Introduction to Hypoelliptic Schrödinger type operators

Lectures by

Martin Hairer

Notes taken by

Piotr Ługiewicz

August 31, 2007

Contents

1 **Introduction and Motivation** 5

2 **The Strategy of our Analysis** 11

3 **Some Remarks on Compactness** 15

4 **Hörmander's Theorem** 19

5 **Hörmander-like Estimates in Non-Compact Domains** 27

1 Introduction and Motivation

This series of lectures is devoted to the study of the long-time behaviour of semigroups generated by (the appropriate closure of) operators given informally by the expression:

$$Lf(x) = \sum a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum b_i(x) \frac{\partial f(x)}{\partial x_i} + V(x) f(x) \;, \tag{1.1}$$

where the functions $a_{ij}(x)$, $b_j(x)$ and $V(x)$ are smooth and the operator L is well defined initially on the set of functions f which are smooth and compactly supported. Any operator of this form will be called a Schrödinger-type operator. The reason for this comes from physics: in quantum mechanics, the Schrödinger operator H is the operator

$$Hf(x) = -\sum \frac{\partial^2 f(x)}{\partial x_i^2} + V(x) f(x).$$

This is of the form (1.1) with $b = 0$ and $a_{ij}(x) = -\delta_{ij}$ for each x, i.e. the symmetric matrix function $a_{ij}(x)$ is uniformly (strictly) negative definite, so that L is a self-adjoint elliptic operator. Relaxing this condition we may lose both ellipticity and self-adjointness, but expect that our operator (1.1) will still have nice properties which enables us to do an interesting analysis. Without any additional assumption, the class of operators L specified in this way is too large to be interesting. Throughout these lectures, we will always deal with hypoelliptic operators (instead of elliptic ones) and we will make further structural ssumptions further on.

Information about the long-time behaviour of the underlying diffusion processes can be captured by the detailed analysis of the generator L. More precisely, we are going to investigate the resolvent of L and deduce from its properties relevant features of the spectrum $\sigma(L)$.

It is worth mentioning some physical context before we start our discussion of the main issues in the next sections. Operators (1.1) appear in the study of non-equilibrium statistical physics. As a simplified example, consider first the system with one particle and the law of motion governed by the Hamiltonian function H_S given by

$$H_S(q,p) = \frac{p^2}{2} + V(q) \;. \tag{1.2}$$

The states of the particle are described by points (q, p) in Euclidean space. The function $V(q)$ describes the potential (external fields acting on the particle). We now wish to introduce a 'heat bath' at some temperature T with which our particle interacts. This will be done in the spirit of [FKM65, FK87]. We follow very closely the exposition of [EPR99b].

We introduce the temperature in two steps. First we consider an infinite reservoir which is deterministic (so no statistical aspects are involved for a while). The law of motion for the reservoir is given by the Hamiltonian function

$$H_B(\varphi) = \frac{1}{2}\|\varphi\|^2_{\mathcal{H}} ,$$

where we assumed that the state φ of the reservoir takes values in some appropriate Hilbert space $\mathcal{H}$. The equation of motion of the free reservoir can in general be given by

$$\frac{d}{dt}\varphi = \mathcal{L}\varphi ,$$

for some linear operator $\mathcal{L} : \mathcal{H} \to \mathcal{H}$. Since one has the identity

$$\frac{d}{dt}H_B(\varphi) = \langle \mathcal{L}\varphi, \varphi\rangle + \langle \varphi, \mathcal{L}\varphi\rangle = \langle \varphi, (\mathcal{L} + \mathcal{L}^*)\varphi\rangle ,$$

we will require $\mathcal{L}^* = -\mathcal{L}$, so that the energy of isolated reservoir is conserved. To be more specific one can think about φ as an electro-magnetic field

$$\varphi = \begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix} \quad : \ \mathbb{R}^3 \longrightarrow \mathbb{R}^6$$

in a vacuum, with the additional constraints $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{B} = 0$. The Hilbert space $\mathcal{H}$ is given by the scalar product

$$\|\varphi\|^2_{\mathcal{H}} = \int \left(\mathbf{E}^2 + \mathbf{B}^2\right) dx$$

and the generator of the evolution is given by Maxwell's equations:

$$\mathcal{L}\begin{pmatrix} \mathbf{E} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} \nabla \wedge \mathbf{E} \\ \nabla \wedge \mathbf{B} \end{pmatrix} .$$

Finally, we add some interaction between the particle and the heat reservoir. We assume that the interaction is of dipole form, so that the total Hamiltonian of the system is given by:

$$H(q, p, \varphi) = \frac{p^2}{2} + V(q) + \frac{1}{2}\|\varphi + q\alpha\|^2_{\mathcal{H}} .$$

Here $q\alpha \in \mathcal{H}$ can be interpreted as the field which describes the contribution coming from the charged particle.

Having done this, one can write down the equations for the total system (particle + reservoir) as follows:

$$\begin{cases} \frac{d}{dt} q = p \\ \\ \frac{d}{dt} p = -\nabla V(q) - \langle \varphi + q\alpha, \alpha \rangle \\ \\ \frac{d}{dt} \varphi = \mathcal{L}(\varphi + q\alpha) \end{cases} \tag{1.3}$$

We want somehow to "forget" about the big system and get an effective law of motion for our particle. This can be realized by solving the last equation in (1.3)

$$\varphi(t) = e^{t\mathcal{L}}\varphi(0) + \int_0^t q(s)\ \mathcal{L}e^{(t-s)\mathcal{L}}\alpha\ ds.$$

We integrate by parts to get

$$\varphi(t) = e^{t\mathcal{L}}\varphi(0) + q(0)e^{t\mathcal{L}}\alpha - q(t)\alpha + \int_0^t p(s)\ e^{(t-s)\mathcal{L}}\alpha\ ds.$$

We plug this expression into the second equation in (1.3) to get

$$\frac{d}{dt}p = -\nabla V(q) - \langle \varphi(0), e^{-t\mathcal{L}}\alpha \rangle - q(0)\langle e^{t\mathcal{L}}\alpha, \alpha \rangle - \int_0^t p(s)\langle e^{(t-s)\mathcal{L}}\alpha, \alpha \rangle ds \tag{1.4}$$

By a simple change of co-ordinates, we can and do assume $q(0) = 0$. Then we have

$$\frac{d}{dt}p = -\nabla V(q) - \langle \varphi(0), e^{-t\mathcal{L}}\alpha \rangle - \int_0^t p(s)\langle e^{(t-s)\mathcal{L}}\alpha, \alpha \rangle ds \tag{1.5}$$

