ancient if it is defined for all prior time. That is, it is definded on time $(-\infty, c]$ for some $c \in \mathbb{R}$. This is, in a sense, a natural generalization of harmonic functions.

13.1. **Reverse Poincaré Inequality.** Let (M,g) be a Riemannian manifold, and u be an ancient solution to the heat equation on M. We want to obtain a parabolic version of the reverse Poincaré inequality for u. Let $\varphi\colon M \to \mathbb{R}$ be a non-negative function with compact support. We would like to look at

$$\begin{aligned} \frac{d}{dt}\int u^2\varphi^2 dV_g &= 2\int u u_t \varphi^2 dV_g \\ &= 2\int u\Delta u \varphi^2 dV_g \\ &= -2\int \left(|\nabla u|^2\varphi^2 + 2u\varphi\left\langle \nabla u, \nabla\varphi\right\rangle\right) dV_g, \end{aligned}$$

which, by the AM-GM inequality (absorbing inequality), is less than

$$-2\int |\nabla u|^2\varphi^2 dV_g + 2\int\left(\frac{1}{2}|\nabla u|^2\varphi^2 + 2u^2|\nabla\varphi|^2\right) dV_g.$$

Thus

$$\frac{d}{dt}\int u^2\varphi^2 dV_g \leq -\int |\nabla u|^2\varphi^2 dV_g + 4\int u^2|\nabla\varphi|^2 dV_g.$$

Hence after integration, we derive

$$\int u^2\varphi^2 dV_g|_{t=0} - \int u^2\varphi^2 dV_g|_{t=-T} \leq -\int_{-T}^0\int |\nabla u|^2\varphi^2 dV_g + 4\int_{-T}^0\int u^2|\nabla\varphi|^2 dV_g.$$

Therefore,

$$\int_{-T}^{0}\int |\nabla u|^2\varphi^2 dV_g \le 4\int_{-T}^{0}\int u^2|\nabla\varphi|^2 dV_g + \int u^2\varphi^2 dV_g|_{t=-T}.$$

We want to take the cut-off φ only depending on the distance to the origin (in the Euclidean case), which is 1 on B_R and 0 outside B_{2R}. Then we get

$$\int_{-T}^{0}\int_{B_R} |\nabla u|^2 \le 4\int_{-T}^{0}\int_{B_{2R}} \frac{u^2}{R^2} + \int_{B_{2R}} u(\cdot,-T)^2.$$

Note that we have

$$\int_{B_{2R}} u(\cdot,-T)^2 \le \frac{1}{R^2}\int_{-2R^2}^{0}\int_{B_{2R}} u^2$$

for some $-T\in[-2R^2,-R^2]$ (just taking the mean). Thus

$$\int_{-R^2}^{0}\int_{B_R} |\nabla u|^2 \le \frac{4}{R^2}\int_{-2R^2}^{0}\int_{B_{2R}} u^2 + \frac{1}{R^2}\int_{-2R^2}^{0}\int_{B_{2R}} u^2 = \frac{5}{R^2}\int_{-2R^2}^{0}\int_{B_{2R}} u^2. \tag{13.1}$$

Next, we look at

$$\begin{aligned}\frac{d}{dt}\int |\nabla u|^2\varphi^2 &= 2\int \langle\nabla u_t,\nabla u\rangle\varphi^2\\ &= -2\int u_t\Delta u\varphi^2 - 4\int u_t\varphi\langle\nabla u,\nabla\varphi\rangle\\ &\le -2\int u_t^2\varphi^2 + \left(\int u_t^2\varphi^2 + 4\int |\nabla\varphi|^2|\nabla u|^2\right)\\ &= -\int u_t^2\varphi^2 + 4\int |\nabla\varphi|^2|\nabla u|^2\end{aligned}$$

using the AM-GM inequality again. After integration, we get

$$\int |\nabla u|^2\varphi^2|_{t=0} - \int |\nabla u|^2\varphi^2|_{t=-T} \le -\int_{-T}^{0}\int u_t^2\varphi^2 + 4\int_{-T}^{0}\int |\nabla\varphi|^2|\nabla u|^2,$$

so

$$\int_{-T}^{0}\int u_t^2\varphi^2 \le \int |\nabla u|^2\varphi^2|_{t=-T} + 4\int_{-T}^{0}\int |\nabla\varphi|^2|\nabla u|^2.$$

Again, for some $T\in[R^2,2R^2]$, we have

$$\int |\nabla u|^2\varphi^2|_{t=-T} \le \frac{1}{R^2}\int_{-2R^2}^{0} |\nabla u|^2\varphi^2,$$

so

$$\int_{-T}^{0}\int u_t^2\varphi^2 \le 4\int_{-T}^{0}\int |\nabla\varphi|^2|\nabla u|^2 + \frac{1}{R^2}\int_{-2R^2}^{0} |\nabla u|^2\varphi^2.$$

Taking the same cut-off φ, it becomes

$$\int_{-R^2}^{0}\int_{B_R} u_t^2 \le \frac{4}{R^2}\int_{-2R^2}\int_{B_{2R}} |\nabla u|^2 + \frac{1}{R^2}\int_{-2R^2}\int_{B_{2R}} |\nabla u|^2 = \frac{5}{R^2}\int_{-2R^2}\int_{B_{2R}} |\nabla u|^2. \tag{13.2}$$

We are going to combine (13.1) and (13.2). The reason why they are useful is that their right hand sides are of the order R^{-2}. Thus in some integral sense, we could say

$$|u_t| \lesssim \frac{C}{R^2}|\nabla u| \text{ and } |\nabla u| \lesssim \frac{C}{R^2}|u|.$$

In fact, putting them together after doubling gives

$$R^2 \int_{-R^2}^{0} \int_{B_R} u_t^2 + \int_{-R^2}^{0} \int_{B_R} |\nabla u|^2 \leq \frac{C}{R^2} \int_{-4R^2}^{0} \int_{B_{4R}} |u|^2, \tag{13.3}$$

called the (parabolic-version) reverse Poincaré inequality.

