Lecture 12.

Log-Sobolev Inequality. Let us consider on R the generator

$$\mathcal{L} = \frac{1}{2}D_x^2 + \frac{1}{2}b(x)D_x$$

where $b(x) = \frac{\phi'(x)}{\phi(x)}$. Clearly $\mathcal{L}$ is self adjont with respect to the weight $\phi(x)$ and

$$\mathcal{L}u = \frac{1}{2\,\phi}D_x\phi(x)D_x u$$

or

$$\langle \mathcal{L}u, v\rangle = \int (\mathcal{L}u)(x)\,v(x)\,\phi(x)dx = -\frac{1}{2}\int u_x\,v_x\phi(x)dx$$

We will take $\phi(x) = \exp[-\psi(x)]$ with a uniformly convex ψ, i.e $\psi_{xx} \geq c > 0$. $b = -\psi'(x)$ and $b_x \leq -c < 0$. We let $f(t,x) \geq 0 \in L_1(\phi\,dx)$ evolve according to the equation

$$f_t(t,x) = (\mathcal{L}f)(t,x)$$

Since $\mathcal{L}$ is self adjoint $L^* = L$ with respect to the weight ϕ. We denote by

$$H(t) = \int f(t,x)\log f(t,x)\,\phi(x)dx$$

Then

$$\frac{dH}{dt} = -\frac{1}{2}\int \frac{[f_x(t,x)]^2}{f(t,x)}\phi(x)\,dx$$

We are interested in calculating

$$\frac{d^2H(t)}{dt^2} = -\frac{d}{dt}\frac{1}{2}\int \frac{[f_x(t,x)]^2}{f(t,x)}\phi(x)dx$$

which equals

$$\frac{1}{2}\int \frac{[f_x(t,x)]^2}{[f(t,x)]^2}(\mathcal{L}f)(t,x)\phi(x)dx - \int \frac{f_x(t,x)}{f(t,x)}(\mathcal{L}f)_x(t,x)\phi(x)dx$$

We note that

$$(\mathcal{L}f)_x = \mathcal{L}f_x + \frac{1}{2}b(x)f_x$$

Moreover

$$\begin{aligned}&\frac{1}{2}\int \frac{[f_x(t,x)]^2}{[f(t,x)]^2}(\mathcal{L}f)(t,x)\phi(x)dx\\ &\quad= \frac{1}{2}\langle \frac{f_x^2}{f^2}, \mathcal{L}f\rangle = \frac{1}{2}\langle \mathcal{L}(\frac{f_x^2}{f^2}), f\rangle\\ &\quad\geq \langle \frac{f_x}{f}\mathcal{L}(\frac{f_x}{f}), f\rangle = \langle \mathcal{L}(\frac{f_x}{f}), f_x\rangle = \langle (\frac{f_x}{f}), \mathcal{L}f_x\rangle\end{aligned}$$

Therefore

$$\frac{d^2H(t)}{dt^2} \ge -\frac{1}{2}\int \frac{b_x f_x^2}{f}\phi(x)dx \ge \frac{c}{2}\int \frac{[f_x(t,x)]^2}{f(t,x)}\phi(x)dx = -c\frac{dH(t)}{dt}$$

If we denote by $I(t) = -\frac{dH(t)}{dt}$, then $\frac{dI(t)}{dt} \le -cI(t)$, providing $\int_0^\infty I(s)ds \le \frac{1}{c}I(0)$. But

$$H(0) = \int_0^\infty I(t)dt \le \frac{1}{c}I(0)$$

We have assumed $H(\infty) = 0$. True for a dense set.

Suppose we are in R^d and we have a generator of the type

$$\mathcal{L}u = \frac{1}{2\phi}\nabla \cdot \phi C \nabla u$$

with a positive definite symmetric C, (independent of x) which is self adjoint with respect to the weight $\phi(x) = e^{-\psi(x)}$. Then

$$H(t) = \int f(t,x)\log f(t,x)\,\phi(x)\,dx$$

$$I(t) = \frac{1}{2}\int \frac{\langle C\nabla f, \nabla f\rangle}{f}\phi dx$$

The crucial step is to estimate

$$\frac{dI(t)}{dt} = -\frac{1}{2}\langle \mathcal{L}f, \frac{\langle C\nabla f, \nabla f\rangle}{f^2}\rangle + \int \frac{\langle C\nabla \mathcal{L} f, \nabla f\rangle}{f}\phi dx$$

and note that $\nabla \mathcal{L} f = \nabla \mathcal{L} f + (Db)\nabla f$. We need to bound

$$\langle C\nabla f, (Db) C\nabla f\rangle \le -c\langle C\nabla f, \nabla f\rangle$$