Up to now, everything has been deterministic, as the equation of motion were simply derived by the use of Newton's and Maxwell's laws. Now, we are in a position to put the reservoir at temperature T - the second announced step. What we do is simple: we just postulate that the deterministic field φ is replaced by a random field $\phi(0)$ with initial probability distribution describing the statistical behaviour of the physical reservoir at the temperature T. More precisely, we assume that the probability distribution of the initial data $\phi(0)$ is Gaussian with the covariance given by

$$\mathbb{E}\langle \phi(0), f \rangle_{\mathcal{H}} \cdot \langle \phi(0), g \rangle_{\mathcal{H}} = T\langle f, g \rangle_{\mathcal{H}}. \tag{1.6}$$

The measure corresponding to (1.6) describes the distribution of the physical ensemble (the canonical Gibbs distribution) and has density given informally by the expression:

$$\text{“}\ \frac{e^{-\frac{\|\varphi\|^2_{\mathcal{H}}}{2T}}}{Z}\ \text{”}$$

It is natural in view of (1.5) to introduce a centred Gaussian process ξ defined by

$$\xi(t) = \langle \phi(0), e^{-t\mathcal{L}}\alpha\rangle_{\mathcal{H}} \,.$$

The covariance of ξ can be computed easily from (1.6):

$$\mathbb{E}\xi(s)\xi(t) = T\langle e^{-t\mathcal{L}}\alpha, e^{-s\mathcal{L}}\alpha\rangle_{\mathcal{H}} = T\langle \alpha, e^{(t-s)\mathcal{L}}\alpha\rangle_{\mathcal{H}} \,,$$

which shows that ξ is a stationary stochastic process.
Using this notation, the system of equations (1.3) can be rewritten as:

$$\begin{cases} \frac{d}{dt}q = p \\ \\ \frac{d}{dt}p = -\nabla V(q) - \xi(t) - \int_0^t p(s)C(t-s)ds \\ \\ \text{with} \quad C(t-s) = \frac{1}{T}\mathbb{E}\xi(t)\xi(s) \end{cases} \tag{1.7}$$

Here, the fact that the covariance of ξ also appears in the memory kernel of the ‘friction term’ is the celebrated fluctuation - dissipation theorem.

The last simplification we consider is

$$C(t) \sim 2\gamma\delta(t) \tag{1.8}$$

where $\delta(t)$ - is the Dirac “delta function” and a constant $\gamma > 0$. Thus ξ corresponds just to the white noise.

One can argue that (1.8) is a reasonable assumption since the covariance $C(t)$ corresponds to the scalar product of particle fields taken at time 0 and t where the initial field of the particle is subjected to the free dynamics; hence the Dirac delta function is a good approximation to $C(t)$ as soon as the field generated by the particle at time 0 is strongly localized in space directions, so that the time it takes for this field to radiate away from the particle is small compared to any other timescale of the system.

Summarizing we have

$$\begin{cases} \frac{d}{dt}q = p \\ \\ \frac{d}{dt}p = -\nabla V(q) + \xi(t) - \gamma p(t) \\ \\ \text{with} \quad C(t-s) = \frac{1}{T}\mathbb{E}\xi(t)\xi(s) = \frac{2\gamma}{T}\delta(t-s) \end{cases} \tag{1.9}$$

It is convenient to introduce properly re-scaled process $\xi \longrightarrow \sqrt{2\gamma T}\xi$. After this modification the equation (1.9) reads

$$\begin{cases} \frac{d}{dt}q = p \\ \\ \frac{d}{dt}p = -\nabla V(q) - \gamma p(t) + \sqrt{2\gamma T}\xi(t) \\ \\ \text{with} \quad \mathbb{E}\xi(t)\xi(s) = \delta(t-s) \end{cases} \tag{1.10}$$

i.e. $\xi(t)$ is the standard white noise. Note that the equation of motion (1.10) is *stochastic* even though it was derived from a *deterministic* model with random initial conditions. The fact that 'new randomness' gets injected into (1.10) for all times is a consequence of the fact that we assumed C to decay (quickly) for large times, which implies that the free fields sends disturbances out to infinity. The counterpart of this statement is that 'fresh randomness' can also come in from infinity.

Finally, we remark that the physical model described above can easily be generalized to two (or more) heat baths at different temperatures $T_{_L}$, $T_{_R}$ and moreover one can consider arbitrary many (but finite) number of particles instead of one. Let us consider a very simple model of heat conduction for which the Hamiltonian is given by

$$H \equiv H(q,p) = \sum_i \left(\frac{p_i^2}{2} + V_1(q_i) + V_2(q_{i+1} - q_i) \right) .$$

The potential V_1 is a 'pinning term' that keeps the particles around their equilibrium position, whereas the potential V_2 is an 'interaction term' which is responsible for the transmission of energy / information along the chain. The corresponding system of stochastic equations is as follows (in the case of two heat baths)

$$\begin{cases} \frac{d}{dt}q_i = p_i \qquad for\ i \in \{1,..,N\} \\ \\ \frac{d}{dt}p_0 = -\nabla_0 V_1(q_0) - \nabla_0 V_2(q_1 - q_0) - \gamma p_0(t) + \sqrt{2\gamma T_{_L}}\xi_0(t) \\ \\ \frac{d}{dt}p_i = -\nabla_i V_1(q_i) - \nabla_i V_2(q_{i+1} - q_i) - \nabla_i V_2(q_i - q_{i-1}) \qquad for\ i \in \{1,..,N-1\} \\ \\ \frac{d}{dt}p_N = -\nabla_N V_1(q_N) - \nabla_N V_2(q_N - q_{N-1}) - \gamma p_N(t) + \sqrt{2\gamma T_{_R}}\xi_N(t) \\ \\ \text{with} \quad \mathbb{E}\xi_k(t)\xi_l(s) = \delta_{kl}\delta(t-s) \qquad\qquad k,l = 0,N \end{cases} \tag{1.11}$$

The only directions, in which the ellipticity appears, are p_0 and p_N. In the remaining directions the dynamic is only deterministic hence we have a "big"

non-ellipticity. The system of stochastic equations (1.11) provides us with an example of the operator L of the type (1.1) which is hypoelliptic.
One of the surprising features of this type of operators is that their spectral properties are extremely sensitive to the fine structure of their first-order part. In particular, we will see that if we choose

$$\begin{cases} V_1(q) = q^2 \\ V_2(q) = q^4 \end{cases}$$

it is possible to show that the generator L of (1.11) has compact resolvent in a suitably chosen Hilbert space. This will allow us to conclude the existence of a unique stationary distribution for (1.11) and establish the exponential convergence to it. On the other hand, if one were to choose

$$\begin{cases} V_1(q) = q^4 \\ V_2(q) = q^2 \end{cases}$$

it is still an open problem to show the *existence* of a stationary solution. In particular, it is possible to show that if the chain is long enough, the generator L does *not* have compact resolvent in any of the natural spaces associated with it. Intuitively, the reason behind this fact is that at high energies, the interaction between neighbouring particles becomes weaker and weaker, so that the chain appears more and more 'broken'.