13.2. **The Space of Ancient Solutions.** Now we assume our manifold (M, g) has polynomial volume growth, in the sense that fo any $p \in M$ and $r > 0$,

$$\mathrm{Vol}(B_r(p)) \leq C_M(1+r)^{k_M},$$

which is clearly a vector space. Other objects of interest form the following set. Let

$$\mathcal{H}_d(M) := \left\{u \colon M \to \mathbb{R} : \Delta u = 0 \text{ and } |u(x)| \leq C_u(1 + d(x,p))^d\right\},$$

whose elements are called harmonic functions of polynomial growth at most d. A classical result is that

$$\mathcal{H}_d(\mathbb{R}^n) = \{\text{harmonic ploynomials of degree at most } d\}.$$

In general, one can very efficiently bound the dimension of this kind of spaces, which is a conjecture of Yau. (See [CM97].)

Here we want to look at the parabolic version, i.e.,

$$\mathcal{P}_d(M) := \left\{u \colon M \times (-\infty, 0] \to \mathbb{R} : (\partial_t - \Delta)u = 0 \text{ and } |u(x,t)| \leq C_u\left(1 + d(x,p) + \sqrt{-t}\right)^d\right\}.$$

Here $d(x,p) + \sqrt{-t}$ is some kinds of "parabolic distance" to $(p, 0)$ in spacetime.

Theorem 13.4. Let (M, g) be a Riemannian manifold with polynomial volume growth. Then

$$\dim \mathcal{P}_{2d} \leq \sum_{k=0}^{d} \dim \mathcal{H}_{2k}(M).$$

We remark that this inequality is sharp, in the sence that the equality holds when (M, g) is the standard Euclidean space.

First we use the reverse Poincaré inequality to show the following lemma.

Lemma 13.5. If (M, g) has polynomial volume growth anad $u \in \mathcal{P}_d(M)$, then there exists $k > 0$ such that $\partial_t^k u = 0$.

(Proof.) By the reverse Poincaré inequality (13.3), we in particular have

$$r^4 \int_{-r^2}^{0} \int_{B_r} u_t^2 \leq C \int_{-4r^2}^{0} \int_{B_{2r}} |u|^2.$$

Since u_t is also an ancient solution, applying the inequality again gives

$$r^4 \int_{-r^2}^0 \int_{B_r} (\partial_t^2 u)^2 \le C \int_{-4r^2}^0 \int_{B_{2r}} |u_t|^2.$$

As a result,

$$\begin{aligned} r^8 \int_{-r^2}^0 \int_{B_r} (\partial_t^2 u)^2 &\le C \int_{-4r^2}^0 \int_{B_{2r}} |u_t|^2 \\ &\le C^2 \int_{-16r^2}^0 \int_{B_{4r}} u^2. \end{aligned}$$

Inductively, we have

$$r^{4k} \int_{B_r \times [-r^2,0]} (\partial_t^k u)^2 \le C^k \int_{B_{2^k r} \times [-2^{2k} r^2, 0]} u^2.$$

Since $u \in \mathcal{P}_d$, we have

$$|u(x,t)| \le C_u \left(1 + d(x,p) + \sqrt{-t}\right)^d.$$

Thus on $B_{2^k r} \times [-4^k r^2, 0]$,

$$|u(x,t)| \le C_u (1 + 2^k r + 2^k r)^d \le \widetilde{C}(1 + 2^k r)^d.$$

Therefore,

$$r^{4k} \int_{B_r \times [-r^2,0]} (\partial_t^k u)^2 \le \widetilde{C} C^k \mathrm{Vol}(B_{2^k r}) \cdot 2^{2k} r^2 (1 + 2^k r)^d.$$

By the polynomial volume growth, it is bounded by

$$\widetilde{C} C^k C_V (1+r)^\nu \cdot 2^{2k} r^2 (1 + 2^k r)^d.$$

In conclusion, we could write

$$r^{4k} \int_{B_r \times [-r^2,0]} (\partial_t^k u)^2 \le C_0 (4C)^k (1+r)^{\nu+2} (1 + 2^k r)^d.$$

Note that we are only interested in the case when r is large, so $1 + Cr \approx Cr$, which implies

$$r^{4k} \int_{B_r \times [-r^2,0]} (\partial_t^k u)^2 \le C_0 (4C)^k r^{\nu+d+2} \cdot 2^{kd}$$

and hence

$$\begin{aligned} \int_{B_r \times [-r^2,0]} (\partial_t^k u)^2 &\le C_0 r^{-4k} (2^{d+2} C)^k r^{\nu+d+2} \\ &= C_0 \left(\frac{2^{d+2} C}{r^4} \right)^k r^{\nu+d+2}. \end{aligned}$$

As a result, for k large (in fact, so large that $4k > \nu + d + 2$), we have

$$\int_{M \times (-\infty, 0]} (\partial_t^k u)^2 = 0,$$

and the conclusion of the lemma follows. □

(Proof of theorem 13.4.) Based on the lemma, for each fixed x, we know that the function $u(x,t)$ is a polynomial in t with degree at most $k-1$. Thus we can write

$$u(x,t) = p_0(x) + p_1(x)t + \cdots + p_{k-1}(x)t^{k-1}$$

for some smooth functions p_i's. Since $u \in \mathcal{P}_d$, for fixed x we have

$$|u(x,t)| \le C(1+\sqrt{-t})^d,$$

so $k-1$ could at most be d. We may just write $k = d$ (with probably some $p_i = 0$).

Since $(\partial_t - \Delta)u = 0$, the polynomial expression tells us

$$\sum_i ip_i(x)t^{i-1} = \sum_i (\Delta p_i)t^i.$$

This holds for all $t \le 0$, so

$$(\Delta p_i) = (i+1)p_{i+1}$$

for all i. In particular, $\Delta p_d = 0$. Thus we could consider

$$\Phi_d \colon \mathcal{P}_d(M) \to \{\text{harmonic functions on } M\}, u \mapsto p_d.$$

We claim that

$$|p_i(x)| \le C(1+r_p)^{2d-2i} \tag{13.6}$$

where $r_p = d(p,\cdot)$. If so, p_d is a bounded harmonic function. Thus the image of Φ_d is contained in $\mathcal{H}_0(M)$. Therefore,

$$\dim \mathcal{P}_d \le \dim\ker\Phi_d + \dim\mathcal{H}_0. \tag{13.7}$$

If $u \in \ker\Phi_d$, i.e., $p_d = 0$, then

$$u = p_0(x) + p_1(x)t + \cdots + p_{d-1}(x)t^{d-1}.$$

Then we consider

$$\Phi_{d-1}(u) := p_{d-1}$$

for $u \in \ker\Phi_d$. Then (13.6) implies $\Phi_{d-1}(u) \in \mathcal{H}_2$. In particular,

$$\dim\ker\Phi_d \le \dim\mathcal{H}_2 + \dim\ker\Phi_{d-1}.$$

Inductively, by considering Φ_{d-2} on $\ker\Phi_{d-1}$, we have

$$\dim\ker\Phi_{d-1} \le \dim\mathcal{H}_4 + \dim\ker\Phi_{d-2}.$$

Summing over these inequalities and combining with (13.7), the conclusion follows. $\square$

We remark that ancient solutions to the heat equation come up naturally in blow-up arguments.