In our particular context we are looking at R^{N-1} represented as the hyperplane $\sum_{i=1}^N x_i = Nm$. C is the quadratic form

$$\langle Cu, u\rangle = \sum_{i=2}^N (u_i - u_{i-1})^2$$

$\widehat{\psi}$ is the restriction of $\sum_i \psi(x_i)$ to the hyperplane. $-Db$ is the Hessian of ψ. The matrix representing C is

$$\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 & 0 \\ \cdots & & & \cdots & & & \cdots \\ \cdots & & & \cdots & & & \cdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

It is not hard to see that $c = c_N = aN^{-2}$ where a is a lower bound on $\psi''(x)$.

Our density $f(x)$ on R^N is written as the superposition of densities $\mu_m = f_m(x)\lambda_m(dx)$ on the hyperplane relative to the conditionals λ_m, which are just $\pi\phi(x_i)$ normalized on the hyperplane.

$$f(x)dx = \int \mu_m(x)d\nu(m)$$

The entropy under control is

$$\int H(\mu_m|\lambda_m)d\nu(m) \le cN^2 I(f)$$

which is obtained on each hyperplane and integrated with respect to m.

Large deviation estimates. With respect to λ_m the probability

$$\lambda_m\Big[|\frac{1}{N}\sum g(x_i) - \widehat{g}(m)| \ge \delta] = \exp[-c(\delta)N]$$

where

$$\widehat{g}(m) = \frac{1}{M(\lambda)}\int g(x)e^{\lambda x}\phi(x)dx$$

with $\lambda = \lambda(m)$. Note that $N = N\epsilon$, $I(f) = \frac{\epsilon}{N}$. Therefore

$$\int H(\mu_m|\lambda_m)d\nu(m) \le cN\epsilon^3$$

Since $m = \bar{x}$, with respect to λ_m and μ_m, we have

$$\mu_m\Big[|\frac{1}{N}\sum g(x_i) - \widehat{g}(m)| \ge \delta] \le c(\delta)\epsilon^2$$

This estimate is valid uniformly over bounded set of values of m. Integrate w.r.t ν. Entropy controls integrability. For instance, since $\psi(x) \ge cx^2$, with $w(x) = \sqrt{1+x^2}$,

$$\int \exp[\sum_{i=1}^{N} w(x_i)]e^{-\sum_{i=1}^{N}\psi(x_i)}dx \le e^{CN}$$

If $f(t,x)$ is the density at time t, then

$$H(t) = \int f(t,x)\log f(t,x)\Pi\phi(x_i)dx \le H(0) \le CN$$

The entropy inequality states

$$\int Fd\lambda \le H(\lambda|\mu) + \log\big[\int e^F d\mu]$$

Proof:

$$\sup_x [xy - x\log x + x] = ye^y - ye^y + e^y = e^y$$

Therefore

$$\int F f d\mu \leq \int e^F d\mu + \int [f \log f - f] d\mu$$

Since $\int f d\mu = 1$, we can replace F by $F + c$ to get

$$\int F f d\mu \leq \int e^{F+c} d\mu - c - 1 + \int f \log f d\mu = e^c \int e^F d\mu - c - 1 + H(\lambda|\mu)$$

Minimize with respect to c. $c = -\log \int e^F d\mu$. Provides control. In particular taking $F = k\mathbf{1}_A(x)$,

$$k\lambda(A) \leq H(\lambda|\mu) + \log[e^k \mu(A) + (1 - \mu(A)]$$

with $k = \log \frac{1}{\mu(A)}$, we get

$$\lambda(A) \leq \frac{2 + H(\lambda|\mu)}{\log \frac{1}{\mu(A)}}$$

Convexity of $Q(f) = D(\sqrt{f})$ is an immediate consequence of the variational formula

$$Q(f) = -\inf_{u>0} \int \frac{(\mathcal{L}u)(x)}{u(x)} f(x) d\mu$$