2 The Strategy of our Analysis

Let $\mathcal{P}_t$ be the Markov semi-group generated by the system of stochastic differential equations (1.11). In other words, for any bounded and measurable function

$$\psi : \quad \mathbb{R}^{2(N+1)} \longrightarrow \mathbb{R}$$

we define

$$(\mathcal{P}_t\psi)(x) = \mathbb{E}\left[\psi(x(t))\,;\; x(0) = x\right] \tag{2.1}$$

where $\mathbb{E}$ is the mean value with respect to the Wiener measure on the sample space of Brownian motion (in $\mathbb{R}^{2(N+1)}$).

The family of operators $\{\mathcal{P}_t\}$ satisfies the following list of properties which are easy to verify

$$\begin{cases} \mathcal{P}_t \circ \mathcal{P}_s\psi &= \mathcal{P}_{t+s}\psi \qquad \text{Markov property} \\ \partial_t\mathcal{P}_t\psi &= \mathcal{L}\mathcal{P}_t\psi \\ \mathcal{L} &= X_H - \gamma p_0\frac{\partial}{\partial p_0} + \gamma T_{{}_L}\frac{\partial^2}{\partial p_0^2} - \gamma p_N\frac{\partial}{\partial p_N} + \gamma T_{{}_R}\frac{\partial^2}{\partial p_N^2} \\ X_H H &= 0 \quad \text{(conservation of energy)} \\ X_H &= \sum_i \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i} \quad \text{(Liouville operator)} \end{cases}$$

Let us introduce now the concept of an *invariant measure.* A probability measure μ on $\mathbb{R}^{2(N+1)}$ is *invariant* with respect to semi-group $\mathcal{P}_t$ iff

$$\int (\mathcal{P}_t\psi)(x)\mu(dx) = \int \psi(x)\mu(dx) \tag{2.2}$$

holds true for any $t > 0$ and any ψ as described in (2.1) above. We also write this as $\mu\mathcal{P}_t = \mu$ for any $t > 0$.

Assume that $\mu(dx) = \rho(x)dx$, then the condition (2.2) reads $\partial_t\langle\rho, \mathcal{P}_t\psi\rangle = 0$ where $\langle\ ,\ \rangle$ is the scalar product with respect to Lebesgue measure. Therefore

$$\langle\mathcal{L}^*\rho, \mathcal{P}_t\psi\rangle = \langle\rho, \mathcal{L}\mathcal{P}_t\psi\rangle = 0$$

for all $t > 0$ and all ψ as above. This means that

$$\mathcal{L}^*\rho = 0\ .$$

The question of the existence of an invariant measure now becomes: can we find ρ such that the condition above is satisfied? This leads us naturally to the question of the behaviour of the spectrum of L, i.e. does 0 belong to the spectrum? But even if 0 belongs to the spectrum, this doesn't guarantee eigenfunctions to exist which correspond to probability densities. Our strategy is then as follows:

(1) find a Hilbert space $\mathcal{H}$ such that all bounded (measurable) functions ψ belong to $\mathcal{H}$;

(2) Show that $\mathcal{P}_t$ can be extended to $\mathcal{C}_0$-semi-group of bounded operators on $\mathcal{H}$;

(3) Show that $\mathcal{H}$ is such that 0 is an isolated eigenvalue with finite multiplicity for the generator $\mathcal{L}$ of the extension of $\mathcal{P}_t$ to $\mathcal{H}$.

This strategy does have a good chance of being useful since we have the following:

Proposition
If there exists a Hilbert space $\mathcal{H}$ such that **1**-**3** *hold true, then there exists a $\mathcal{P}_t$-invariant measure μ.*

Proof:
Let $\mathcal{L}$ denote the generator of the extension of $\mathcal{P}_t$ to $\mathcal{H}$, and denote by $\mathcal{L}^*$ the adjoint of $\mathcal{L}$ in the Hilbert space $\mathcal{H}$. Since, by assumption, 0 is an isolated eigenvalue with finite multiplicity for $\mathcal{L}$, it is also so for $\mathcal{L}^*$.
Let now $g \in \mathcal{H}$ be such that $\mathcal{L}^* g = 0$. By Riesz's representation theorem, g can be considered as a linear bounded functional on $\mathcal{H}$. Since we assumed that $\mathcal{H}$ contained all measurable bounded functions, this shows that g can be identified with a signed measure with bounded variation $\|g\|_{Var} < +\infty$. We want to show that one can find a nonnegative g such that $\langle g, \mathcal{P}_t\psi\rangle = \langle g, \psi\rangle$.
Let us recall Jordan's decomposition theorem for signed measures. *Any signed measure with bounded variation can be written uniquely as*

$$\mu = \mu_+ - \mu_- \ ,$$

where the positive measures μ_- and μ_+ are mutually singular.
We also use the following.
Fact: *If $\mu = \tilde{\mu}_+ - \tilde{\mu}_-$, with $\tilde{\mu}_+$ and $\tilde{\mu}_-$ positive (not necessarily mutually singular), then there exists a nonnegative measure δ such that: $\mu_+ = \tilde{\mu}_+ - \delta$ and $\mu_- = \tilde{\mu}_- - \delta$.*
After these preparations we come back to the proof of proposition. Writing $g = g_+ - g_-$, since $\mathcal{L}^* g = 0$ we have

$$\mathcal{P}_t^* g = \mathcal{P}_t^* g_+ - \mathcal{P}_t^* g_- = g$$

where $\mathcal{P}_t^* g_+$, $\mathcal{P}_t^* g_-$ correspond to non-negative measures since $\mathcal{P}_t^*$ is Markov. In this way we get a new decomposition of the density g and we can apply the fact above to infer that there exists a non-negative measure δ such that

$$g_+ = \mathcal{P}_t^* g_+ - \delta \quad \text{and} \quad g_- = \mathcal{P}_t^* g_- - \delta \; . \tag{2.3}$$

Using again the fact that $\mathcal{P}_t$ is a Markov operator, we have

$$\mathbf{mass}\,(\mathcal{P}_t^* g_+) = \mathbf{mass}\,(g_+) \; .$$

If we now insert this in (2.3), we get

$$\mathbf{mass}\,(g_+) = \mathbf{mass}\,(g_+) - \mathbf{mass}(\delta) \; ,$$

which implies $\delta = 0$, so that $\mathcal{P}_t^* g_\pm = g_\pm$. Since $g \neq 0$, one of the two positive measures g_- and g_+ must be non-vanishing, so that we found the requested invariant measure. ∎

3 Some Remarks on Compactness

As we remember from the previous section, the structure of our generator is as follows
$$\mathcal{L} = X_H - \gamma p_0 \frac{\partial}{\partial p_0} + \gamma T_{_L} \frac{\partial^2}{\partial p_0^2} - \gamma p_{_N} \frac{\partial}{\partial p_{_N}} + \gamma T_{_R} \frac{\partial^2}{\partial p_{_N}^2} \ ,$$
where X_H is the first order operator which corresponds to deterministic Hamiltonian dynamics, the second and third terms together represent the generator of the Ornstein-Uhlenbeck process at the temperature $T_{_L}$, and the forth and fifth term represent the generator of the Ornstein-Uhlenbeck process at the temperature $T_{_R}$.