14. **Mean Curvature Flow**

We now go into the topic of mean curvature flow as an example of a nonlinear PDE of the type of heat equations. Let Σ^n be a smooth submanifold in $\mathbb{R}^N$. For vector fields X and Y on Σ, the second fundamental form of Σ is defined by

$$A(X,Y) := (\nabla_X Y)^\perp,$$

whose value at $p \in \Sigma$ depends only on $X(p)$ and $Y(p)$ (so forms a tensor). It is a symmetric 2-tensor. Pointwisely, we can think of it as a map

$$A \colon T_p\Sigma \times T_p\Sigma \to N_p\Sigma.$$

The mean curvature vector of Σ is a normal vector field defined by[1]

$$\mathbf{H} := -\mathrm{tr}A = -\sum_i A(e_i, e_i)$$

for an ONB e_i's at p. When $N = n+1$, i.e., Σ is a hypersurface in $\mathbb{R}^{n+1}$, we would define the mean curvature by

$$H = \langle \mathbf{H}, \mathbf{n} \rangle = -\mathrm{tr}\,\langle A, \mathbf{n} \rangle = -\sum_i \left\langle \nabla^\perp_{e_i} e_i, \mathbf{n} \right\rangle,$$

where $\mathbf{n}$ is the outer unit normal. SInce $\langle e_i, \mathbf{n} \rangle = 0$, the Leibniz rule tells us

$$H = -\sum_i \left\langle \nabla^\perp_{e_i} e_i, \mathbf{n} \right\rangle = \sum_i \langle e_i, \nabla_{e_i}\mathbf{n} \rangle = \mathrm{tr}\,\langle \cdot, \nabla_\cdot \mathbf{n} \rangle.$$

For tangent vectors X and Y, note that the same (Leibniz) reason implies

$$\langle \nabla_X \mathbf{n}, Y \rangle = -\langle \mathbf{n}, \nabla_X Y \rangle,$$

which is symmetric, so $\langle \cdot, \nabla_\cdot \mathbf{n} \rangle$ is also symmetric, called the Weingarten map.

14.1. **First Variation of Volume.**

For a closed submanifold $\Sigma^n \subseteq \mathbb{R}^N$ and a vector field V with compact support consider

$$\Sigma_{s,V} := \{x + sV(x) | x \in \Sigma\},$$

which forms a variation of Σ. We want to calculate the variation of its volume, i.e.,

$$\frac{d}{ds}\mathrm{Vol}(\Sigma_{s,V})|_{s=0}.$$

Looking at the volume element $d\mathrm{Vol}_{\Sigma_{s,V}}$ of $\Sigma_{s,V}$, we can derive a standard formula

$$\frac{d}{ds} d\mathrm{Vol}_{\Sigma_{s,V}} = \mathrm{div}_\Sigma V \cdot d\mathrm{Vol}.$$

(Recall that $\mathrm{div}_\Sigma V = \sum_i \langle \nabla_{e_i} V, e_i \rangle$ for an ONB e_i's.) As a result, (after omitting dVol)

$$\frac{d}{ds}\mathrm{Vol}(\Sigma_{s,V})|_{s=0} = \int_\Sigma \mathrm{div}_\Sigma V.$$

We could write $V = V^T + V^\perp$ with respect to Σ. By Stokes' theorem, since Σ is closed, we have

$$\int_\Sigma \mathrm{div}_\Sigma X = 0$$

[1]There are different conventions about the sign of the mean curvature vector.

for any tangent vector X of Σ. Consequently, we have

$$\frac{d}{ds}\mathrm{Vol}(\Sigma_{s,V})|_{s=0} = \int_\Sigma \mathrm{div}_\Sigma V^\perp.$$

This matches our intuition, since tangential variation would only reparametrize the hypersurface. Moreovre, we have

$$\mathrm{div}_\Sigma V^\perp = \sum_i \left\langle \nabla_{e_i} V^\perp, e_i \right\rangle = -\sum_i \left\langle V^\perp, \nabla_{e_i} e_i \right\rangle = \left\langle V^\perp, \mathbf{H} \right\rangle = \langle V, \mathbf{H} \rangle$$

since $\mathbf{H}$ is already normal. In conclusion, we get the first variation formula

$$\frac{d}{ds}\mathrm{Vol}(\Sigma_{s,V})|_{s=0} = \int_\Sigma \langle V, \mathbf{H} \rangle. \tag{14.1}$$

This formula goes back to Euler (1740) and Lagrange (1755). In fact, what Euler looked at was a special submanifold: catenoids. When he calculated the first variation of catenoids, he found that it is zero! He put this into his books, and later Lagrange read them. Lagrange put this into a more systematic setting, which could be regarded as the beginning of calculus of variation. Lagrange concluded that if the first variation vanishes for all V, in particular when $V = H$, then $H = 0$.

We could do the same thing for all submanifold with V having compact support. A submanifold is said to be **minimal** if the first variation is zero for all V, which is equivalent to $H = 0$ by the first variation formula (14.1). A conspocuous example of this kind of submanifolds are soap films, which is related to the Plateau problem, later solved by Douglas.

14.2. **Flowing Submanifolds.** How are minimal submanifolds related to heat equations? The idea is to flow the submanifold such that the volume decreases as soon as possible. By the first variation formula (14.1), it is reasonable to consider the flow direction as $V = -\mathbf{H}$. As a result, a one-parameter family of submanifolds Σ_t^n in $\mathbb{R}^N$ is said to **evolve by the mean curvature flow (MCF)** if its position vector x satisfies

$$x_t = -\mathbf{H}. \tag{14.2}$$

This is a nonlinear equation. In fact, this is a nonlinear version of the heat equation. In general, for a function $f\colon \mathbb{R}^N \to \mathbb{R}$ and $\Sigma \subseteq \mathbb{R}^N$, we recall

$$\Delta_\Sigma f = \mathrm{div}_\Sigma\left(\nabla^T f\right).$$

By definition, we have

$$\begin{aligned}
\Delta_\Sigma f &= \mathrm{div}_\Sigma\left(\nabla^T f\right) \\
&= \mathrm{div}_\Sigma(\nabla f) - \mathrm{div}_\Sigma(\nabla^\perp f) \\
&= \sum_i \langle \nabla_{e_i} \nabla f, e_i \rangle - \mathrm{div}_\Sigma(\nabla^\perp f) \\
&= \sum_i \mathrm{Hess}_f(e_i, e_i) - \mathrm{div}_\Sigma(\nabla^\perp f)
\end{aligned}$$

for an ONB for $T_p\Sigma$. Note that

$$\mathrm{div}_\Sigma(\nabla^T f) = \sum_i \left\langle \nabla_{e_i}\nabla^\perp f, e_i \right\rangle = -\sum_i \left\langle \nabla^\perp f, \nabla_{e_i} e_i \right\rangle = \left\langle \nabla^\perp f, H \right\rangle = \langle \nabla f, H \rangle .$$

In conclusion, we derive

$$\Delta_\Sigma f = \sum_i \mathrm{Hess}_f(e_i, e_i) - \langle \nabla f, \mathbf{H} \rangle . \tag{14.3}$$

There are some special examples.

1. $f = x_i$, coordinate functions. In this case, $\mathrm{Hess}_{x_i} = 0$, so

$$\nabla_\Sigma x_i = -\langle e_i, \mathbf{H} \rangle .$$

Thus

$$\Delta_\Sigma x = -\sum_i \langle e_i, \mathbf{H} \rangle e_i = -\mathbf{H}. \tag{14.4}$$

2. $f = |x|^2$. In this case, we have $\mathrm{Hess}_{|x|^2} = 2\langle \cdot, \cdot \rangle$ and $\nabla |x|^2 = 2x$, so

$$\Delta_\Sigma |x|^2 = 2n - 2\langle x, \mathbf{H} \rangle . \tag{14.5}$$

When Σ_t is evolving by the MCF, based on (14.4), the MCF equation (14.2) becomes

$$x_t = \Delta_\Sigma x, \tag{14.6}$$

which is just the heat euqation. The point is that here the Laplacian operator depends on t is a fairly complicated way. More generally, for a C^2 function $f\colon \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$, we could consider $f\colon \Sigma_t \to \mathbb{R}$ for any t. Then

$$\partial_t f = \partial_t f(x(t), t) = \langle \nabla f, x_t \rangle + \frac{\partial f}{\partial t} = -\langle \nabla f, \mathbf{H} \rangle + \frac{\partial f}{\partial t}.$$

When $f = |x|^2$, we have

$$\partial_t |x|^2 = -2\langle x, \mathbf{H} \rangle = \Delta_\Sigma |x|^2 - 2n \tag{14.7}$$

based on (14.5). As a consequence, by the parabolic maximum principle (applied to (14.6) (14.7)), if Σ_0 lies in a half-space or a ball, it will remain so for any t.

14.3. **Examples.** We see some classic examples of the MCF.

1. Planes. For planes, $\mathbf{H} = 0$. Thus they form a static solution.

2. Spheres. Let Σ be the sphere of radius r_0. Then $\mathbf{H} = \frac{n}{r_0}$ with outward direction.Thus $-\mathbf{H}$ is pointing inward, i.e., toward the center of the sphere. By symmetry, it will remain spherical, and its radius r satisfies

$$r_t = -\frac{n}{r},$$

which admits the solution

$$r = \sqrt{r_0^2 - 2nt}.$$

Observe that when $2nt = r_0^2$, we have $r = 0$. i.e., it shrinks to a point, so the flow could make sense anymore.

3. Cylinders. In general, we could take $S^k \times \mathbb{R}^{n-k}$. Since the mean curvature is pointing toward the axis, it will shrink to the axis. The speed is given by the spherical factor (as computed in the previous example).

There are many other examples, but it turns out that these are the most important models. Note that these examples have the same feature: they are shrinking homothetically under the MCF. We will see some of other meanings of a shrinker later.

14.4. **Momotonicity and Shrinkers.** Watson first observed the mean value equality. Suppose $u, v\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ satisfies $(\partial_t - \Delta)u = (\partial_t + \Delta)v = 0$ and decay in a suitable sense. Then

$$\frac{d}{dt}\int uv = \int u_t v + \int uv_t = \int v\Delta u + \int uv_t = \int u\Delta v + \int uv_t = 0$$

after integration by parts. Based on this, we have the mean value equality.

Theorem 14.8. Let $v(x,t) = (-4\pi t)^{-\frac{n}{2}} e^{\frac{|x|^2}{4t}}$, which satisfies the backward heat equation. If u is a solution to the heat equation, then

$$u(0,0) = (-4\pi t)^{-\frac{n}{2}} \int u(x,t) e^{\frac{|x|^2}{4t}}.$$

We want to get a more general monotonicity formula. Suppose Σ_t is flowing by the MCF. Then the variation of the volume form gives

$$\begin{aligned}\frac{d}{dt}\int_{\Sigma_t} uv &= \int_{\Sigma_t} u_t v + \int_{\Sigma_t} uv_t - \int_{\Sigma_t} uv \cdot |\mathbf{H}|^2 \\ &= \int u\left((\partial_t + \Delta_{\Sigma_t})v - |\mathbf{H}|^2 v\right)\end{aligned}$$

for any v and a solution u to the heat equation. Thus $\int_{\Sigma_t} uv$ is constant if

$$(\partial_t + \Delta_{\Sigma_t})v - |\mathbf{H}|^2 v = 0,$$

called the conjugate equation to the heat equation. Based on this, Huisken looked at

$$\Phi(x,t) = (-4\pi t)^{-\frac{n}{2}} e^{\frac{|x|^2}{4t}}$$

for $(x,t) \in \mathbb{R}^N \times (-\infty, 0)$, where n is the dimension of the evolving submanifold, and want to know

$$\left(\partial_t + \Delta_{\Sigma_t} - |\mathbf{H}|^2\right)\Phi.$$

Before doing the calculation, first we observe that

$$(\partial_t + \Delta_{\Sigma_t})e^u = e^u u_t + e^u \Delta_{\Sigma_t} u + e^u |\nabla^{\Sigma_t} u|^2$$

so

$$e^{-u}(\partial_t + \Delta_{\Sigma_t})e^u = (\partial_t + \Delta_{\Sigma_t})u + |\nabla^{\Sigma_t} u|^2. \tag{14.9}$$