How shall we choose the auxiliary Hilbert space $\mathcal{H}$ in a proper way? Again, physical intuition is helpful. Observe that if $T_{_L} = T_{_R} = T$, the invariant measure for our system is given by the Gibbs measure

$$\mu(dqdp) \sim e^{-\frac{1}{T}H(q,p)} dqdp$$

This measure is also invariant with respect to Hamiltonian evolution as the energy is conserved, i.e. we also get $X_H \mu = 0$.

For different temperatures $T_{_R} \neq T_{_L}$ it is not possible to guess an explicit formula for the invariant measure. A natural Ansatz for the Hilbert space $\mathcal{H}$ is the following family of weighted spaces

$$\mathcal{H}_\beta = L^2 \left(\mathbb{R}^{2(N+1)};\ e^{-\beta H} dqdp \right)$$

for some $\beta \equiv \frac{1}{T}$ which will be chosen later.

We will deal with adjoint operators, so to make the calculations simpler we use a unitary equivalence between $\mathcal{H}_\beta$ and the 'flat' space $L^2(dx)$. Namely, we introduce the unitary transformation $\mathcal{K} : L^2(dx) \longrightarrow \mathcal{H}_\beta$ given by the formula

$$\left(\mathcal{K}g\right)(x) = \frac{1}{\sqrt{Z}} e^{\frac{\beta}{2}H} g(x)$$

and consider $L = \mathcal{K}^{-1}\mathcal{L}\mathcal{K}$, the image of our generator in the "flat" space $L^2(dx)$. This transformation does not change the overall form of the operator $\mathcal{L}$, but it slightly modifies the 'Ornstein-Uhlenbeck' parts and adds a zero-order term. We get

$$L = X_H - \alpha_0 p_0 \frac{\partial}{\partial p_0} + \overline{\alpha}_0 \frac{\partial^2}{\partial p_0^2} - \alpha_{_N} p_{_N} \frac{\partial}{\partial p_{_N}} + \overline{\alpha}_{_N} \frac{\partial^2}{\partial p_{_N}^2} - c_0 p_0^2 - c_{_N} p_{_N}^2 + c \ ,$$

for some appropriate constants α_i, $\bar{\alpha}_i$ and c_i. To secure the right bound for norms of our semi-group P_t, we would like to have

$$\mathrm{Re}\langle\varphi, L\varphi\rangle \leq C\|\varphi\|^2 \qquad \text{for any} \quad \varphi \in D(L) \equiv \text{domain of} \quad L \tag{3.1}$$

with some constant $C \in (0, \infty)$ independent of the function φ. The condition (3.1) means that the constants c_0 and $c_{_N}$ are (strictly) positive which turns out to be equivalent to the following bound we need to impose on β:

$$\beta < 2 \min\left\{\frac{1}{T_{_L}}, \frac{1}{T_{_R}}\right\}$$

Reminder: Recall the following points

- L is m-accretive iff it satisfies (3.1) and there is no extension of L satisfying (3.1);
- The resolvent is defined by $R_\lambda = (L - \lambda)^{-1}$ for all $\lambda \notin \sigma(L) \equiv$ *spectrum of* L;
- A bounded operator $A : \mathcal{H} \longrightarrow \mathcal{H}$ is compact iff the set $\{Af \ : \ \|f\| \leq 1\}$ is pre-compact (i.e. its closure is compact);
- If we can show that R_λ is compact for one λ then it is compact for all λ from the resolvent set;
- Also, R_λ being compact implies that its spectrum consists of countably many eigenvalues of finite multiplicities accumulating only at 0.
- This implies that eigenvalues of L are discrete and can only accumulate at infinity
- In particular, if we know that the resolvent of L is compact, 0 is an isolated eigenvalue which coincides with condition **(3)** of our general strategy.

Proposition
Assume that the resolvent set of L is non-empty. Then L has compact resolvent iff the set

$$\{f \ : \ \|Lf\| + \|f\| \leq 1\} \tag{3.2}$$

is pre-compact.

Proof: Assume without loss of generality that $0 \notin \sigma(L)$ (otherwise we consider $L - \lambda$ instead of L). We want to show that the condition (3.2) is equivalent to

$$\left\{L^{-1}f \ : \ \|f\| \leq 1\right\} \qquad \text{is pre-compact}$$

But the set above is the same as the set

$$\{g \;:\; \|Lg\| \leq 1\}$$

and this is equivalent to the re-scaled set (the scaling sends compact sets into compact sets)

$$\{f \;:\; \|Lf\| + \|f\| \leq 1\} \;.$$

Indeed, if $\|Lg\| \leq 1$, then $\|Lg\| + \|g\| \leq \|Lg\| + \|L^{-1}Lg\| \leq C\|Lg\| \leq C$. On the other hand, if $\|Lg\| + \|g\| \leq 1$, then of course $\|Lg\| \leq 1$. ∎

Corollary: *If we know that Λ has compact resolvent and that there exists a constant c such that*

$$\|\Lambda f\| \leq c\,(\|Lf\| + \|f\|)$$

for every $f \in \mathcal{D}(L)$, then L has compact resolvent, provided that its resolvent set is not empty.

This gives us a recipe for showing that L has a compact resolvent by comparing it to an operator Λ which can be better understood. Let us now discuss what it actually means for a subset of a Hilbert space to be compact. The following result is well known:

Arzela-Ascoli Theorem

If X is a compact space and $\Sigma \subset \mathcal{C}(X)$. Then Σ is precompact iff it is equi-bounded and equi-continuous.

Corollary: *Let $\Sigma \subset \mathcal{C}(\mathbb{R}^n)$. Suppose there exists $V : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ such that*

$$\lim_{x\to\infty} V(x) = 0$$

and for any $g \in \Sigma$

$$|g(x)| < |V(x)|$$

If Σ is equi-continuous on every compact set, then Σ is precompact.