Thus we could take

$$u = \log \Phi = -\frac{n}{2}\log(-4\pi t) + \frac{|x|^2}{4t},$$

based on which we have

$$\nabla u = \frac{x}{2t},$$

$$|\nabla^T u|^2 = \frac{|x^T|^2}{4t^2},$$

$$\Delta_{\Sigma_t} u = \frac{n}{2t} - \left\langle \frac{x}{2t}, \mathbf{H} \right\rangle$$

by (14.3) since $\operatorname{Hess}_{\frac{|x|^2}{4t}} = \frac{\delta_{ij}}{2t}$, and finally

$$\partial_t u = -\frac{n}{2t} - \frac{|x|^2}{4t} - \frac{2\langle x, \mathbf{H}\rangle}{4t}$$

since $x_t = -\mathbf{H}$. In conclusion, we have

$$\partial_t u + \Delta_{\sigma_t} u + |\nabla^T u|^2 = -\frac{|x|^2}{4t^2} - \left\langle \frac{x}{t}, \mathbf{H} \right\rangle + \frac{|x^T|^2}{4t^2} = -\frac{|x^\perp|^2}{4t^2} - \left\langle \frac{x}{t}, \mathbf{H} \right\rangle.$$

Therefore,

$$\begin{aligned}
\partial_t u + \Delta_{\sigma_t} u + |\nabla^T u|^2 - |\mathbf{H}|^2 &= -\frac{|x|^2}{4t^2} - \left\langle \frac{x}{t}, \mathbf{H} \right\rangle + \frac{|x^T|^2}{4t^2} \\
&= -\frac{|x^\perp|^2}{4t^2} - \left\langle \frac{x^\perp}{t}, \mathbf{H} \right\rangle - |\mathbf{H}|^2 \\
&= -\left| \mathbf{H} + \frac{x^\perp}{2t} \right|^2.
\end{aligned}$$

Then based on (14.9), we have

$$\left(\partial_t + \Delta_{\Sigma_t} - |\mathbf{H}|^2\right)\Phi = -\left|\mathbf{H} + \frac{x^\perp}{2t}\right|^2 \Phi. \tag{14.10}$$

This was first done by Huisken [Hui90]. Based on this, we can have

$$\begin{aligned}
\frac{d}{dt}\int u\Phi &= \int u_t \Phi + \int u\Phi_t - \int u\Phi|\mathbf{H}|^2 \\
&= \int \Phi \Delta_{\Sigma_t} u + \int u\Phi_t - \int u\Phi|\mathbf{H}|^2 \\
&= \int u \Delta_{\Sigma_t} \Phi + \int u\Phi_t - \int u\Phi|\mathbf{H}|^2 \\
&= \int u\left(\partial_t + \Delta_{\Sigma_t} - |\mathbf{H}|^2\right)\Phi \\
&= -\int u \left|\mathbf{H} + \frac{x^\perp}{2t}\right|^2 \Phi.
\end{aligned}$$

if u satisfies the heat equation. This is called Huisken's monotonicity. In particular, if $u \geq 0$ (or just $u(\cdot, 0) \geq 0$ by the parabolic maximum principle), then $\int_{\Sigma_t} u\Phi$ is increasing. The most important case is when $u = 0$. By a change of variables, if we consider

$$\Phi_{x_0,t_0} = (4\pi(t_0 - t))^{-\frac{n}{2}} e^{\frac{|x-x_0|^2}{4(t-t_0)}},$$

we in general have

$$\frac{d}{dt}\int_{\Sigma_t} \Phi_{x_0,t_0} = -\int \left| H + \frac{(x-x_0)^\perp}{2(t-t_0)} \right|^2 \Phi_{x_0,t_0}. \tag{14.11}$$

When is this monotone quantity constant? Consider $\Phi = \Phi_{0,0}$. Since

$$\frac{d}{dt}\int \Phi = -\int \left| \mathbf{H} + \frac{x^\perp}{2t} \right|^2 \Phi,$$

it is constant if and only if

$$\mathbf{H}_{\Sigma_t} + \frac{x^\perp}{2t} = 0. \tag{14.12}$$

This is equivalent to the condition

$$\mathbf{H}_{\Sigma_{-1}} = \frac{x^\perp}{2}. \tag{14.13}$$

In fact, once (14.13) holds, the flow $M_t := \sqrt{-t}\Sigma_{-1}$ evolves by the MCF because $x(t) = \sqrt{-t}y$ ($y \in \Sigma_{-1}$) satisfies

$$x_t = -\frac{1}{2\sqrt{-t}} y^\perp$$

and hence

$$\mathbf{H}_{M_t} = \frac{\mathbf{H}_{\Sigma_{-1}}}{\sqrt{-t}} = \frac{y^\perp}{2\sqrt{-t}} = -x_t.$$

Thus (14.12) holds. As a result, if Σ satisfies

$$\mathbf{H}_\Sigma = \frac{x^\perp}{2}, \tag{14.14}$$

we call Σ a shrinker. We have seen some examples of shrinkers in the section 14.3.

15. **Shrinkers**

This section is based on [CM12].

15.1. **F-functional and Entropy.** First, we define an important quantity for shrinkers. We could define it in a more general setting. For a submanifold Σ^n of $\mathbb{R}^N$, we define

$$F(\Sigma) := (4\pi)^{-\frac{n}{2}} \int_\Sigma e^{-\frac{|x|^2}{4}},$$

called the F-functional (a kind of Lyapunov functional in this dynamic system), and

$$\lambda(\Sigma) := \sup_{t_0>0, x_0\in\mathbb{R}^N} F(t_0\Sigma + x_0),$$

called the entropy of Σ. The crucial property of λ is its monotonicity under the MCF. In general, thes quantities could be infinity. Thus we consider $\Sigma \subseteq \mathbb{R}^N$ with Euclidean volume growth, in the sense that

$$\mathrm{Vol}(B_r(x) \cap \Sigma) \le C_\Sigma r^n$$

for all $r > 0$ and $x \in \mathbb{R}^N$.

Lemma 15.1. If $\Sigma^n \subseteq \mathbb{R}^N$ has Euclidean volume growth, then $F(\Sigma) < \infty$. In fact, we have

$$F(\Sigma) \le C_\Sigma \cdot \widetilde{c}_n$$

with C_Σ the constant in the definition of the volume growth.

(Proof.) By definition,

$$F(\Sigma) = (4\pi)^{-\frac{n}{2}} \sum_{i=0}^\infty \int_{B_{2^{i+1}}\setminus B_{2^i}\cap\Sigma} e^{-\frac{|x|^2}{4}} + (4\pi)^{-\frac{n}{2}} \int_{B_1\cap\Sigma} e^{-\frac{|x|^2}{4}}.$$

Note that

$$\begin{aligned}\int_{B_{2^{i+1}}\setminus B_{2^i}\cap\Sigma} e^{-\frac{|x|^2}{4}} &\le \mathrm{Vol}(B_{2^{i+1}} \cap \Sigma) \cdot e^{-\frac{2^{2i}}{4}} \\ &\le C_\Sigma \cdot 2^{n(i+1)} \cdot e^{-\frac{2^{2i}}{4}},\end{aligned}$$

so

$$F(\Sigma) \le (4\pi)^{-\frac{n}{2}} \sum_{i=0}^\infty C_\Sigma \cdot 2^{n(i+1)} \cdot e^{-\frac{2^{2i}}{4}} + (4\pi)^{-\frac{n}{2}} \cdot \mathrm{Vol}(B_1) = C_\Sigma \cdot \widetilde{c}_n. \quad \square$$

Now suppose Σ has Euclidean volume growth. Then note that $t_0\Sigma$ and $\Sigma + x_0$ also have Euclidean volume growth with the same constant for any $t_0 > 0$ and $x_0 \in \mathbb{R}^N$. As a result, lemma 15.1, in particular, implies that $\lambda(\Sigma) < \infty$.

We mention one particular example. If Σ is a closed submanifold, clearly

$$\mathrm{Vol}(B_r(x) \cap \Sigma) \le \mathrm{Vol}(\Sigma)$$

for all $r > 0$ and $x \in \mathbb{R}^N$. Also, when r is small, $B_r(x) \cap \Sigma$ is close to an n-dimensional Euclidean ball, so we get the volume growth bound.

Surprisingly, we have the following converse statement.

Proposition 15.2. If $\lambda(\Sigma) < \infty$, then Σ has Euclidean volume growth.

(Proof.) By assumption, for any x_0 and r, $F(\Sigma - x_0) < \infty$. However,

$$F(\Sigma - x_0) = (4\pi)^{-\frac{n}{2}} \int_{\Sigma - x_0} e^{-\frac{|x|^2}{4}},$$

so

$$\mathrm{Vol}(B_1(x_0) \cap \Sigma) \leq (4\pi)^{-\frac{n}{2}} \cdot e^{-\frac{1}{4}} \cdot F(\Sigma - x_0) \leq C\lambda(\Sigma).$$

Thus the conclusion follows from scaling. □

Note that after a change of variables, we can rewrite

$$F(t_0\Sigma) = (4\pi t_0)^{-\frac{n}{2}} \int_{\Sigma} e^{-\frac{|x|^2}{4t_0}}$$

and

$$F(\Sigma + x_0) = (4\pi)^{-\frac{n}{2}} \int_{\Sigma} e^{-\frac{|x-x_0|^2}{4}}.$$

These provide another way to think of entropy as the supremum after taking integrals of different Gaussians, i.e.,

$$\lambda(\Sigma) = \sup_{t_0 > 0, x_0 \in \mathbb{R}^N} (4\pi t_0)^{-\frac{n}{2}} \int_{\Sigma} e^{-\frac{|x-x_0|^2}{4t_0}}. \tag{15.3}$$

This interpretation is nice when combining with Huisken's monotonicity (14.11), which says that the quantity

$$(4\pi(t_0 - t))^{-\frac{n}{2}} \int_{\Sigma_t} e^{-\frac{|x-x_0|^2}{4(t_0-t)}}$$

is decreasing along the MCF. This, along with (15.3), proves that $\lambda(\Sigma_t)$ is also decreasing along the MCF, and in particular, is constant if and only if Σ is a shrinker (i.e., satisfying (14.14)).

15.2. **Operators.** Next, we could ask the following euqation. What are the natural operators on shrinkers? One is an analog of Ornstein-Uhlenbeck opeartors. For a function u on a shrinker Σ, define

$$\mathcal{L}u := \Delta_\Sigma u - \frac{1}{2}\left\langle x, \nabla^\Sigma u \right\rangle := \Delta_\Sigma u - \frac{1}{2}\left\langle x^T, \nabla u \right\rangle.$$

For example,

$$\mathcal{L}x_i = \Delta_\Sigma x_i - \frac{1}{2}\left\langle x, e_i^T \right\rangle = -\left\langle e_i, \frac{x^\perp}{2} \right\rangle - \frac{1}{2}\left\langle x, e_i^T \right\rangle = -\frac{1}{2}x, e_i = -\frac{1}{2}x_i$$

by the shrinker equation and (14.3) since $\mathrm{Hess}_{x_i} = 0$. In other words, x_i's are eigenfunctions for $\mathcal{L}$ with eigenvalue $\frac{1}{2}$. Another nice function is $|x|^2$. Similar argument gives

$$\begin{aligned}
\mathcal{L}|x|^2 &= \Delta_\Sigma |x|^2 - \frac{1}{2}\left\langle x, \nabla^T |x|^2 \right\rangle \\
&= \sum_i \mathrm{Hess}_{|x|^2}(e_i, e_i) - \left\langle \nabla |x|^2, \mathbf{H} \right\rangle - \frac{1}{2}\left\langle x, \nabla^T |x|^2 \right\rangle \\
&= 2n - 2\left\langle x, \frac{x^\perp}{2} \right\rangle - \frac{1}{2}\left\langle x, 2x^T \right\rangle \\
&= 2n - |x|^2.
\end{aligned}$$

As a result, we can rewrite it as

$$\mathcal{L}(2n - |x|^2) = -(2n - |x|^2).$$

That is to say, $2n - |x|^2$ is an eigenfunction with eigenvalue 1.