The moral of the Arzela-Ascoli Theorem is that in order to obtain compactness, one needs both *confinement* and *regularity*. In view of this, the following criterion for compactness in L^2 should not come as a surprise:

Rellich's criterion:

Take $V,\ W\ : \mathbb{R}^n \longrightarrow [1,\infty)$ such that

$$\begin{cases} \lim_{|x|\to\infty} V(x) & = +\infty \qquad \text{(decay)} \\ \lim_{|x|\to\infty} W(x) & = +\infty \qquad \text{(regularity)} \end{cases}$$

Then the set

$$\left\{ f \ : \ \|Vf\| \le 1 \quad \text{and} \quad \|W\hat{f}\| \le 1 \right\}$$

is pre-compact in $L^2(\mathbb{R}^n)$*. Here, we denoted by* $\hat{f}$ *the Fourier transform of* f*.*

Corollary:
Consider $\Lambda = -\Delta + V(x)$ *with* V *continuous and* $V(x) \to +\infty$ *as* $|x| \to +\infty$*. Then* Λ *has compact resolvent.*

The following example shows that although the above growth condition is sufficient, it is not necessary to obtain compactness of the resolvent.
Example:
Consider the Schrödinger operator

$$\Lambda = -\partial_x^2 - \partial_y^2 + x^2y^2 \ .$$

The potential $V(x,y) = x^2y^2$ doesn't grow to infinity in all directions but it is nevertheless possible to show that Λ has compact resolvent. To see this we can rewrite the operator Λ in the following way

$$\Lambda = -\frac{1}{2}\Delta + \left(-\frac{1}{2}\partial_x^2 + \frac{y^2}{2}x^2\right) + \left(-\frac{1}{2}\partial_y^2 + \frac{x^2}{2}y^2\right) \ge -\frac{1}{2}\Delta + |y| + |x| \equiv A$$

where the last inequality is understood in the quadratic forms sense. (*Think of the two operators in the brackets as one dimensional quantum harmonic oscillators for which the bottom of spectrum is given by the two expressions on the right hand side of the inequality.*)

This shows that one has the bound

$$\|A^{1/2}f\| = \sqrt{\langle f, Af\rangle} \le \sqrt{\langle f, \Lambda f\rangle} \le \frac{1}{2}\left(\|\Lambda f\| + \|f\|\right) \ .$$

By Rellich's criterion, A (and therefore also $A^{1/2}$) has compact resolvent, and hence Λ has compact resolvent.

4 Hörmander's Theorem

There are two main problems with the operator L that we wish to analyse. First, it is not elliptic, so that it is not obvious *a priori* how we can obtain the regularity estimates. Second, the 'potential' $p_0^2 + p_N^2$ does not tend to infinity in all directions, so that it is not obvious how to obtain the necessary confinement bounds. In this section, we show how to address the first problem.

First we introduce the appropriate notion of regularity. Let $L: \mathcal{S}' \longrightarrow \mathcal{S}'$ be a linear operator defined on the space of tempered distributions $\mathcal{S}'$. We say that L is *hypoelliptic* if, whenever $Lf = g$ and $g \in C^\infty$, then f is in C^∞.

To become more familiar with this notion we study two examples:

1. $L_1 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2}$,

2. $L_2 f(x,y) = \frac{\partial f(x,y)}{\partial y} + \frac{\partial^2 f(x,y)}{\partial x^2}$.

Both examples are similar. The property they have in common is that they have second order derivatives only in one of the two spatial directions. The first example is not a hypoelliptic operator: if $f(x,y)$ is an arbitrary measurable function depending only on the y-coordinate, then $Lf = 0$, which is clearly a smooth function, even though f itself is not smooth. In the case of L_2 however, we can interpret the y-coordinate as a 'time', so that the equation $L_2 f = g$ can be viewed as a heat equation with a source term. We know from the theory of parabolic PDEs that the solutions to such an equation are smooth, provided that the source term is smooth.

In general, it turns out that the first example is, up to change of coordinates, pretty close to being the only way in which a second-order differential operator can fail to be hypoelliptic. The precise answer is given by the Hörmander Theorem which we are going to study in this section.

Before we state this Theorem let us look at an intuitive geometrical picture which is behind the Frobenius Theorem described below. Let M be a smooth finite dimensional manifold and let E be a smooth sub-bundle $M \ni p \to E_p \subset T_pM$ of the tangent bundle TM. We say that E is *integrable* iff for any two vectors fields X_1 and X_2 such that $X_i\big|_p \in E_p$, $i = 1,2$ one has $[X_1, X_2]\big|_p \in E_p$

for any $p \in M$. Here, we denote by $[X_1, X_2]$ the *Lie bracket* of the two vector fields X_1 and X_2.

Frobenius Theorem:
E is integrable iff it arises from tangent sub-bundles generated by a (regular) family of submanifolds.

Keeping in mind this geometrical picture we see that the following criterion of hypoellipticity due to Hörmander is "almost" iff:

Hörmander's Theorem:
Let M be a smooth finite-dimensional manifold. Suppose that at each point $x \in M$ the tangent space T_xM can be spanned by the fields $X_0, X_1, \dots, X_n$ together with their Lie brackets, i.e.

$$T_xM = Span\left\{X_j\big|_x, [X_j, X_k]\big|_x, [X_j, [X_k, X_l]]\big|_x, ...\right\} \qquad \textbf{(Hö)}$$

with indices $j, k, l, ... \in \{0, 1, ..., n\}$. Then the operator

$$L = X_0 + \sum_{i=1}^{n} X_i^2$$

is hypoelliptic.

In the sequel, an operator L which satisfies the condition described in the Theorem above will be called a *Hörmander operator*. Note that we do not impose any restriction on n; in particular one can give non-trivial examples of Hörmander operators with $n = 2$ on $\mathbb{R}^d$ for arbitrary d.

The main subject of this section is to give an idea of the proof that Hörmander operators are hypoelliptic. To get started with the proof of this statement we need some preliminaries: we introduce *pseudo-differential operators*. To make the discussion less technical at this point one can forget about general manifolds and just think about $\mathbb{R}^d$ since the statement of Hörmander's Theorem has a local character.

Given a smooth function $\mathbf{a} : \mathbb{R}^{2d} \longrightarrow \mathbb{C}$ we define a *pseudo-differential operator* $Op\mathbf{a} : \mathcal{S} \longrightarrow \mathcal{S}$ by

$$(Op\mathbf{a}\ u)(x) = \frac{1}{(2\pi)^d} \int e^{\mathbf{i}\langle x, \xi\rangle} \mathbf{a}(x, \xi) \hat{u}(\xi) d\xi$$

The function $\mathbf{a}(x,\xi)$ is called the *symbol* of the operator $Op\mathbf{a}$.

We say that $\mathbf{a} \in \mathbf{S}^m$ (or that $\mathbf{a}$ is of order m) with $m \in \mathbb{R}$ if $\forall \alpha, \beta \in \mathbb{N} \cap \{0\}$

$$\left| \partial_x^\alpha \partial_\xi^\beta \mathbf{a}(x,\xi) \right| \leq C_{\alpha,\beta} \left(1 + \|\xi\|\right)^{m-|\beta|} \ .$$

The classes of symbols have the following properties:

- If $\mathbf{a} \in \mathbf{S}^m$ and $\mathbf{b} \in \mathbf{S}^n$, then the symbol of $Op\mathbf{a} \cdot Op\mathbf{b}$, (which is just the composition of operators), belongs to $\mathbf{S}^{m+n}$ and

$$Op\mathbf{a} \cdot Op\mathbf{b} = Op\mathbf{ab} + Op\mathbf{c} \ ,$$

  with $\mathbf{c} \in \mathbf{S}^{m+n-1}$. In particular the commutator $[Op\mathbf{a}, Op\mathbf{b}]$ has a symbol in the space $\mathbf{S}^{m+n-1}$, i.e. is of one order less.

- If $\mathbf{a} \in \mathbf{S}^0$, then $Op\mathbf{a}$ is a bounded operator in L^2.

- Finally, we have the following rule for the adjoint operator. If $\mathbf{a} \in \mathbf{S}^m$, then

$$(Op\mathbf{a})^* = Op\overline{\mathbf{a}} + Op\mathbf{c} \ ,$$

  where $\mathbf{c} \in \mathbf{S}^{m-1}$.

In the considerations below we use $\Lambda = (1-\Delta)^{\frac{1}{2}}$. Obviously, the symbol of Λ^α belongs to $\mathbf{S}^\alpha$ for every $\alpha \in \mathbb{R}$.

Main Step in the Proof of Hörmander Theorem:
For every compact set K, there exists $\varepsilon > 0$ and a constant $c > 0$ such that

$$\|\Lambda^\varepsilon u\| \leq c \left(\|Lu\| + \|u\|\right) \qquad (*)$$

for every $u \in C_0^\infty(K)$

Overall strategy: *Suppose that we can find smooth vector fields $\{Y_j(x)\}$ such that*

1° *span* $\{Y_j(x)\} = \mathbb{R}^d$ *for every $x \in K$ and*

2° *there exists $0 < \varepsilon < 1$ and $c > 0$ such that $\forall u \in C_0^\infty(K)$ and $\forall j$*

$$\|\Lambda^{\varepsilon-1} Y_j u\| \leq c \left(\|Lu\| + \|u\|\right) \ .$$

Then the inequality $(*)$ *holds.*

Note that since we only care about what happens in the compact set K, we can modify L outside K to have all of its co-efficient and all their derivatives bounded for all x whenever we need this. We assume from now one that we have performed this modification, so that L has a symbol in $\mathbf{S}^2$.

Proof :
Before we start the proof we introduce the following convention:
After the symbol $\preceq$ *we omit all expressions of the type* $\lambda\left(\|Lu\|^2 + \|u\|^2\right)$ *with some positive constant* λ *independent of* u.

By $1°$, there exists a function a_{ij} assuming values in the set of matrices, which is smooth and such that

$$\Lambda^2 = \sum_{i,j} a_{ij}(x) Y_i Y_j + 1.$$

Given this, with some constant $C > 0$, one has

$$\|\Lambda^\varepsilon u\|^2 = \langle u, \Lambda^{2\varepsilon-2}\Lambda^2 u\rangle \leq C\ \|u\|^2 + \sum \langle u, \Lambda^{2\varepsilon-2} a_{ij} Y_i Y_j u\rangle$$

In order to use $2°$, we intend to put Y_i in the front of the first u in the second term of the expression above to bound it by

$$\preceq \sum -\langle Y_i u, \Lambda^{2\varepsilon-2} a_{ij} Y_j u\rangle + \langle u, T Y_j u\rangle$$

with commutator $T \equiv \left[\Lambda^{2\varepsilon-2} a_{ij}, Y_i\right] \in Op\mathbf{S}^{2\varepsilon-2}$. The last expression can be rewritten as follows

$$= -\sum \langle \Lambda^{\varepsilon-1} Y_i u, \left(\Lambda^{\varepsilon-1} a_{ij} \Lambda^{1-\varepsilon}\right) \Lambda^{\varepsilon-1} Y_j u\rangle + \langle u, \left(T\Lambda^{1-\varepsilon}\right) \Lambda^{\varepsilon-1} Y_j u\rangle \preceq 0$$

The operators $\left(\Lambda^{\varepsilon-1} a_{ij} \Lambda^{1-\varepsilon}\right)$ and $\left(T\Lambda^{1-\varepsilon}\right)$ are bounded (to check this use the rules for pseudo-differential operators). Thus we need to bound the expression $\|\Lambda^{1-\varepsilon} Y_i u\|$ which we can do by using $(2°)$. From that the inequality $(*)$ follows.
■

To prove the Hörmander's Theorem we need only to prove $2°$ and we do this below by induction. The idea will be to chose the fields as in the condition (**Hö**).

Proof of $(2°)$ **:**
The proof is divided into several steps. Since by the Hörmander condition (**Hö**), the first part of the family of spanning fields Y_j is given by the basic fields

X_0, X_i, $i = 1, .., n$, in the first two steps we prove 2° for them. The remaining steps provide inductive machinery which allows us to prove the desired condition for a field which is given as a commutator of a basic field and a field for which 2° is known (possibly at a cost of taking smaller $\varepsilon > 0$).

STEP 1. Note that $L = \sum X_i^* X_i + \tilde{X}_0$ with $\tilde{X}_0 \equiv \sum_i f_i X_i + X_0$ for some smooth functions f_i; (the antisymmetric part of $\tilde{X}_0$ really does not matter while the symmetric part of the operator $\tilde{X}_0$ corresponds to the multiplication by a bounded function and hence the contribution coming from this part can be bounded by $c\|u\|$). We therefore have for $i \neq 0$:

$$\|X_i u\|^2 = \langle u, X_i^* X_i u\rangle \leq \sum_i \langle u, X_i^* X_i u\rangle \leq \mathrm{Re}\langle u, Lu\rangle + c\|u\|^2$$

with some constant $c > 0$, and using Cauchy-Schwartz inequality, we get

$$\|X_i u\| \leq C\left(\|Lu\| + \|u\|\right)$$

with some constant $C > 0$.

STEP 2. For $\varepsilon \in (0, \frac{1}{2}]$, using $X_0 = L - \sum X_i^2$, one has

$$\|\Lambda^{\varepsilon-1} X_0 u\|^2 = \langle \Lambda^{2\varepsilon-2} X_0 u, \left(L - \sum X_i^2\right) u\rangle \leq$$

$$\leq c\|u\| \cdot \|Lu\| - \sum \langle \Lambda^{2\varepsilon-2} X_0 u, X_i^2 u\rangle \preceq ...$$

(in the first term, to bound the norm of $\Lambda^{2\varepsilon-2} X_0 u$ we used $\varepsilon \leq \frac{1}{2}$),

$$... \preceq -\sum \langle \Lambda^{2\varepsilon-2} X_0 X_i u, X_i u\rangle + |\langle Tu, X_i u\rangle|$$

with $\Lambda^{2\varepsilon-2} X_0$ and the commutator $T \equiv \left[X_i^*, \Lambda^{2\varepsilon-2} X_0\right] \in Op\mathbf{S}^{2\varepsilon-1}$, and hence bounded for $\varepsilon \leq \frac{1}{2}$. We conclude using estimates from Step 1 (to bound $\|X_j u\|$ and $\|X_j^* u\|$).