15.3. **Rescaled MCF and the First Variational Formula for F.** Suppose Σ_t is evolving by the MCF (for $t < 0$). We consider a scaling

$$\frac{1}{\sqrt{-t}}\Sigma_t$$

and a reparametrization $t = -e^{-s}$. i.e., the new flow is

$$\Gamma_s = e^{\frac{s}{2}} \Sigma_{-e^{-s}}.$$

Then

$$y(s) = e^{\frac{s}{2}} x(-e^{-s}),$$

and hence

$$\begin{aligned}
y_s^\perp &= \frac{1}{2} e^{\frac{s}{2}} x^\perp + e^{\frac{s}{2}} \cdot e^{-s} x_t^\perp \\
&= \frac{1}{2} y^\perp - e^{-\frac{s}{2}} \cdot \mathbf{H}_{\Sigma_t} \\
&= \frac{1}{2} y^\perp - \mathbf{H}_{\Gamma_s}
\end{aligned}$$

since $\mathbf{H}_{\Gamma_s} = e^{-\frac{s}{2}} \mathbf{H}_{\Sigma_t}$ by definition. This is called the **rescaled MCF equation**. This provides another viewpoint of shrinkers, that is, fixed points for the rescaled MCF.

In fact, there is another point of view to the rescaled MCF. Consider a variation of Σ as

$$\Sigma_{s,V} := \{x + sV(x) | x \in \Sigma\}.$$

Then by the first variational formula for volume,

$$\begin{aligned}
\frac{d}{ds} F(\Sigma_{V,s})|_{s=0} &= (4\pi)^{-\frac{n}{2}} \int_\Sigma \left(\langle \mathbf{H}, V \rangle - \left\langle \frac{x^\perp}{2}, V \right\rangle \right) e^{-\frac{|x|^2}{4}} dx \\
&= (4\pi)^{-\frac{n}{2}} \int_\Sigma \left\langle V, \mathbf{H} - \frac{x^\perp}{2} \right\rangle e^{-\frac{|x|^2}{4}} dx.
\end{aligned}$$

Thus Σ is a shrinker if and only if it is a critical point for the F-functional. Moreover, the negative gradient vector field for F is

$$V = -\left(\mathbf{H} - \frac{x^\perp}{2}\right),$$

which is just given by the rescaled MCF. In particular, the F-functional is decreasing along the rescaled MCF.

In general, for a submanifold $\Sigma^n \subseteq \mathbb{R}^N$, recall that we have (14.3)

$$\Delta_\Sigma f = \sum_i \mathrm{Hess}_f(e_i, e_i) - \langle \nabla f, \mathbf{H} \rangle,$$

which helps us to derive

$$\Delta_\Sigma x_i = -\left\langle e_i^\perp, \mathbf{H} \right\rangle.$$

If Σ_t flows by the rescaled MCF, i.e.,

$$x_t = -\mathbf{H} + \frac{x^\perp}{2},$$

then we have

$$\partial_t x_i = -\left\langle e_i^\perp, \mathbf{H} \right\rangle + \frac{1}{2}\left\langle x, e_i^\perp \right\rangle$$

and

$$\mathcal{L} x_i = -\left\langle x_i^\perp, \mathbf{H} \right\rangle - \frac{1}{2}\left\langle x, e_i^\perp \right\rangle.$$

As a result, we have

$$(\partial_t - \mathcal{L}) x_i = \frac{1}{2}\langle x, e_i \rangle = \frac{1}{2} x_i.$$

Similarly, we can derive

$$(\partial_t - \mathcal{L})|x|^2 = |x|^2 - 2n.$$

16. **Pseudolocality**

Consider a submanifold $\Sigma^n \subseteq \mathbb{R}^N$. Recall that

$$F(\Sigma) = (4\pi)^{-\frac{n}{2}} \int_\Sigma e^{-\frac{|x|^2}{4}}$$

and

$$\lambda(\Sigma) = \sup_{t_0 > 0, x_0 \in \mathbb{R}^N} F(t_0 \Sigma + x_0) = \sup_{t_0 > 0, x_0 \in \mathbb{R}^N} (4\pi t_0)^{-\frac{n}{2}} \int_\Sigma e^{-\frac{|x-x_0|^2}{4t_0}}$$

as we mentioned in the preceding section. Based on Huisken's monotonicity, we know that $\lambda(\Sigma_t)$ is decreasing if Σ_t evolves under the MCF. Also, we have seen that Σ has Euclidean volume growth if and only if $\lambda(\Sigma) < \infty$ (lemma 15.1 and proposition 15.2).

16.1. **First Observation.** More generally, we can consider

$$F_{x_0,t_0}(\Sigma) := (4\pi t_0)^{-\frac{n}{2}} \int_\Sigma e^{-\frac{|x-x_0|^2}{4t_0}}$$

for the sake of convenience. We would like to compare $F_{x_0,t_0}(\Sigma)$ with $F_{0,1}(\Sigma)$ (which is just $F(\Sigma)$). We may expet that $F_{x_0,t_0}(\Sigma)$ is close to $F(\Sigma)$ if (x_0, t_0) is close to $(0, 1)$ (and hopefully this closeness is independent of Σ).

Now suppose $\lambda(\Sigma) \le \lambda_0 < \infty$. Then

$$\mathrm{Vol}(B_r(x) \cap \Sigma) \le c(n)\lambda_0 r^n,$$

for any x and r based on proposition 15.2. Then we can write

$$\begin{aligned} F(\Sigma) &= (4\pi)^{-\frac{n}{2}} \int_{\Sigma \cap B_R} e^{-\frac{|x|^2}{4}} + (4\pi)^{-\frac{n}{2}} \int_{\Sigma \setminus B_R} e^{-\frac{|x|^2}{4}} \\ &\le (4\pi)^{-\frac{n}{2}} \int_{\Sigma \cap B_R} e^{-\frac{|x|^2}{4}} + (4\pi)^{-\frac{n}{2}} \sum \int_{(B_{2^{i+1}R} \setminus B_{2^i R}) \cap \Sigma} e^{-\frac{|x|^2}{4}} \\ &\le (4\pi)^{-\frac{n}{2}} \int_{\Sigma \cap B_R} e^{-\frac{|x|^2}{4}} + (4\pi)^{-\frac{n}{2}} \sum_{i=0}^{\infty} e^{-\frac{(2^i R)^2}{4}} \mathrm{Vol}(B_{2^{i+1}} \cap \Sigma) \\ &\le (4\pi)^{-\frac{n}{2}} \int_{\Sigma \cap B_R} e^{-\frac{|x|^2}{4}} + C\lambda_0 \cdot e^{-cR^2} \end{aligned}$$

by the polynomial growth. Thus by chopping out, we only need to care about the part inside a large ball, i.e., the difference

$$\int_{\Sigma \cap B_R} e^{-\frac{|x|^2}{4}} - t_0^{-\frac{n}{2}} \int_{\Sigma \cap B_R} e^{-\frac{|x-x_0|^2}{4t_0}}.$$

However, the two integrands are close when (x_0, t_0) is close to $(0, 1)$. Since R is fixed, we can see that $F(\Sigma)$ and $F_{x_0,t_0}(\Sigma)$ are close to each other.