STEP 3. Suppose that, for some $\varepsilon < \frac{1}{2}$, one has $\|\Lambda^{2\varepsilon-1} Y u\|^2 \preceq 0$. We show that this implies that $\|\Lambda^{\varepsilon-1}[X_i, Y]\|^2 \preceq 0$ for $i = 1, ..., n$ ($i \neq 0$). Given i, we set $Z \equiv [X_i, Y]$. Then

$$\|\Lambda^{\varepsilon-1} Z u\|^2 = \langle \Lambda^{2\varepsilon-2} Z u, X_i Y u\rangle - \langle \Lambda^{2\varepsilon-2} Z u, Y X_i u\rangle$$

The first term in this equation can be rewritten as

$$\begin{aligned}
\langle \Lambda^{2\varepsilon-2} Z u, X_i Y u\rangle &= \langle X_i^* \Lambda^{2\varepsilon-2} Z u, Y u\rangle = \\
&= \langle \Lambda^{2\varepsilon-2} Z X_i^* u, Y u\rangle + \langle \left[X_i^*, \Lambda^{2\varepsilon-2} Z\right] u, Y u\rangle =
\end{aligned}$$

$$= \langle \Lambda^{-1} Z X_i^* u, \Lambda^{2\varepsilon-1} Y u\rangle + \langle \Lambda^{-2\varepsilon+1}\left[X_i^*, \Lambda^{2\varepsilon-2} Z\right] u, \Lambda^{2\varepsilon-1} Y u\rangle \; .$$

Setting $T = \Lambda^{-2\varepsilon+1}\left[X_i^*, \Lambda^{2\varepsilon-2} Z\right]$ and performing similar manipulations for the second term, we get

$$\|\Lambda^{\varepsilon-1} Z u\|^2 = \langle \Lambda^{-1} Z X_i^* u, \Lambda^{2\varepsilon-1} Y u\rangle + \langle T u, \Lambda^{2\varepsilon-1} Y u\rangle \qquad (4.1)$$
$$- \; \langle \left(\Lambda^{2\varepsilon-2} Z \Lambda^{1-2\varepsilon}\right)\left(\Lambda^{2\varepsilon-1} Y\right) u, X_i u\rangle + \langle T' u, X_i u\rangle$$

with $\Lambda^{-1} Z \in Op\mathbf{S}^0$ and appropriate commutators $T \in Op\mathbf{S}^0$ and $T' \in Op\mathbf{S}^{2\varepsilon-1}$. Using Cauchy-Schwartz and using the bounds on the terms $\|X_i^* u\|$ and $\|X_i u\|$ obtained in Step 1, we get the required bound.

STEP 4. Commuting with X_0 is more difficult. Assume that there exists $\varepsilon \in (0, \frac{1}{4})$ such that $\|\Lambda^{4\varepsilon-1} Y u\|^2 \preceq 0$. We will show that such a bound implies that $\|\Lambda^{\varepsilon-1}[X_0, Y] u\|^2 \preceq 0$. We remark that after inequality $\preceq$ we can drop the adjoint applied to a vector field, if the expression we get after dropping the field is itself $\preceq 0$. We shall use this property to make the notation simpler (in the previous steps we were more explicit and we didn't use this property). Setting $Z \equiv [X_0, Y]$, we have as before

$$\|\Lambda^{\varepsilon-1} Z u\|^2 = \langle \Lambda^{2\varepsilon-2} Z u, Y X_0 u\rangle - \langle \Lambda^{2\varepsilon-2} Z u, X_0 Y u\rangle. \qquad (4.2)$$

Again as in the Step 3, we focus our attention on the first term using the results of the Step 2. Writing $X_0 = L - \sum X_i^2$ as in Step 2, we have

$$\langle \Lambda^{2\varepsilon-2} Z u, Y X_0 u\rangle = -\sum_i \langle \Lambda^{2\varepsilon-2} Z u, Y X_i^2 u\rangle + \langle \Lambda^{2\varepsilon-2} Z u, Y L u\rangle \qquad (4.3)$$

First we show how to manage the 'easy' term, usig the remark above

$$|\langle \Lambda^{2\varepsilon-2} Z u, Y L u\rangle| \preceq |\langle Y \Lambda^{2\varepsilon-2} Z u, L u\rangle| =$$
$$= \langle \left(\Lambda^{2\varepsilon-2} Z \Lambda^{-4\varepsilon+1}\right) \Lambda^{4\varepsilon-1} Y u, L u\rangle + \langle \left[Y, \Lambda^{2\varepsilon-2} Z\right] u, L u\rangle \preceq 0.$$

The operator $\left(\Lambda^{2\varepsilon-2} Z \Lambda^{-4\varepsilon+1}\right)$ in the first term is bounded and the commutator $\left[Y, \Lambda^{2\varepsilon-2} Z\right]$ in the second term is also bounded. Hence we can use the Cauchy-Schwartz inequality and the inductive assumption $\|\Lambda^{4\varepsilon-1} Y u\|^2 \preceq 0$ to conclude estimates of this term.

Below we will treat all "L"-terms in the similar manner without mentioning this. From the bound we just obtained, the first term in (4.2) can be estimated as follows

$$\langle \Lambda^{2\varepsilon-2} Z u, Y X_0 u\rangle \preceq -\sum_i \langle \Lambda^{2\varepsilon-2} Z u, Y X_i^2 u\rangle \; .$$

We have here the square of the fields X_i and that means that we will have more difficulties than in Step 3. We estimate each term of the sum above separately:

$$-\langle \Lambda^{2\varepsilon-2} Z u, Y X_i^2 u\rangle = -\langle X_i^* Y^* \Lambda^{2\varepsilon-2} Z u, X_i u\rangle \preceq \|X_i^* Y^* \Lambda^{2\varepsilon-2} Z u\| \cdot \|X_i u\|$$

We estimate the term containing the square of $X_i^* Y^*$ as follows.