16.2. **Natural Distances on Space-time.** We define the parabolic distance

$$d^P((x_1, t_1), (x_2, t_2)) = \max\{|x_1 - x_2|, \sqrt{|t_1 - t_2|}\},$$

which is a proper metric on $\mathbb{R}^N \times \mathbb{R}$. Based on this, we can define backward parabolic ball

$$B_r^P(x, s) := \{(y, t) \in \mathbb{R}^N \times \mathbb{R} : t \le s \text{ and } d^P((x, s), (y, t)) < r\}.$$

Note that when $N = 0$ and we equip $\mathbb{R}$ with this funny distance, then we can calculate that the Hausdorff dimension of intervals in this real line is 2. As a result, the dimension of $\mathbb{R}^n \times \mathbb{R}$ is $n + 2$ if we equip it with the parabolic distance.

In general, on a metric space (X, d), for $y \in B_r(x) \subseteq X$ and $r_0 = d(y, x) < r$, we have

$$B_{r-r_0}(y) \subseteq B_r(x)$$

based on the triangle inequality. In our case, in particular, for $(y, t) \in B_r^P(x, s)$ and $r_0 = d^P((y, t), (x, s)) < r$, we can conclude

$$B_{r-r_0}^P(y, t) \subseteq B_r^P(x, s)$$

with the trivial fact that the point $(z, u) \in B^P_{r-r_0}(y, t)$ satisfies $u \le t \le s$.

16.3. **Main Theorem.** Let Σ_t be a one-parameter family of submanifolds (which will evolve under the MCF finally but here we could consider a general setting), and f be a non-negative function defined on

$$\{(x, s) \in B^P_1(0, 0) : x \in \Sigma_s\},$$

by

$$f(x, s) := (1 - d^P((x, s), (0, 0)))|A_{\Sigma_s}|^2(x),$$

where $|A|^2$ is the Hilbert-Schmidt norm of the second fundamental form (that is,

$$|A|^2 = \sum_{i,j} |A(e_i, e_j)|^2$$

for any ONB e_i's for $T_x\Sigma_s$). Note that $1 - d^P((x, s), (0, 0))$ is the distance to the boundary of $B^P_1(0, 0)$.

The idea of blow-up (that we are going to introduce) first came from the study of minimal submanifolds. Note that by definition, f is a function on $B^P_1(0, 0)$ that vanished on $\partial B^P_1(0, 0)$. Thus we may assume f achieves its maximum at $(x, s) \in B^P_1(0, 0)$, which must be an interior point (unless $f \equiv 0$, which is not interesting). Now we are in a position to state the pseudolocality theorem, which is essentially due to Brakke [Bra78] and White [Whi05].

Theorem 16.1 (Pseudolocality)**.** Assume Σ_t is flowing by the MCF for $t \le 0$ with $(0, 0) \in \Sigma_0$ and $\lambda(\Sigma_{-1}) \le \lambda_0 < \infty$. If

$$F(\Sigma_{-1}) \le 1 + \varepsilon$$

for sufficiently small $\varepsilon > 0$, then

$$|A_{\Sigma_0}|(0) \le 1.$$

(Ideas of the proof.) We will prove this theorem by contradiction. That is, assume $|A_{\Sigma_0}|(0) > 1$. Then the function

$$f(x, s) = (1 - d^P((x, s), (0, 0)))|A_{\Sigma_s}|^2(x)$$

now satisfies

$$\max f > 1,$$

since $f(0, 0) > 1$. Suppose f achieves its maximum at (x, s). Then after rescaling, we may assume $|A_{\Sigma_0}|(0) = 1$ and $|A| \le 2$ on B^P_1. Restricting to a smaller ball $B^P_\delta(0, 0)$, we adjust f by considering

$$f(x, s) = (\delta - d^P((x, s), (0, 0)))|A_{\Sigma_s}|^2(x).$$

Then we have $F_{x,t}(\Sigma_{-1}) < 1 + \varepsilon$ for $(x, t) \in B^P_\delta(0, 0)$ if δ is small enough by the discussion in the section 16.1. Thus it is virtually constant. Therefore, Huisken's monotonicity implies it should be a shrinker (here to make it rigorous we need to extract a sequence and see its limit). However, on an infinitesimal scale, the shrinker is close to a plane (since the contradiction assumption holds for any small ε thus the limiting submanifold has unit Gaussian integral), which contradicts to the non-vanishing of the second fundamental form. □

References

[Bra78] Kenneth A. Brakke, *The motion of a surface by its mean curvature*, Mathematical Notes, vol. 20, Princeton University Press, Princeton, N.J., 1978. MR0485012

[CM97] Tobias H. Colding and William P. Minicozzi II, *Harmonic functions on manifolds*, Ann. of Math. (2) **146** (1997), no. 3, 725–747. MR1491451

[CM12] ———, *Generic mean curvature flow I: generic singularities*, Ann. of Math. (2) **175** (2012), no. 2, 755–833. MR2993752

[CM21] ———, *Optimal bounds for ancient caloric functions*, Duke Math. J. **170** (2021), no. 18, 4171–4182. MR4348235

[CM20] ———, *Parabolic frequency on manifolds* (2020). Preprint available at `https://arxiv.org/abs/2002.11015`.

[Ham93] Richard S. Hamilton, *The Harnack estimate for the Ricci flow*, J. Differential Geom. **37** (1993), no. 1, 225–243. MR1198607

[Hui90] Gerhard Huisken, *Asymptotic behavior for singularities of the mean curvature flow*, J. Differential Geom. **31** (1990), no. 1, 285–299. MR1030675

[LY86] Peter Li and Shing-Tung Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** (1986), no. 3-4, 153–201.

[Whi05] Brian White, *A local regularity theorem for mean curvature flow*, Ann. of Math. (2) **161** (2005), no. 3, 1487–1519. MR2180405