$$\begin{aligned} \|X_i^* Y^* \Lambda^{2\varepsilon-2} Z u\|^2 &\preceq \|X_i Y^* \Lambda^{2\varepsilon-2} Z u\|^2 + C\|Y^* \Lambda^{2\varepsilon-2} Z u\|^2 \\ &\preceq \|X_i Y^* \Lambda^{2\varepsilon-2} Z u\|^2 + C\|Y \Lambda^{2\varepsilon-2} Z u\|^2 \end{aligned} \tag{4.4}$$

with some constant $C > 0$. The estimate of the second term goes as follows

$$\|Y \Lambda^{2\varepsilon-2} Z u\|^2 \leq \| \left[\Lambda^{2\varepsilon-2} Z \Lambda^{-4\varepsilon+1}\right] \Lambda^{4\varepsilon-1} Y u\|^2 + \| \left[Y, \Lambda^{2\varepsilon-2} Z\right] u\|^2 \preceq 0$$

where we have used our current inductive assumption $\|\Lambda^{4\varepsilon-1} Y u\|^2 \preceq 0$. To estimate the first term in (4.4) it is sufficient to estimate $\|X_i Y \Lambda^{2\varepsilon-2} Z u\|$, (since the difference is of lower order). We have

$$\|X_i Y \Lambda^{2\varepsilon-2} Z u\|^2 \preceq \mathrm{Re}\langle L Y \Lambda^{2\varepsilon-2} Z u, Y \Lambda^{2\varepsilon-2} Z u\rangle$$

where we pass from X_i to L, (see the similar reasoning in Step 1). Hence

$$\begin{aligned} &\preceq \mathrm{Re}\langle \left(\Lambda^{-2\varepsilon} Y \Lambda^{2\varepsilon-2} Z\right)\ L u, \left(\Lambda^{4\varepsilon-2} Z \Lambda^{1-4\varepsilon}\right) \Lambda^{4\varepsilon-1} Y u\rangle + \\ &+\ \text{terms containing commutators with } \ X_0 \ \text{ and } \ X_i^2 \end{aligned}$$

The operators $\left(\Lambda^{-2\varepsilon} Y \Lambda^{2\varepsilon-2} Z\right)$ and $\left(\Lambda^{4\varepsilon-2} Z \Lambda^{1-4\varepsilon}\right)$ are bounded and we can use our inductive assumption to get that for the first term $\preceq 0$. Finally we need to consider the commutator terms. We consider only commutators with X_i^2, (the analysis for X_0 is simpler as X_0 is a first order operator). We have

$$\begin{aligned} \langle [X_i^2, Y \Lambda^{2\varepsilon-2} Z] u, \Lambda^{2\varepsilon-2} Z Y u\rangle = &\ 2\langle [X_i, Y \Lambda^{2\varepsilon-2} Z] X_i u, \Lambda^{2\varepsilon-2} Z Y u\rangle \\ &+\ \langle [X_i, [X_i, Y \Lambda^{2\varepsilon-2} Z]] u, \Lambda^{2\varepsilon-2} Z Y u\rangle\ . \end{aligned} \tag{4.5}$$

The first term on the r.h.s. of (4.5) is ‘good’, as we have

$$\langle \Lambda^{-4\varepsilon+1} Z^* \Lambda^{2\varepsilon-2}\ [X_i, Y \Lambda^{2\varepsilon-2} Z] X_i u, \Lambda^{4\varepsilon-1} Y u\rangle \preceq 0\ .$$

The last term (4.5) can be represented as follows

$$\langle \Lambda^{-4\varepsilon+1} Z^* \Lambda^{2\varepsilon-2}\ [X_i, [X_i, Y \Lambda^{2\varepsilon-2} Z]] u, \Lambda^{4\varepsilon-1} Y u\rangle \preceq 0\ .$$

This completes the proof of 2°. ∎

FINAL STEP IN THE PROOF OF HÖRMANDER'S THEOREM
We just showed that one has the bound

$$\|\Lambda^{\varepsilon} u\| \leq c\,(\|Lu\| + \|u\|) \ .$$

If we make the substitution $u \longrightarrow \Lambda^{\alpha} u$, $\alpha \in \mathbb{R}$ and use similar arguments to bound the commutator $[\Lambda^{\alpha}, L]$, it is possible to get bounds of the type

$$\|\Lambda^{\varepsilon+\alpha} u\| \leq C\,(\|\Lambda^{\alpha} Lu\| + \|\Lambda^{\alpha} u\|)$$

for every $\alpha > 0$. It is at this stage possible to use a bootstrapping argument to get smoothness of u from the smoothness of Lu. That is, if we know that $Lu = g$ with a smooth function g, starting from some negative K for which $\|\Lambda^{K} u\| < \infty$, we can use the above inequality inductively to get that for any $m \in \mathbb{N}$, $\|\Lambda^{m} u\| < \infty$. Thus u must be smooth and this ends the proof of Hörmander Theorem.

■

The key references for the above route to Hörmander's theorem include [Koh78] and [OR73]. See also the recent book [HN05] for a clear exposition of this and related problems.

In our model one can verify that the property $V_2'' \neq 0$ implies that the Hörmander condition holds. Intuitively, just because the chain is connected, taking commutators gives derivatives in all directions $q_1, ..., q_n, p_1, ..., p_n$; (think what happens when the chain is broken).

5 Hörmander-like Estimates in Non-Compact Domains

The purpose of this section is the study of Hörmander type estimates, but in non-compact regions (actually the whole Euclidean space). The Hörmander Theorem give us bounds of the type

$$\|\Lambda^\varepsilon u\| \le C_K\left(\|Lu\| + \|u\|\right)$$

for each $u \in C^\infty(K)$. It is reasonable to expect that if the coefficients of L are nicely behaved, the constant C_K can grow at most polynomially in the size of the region K. Let us denote by $\{Y_k\}_{k\ge 0}$ the collection of all the vector fields $X_0, \{X_j\}$ making up the differential operator L, together with all their Lie brackets. The following assumption is natural in view of obtaining control over the growth of C_K:

Assumption: *The polynomial Hörmander condition.*

All coefficients of the vector fields X_i (and their derivatives) are bounded by multiples of some polynomial $(1+|x|)^N$. Furthermore, there exists $r > 0$ and a constant C such that

$$\sum_{j\le r}\langle Y_j(x),\xi\rangle^2 \ge \frac{C}{1+|x|^N}|\xi|^2\ ,$$

for every pair $x, \xi \in \mathbb{R}^d$. In short we will refer to this condition as *the poly - Hörmander condition.*

One can show that *if L satisfies the poly-Hörmander condition, then there exist positive constants C, ε and N such that*

$$\|\Lambda^\varepsilon u\| \le C(1+|x|^N)\left(\|Lu\| + \|u\|\right) \tag{5.1}$$

for each $u \in C_0^\infty(\mathbf{B}(x,1))$, where $\mathbf{B}(x,1)$ is a unit ball centred at the point x. We do not enter into a detailed proof of this fact, but it can be proven essentially by retracing the argument of the previous section and keeping track of constants.

For further study we need the following Sobolev spaces. We say that $u \in \mathcal{S}^{\alpha,\beta}$ iff

$\|u\|_{\alpha,\beta} \equiv \|\overline{\Lambda}^{\beta}\Lambda^{\alpha}u\| < +\infty$ where Λ is as before and $\overline{\Lambda}(x) \equiv \left(1+|x|^2\right)^{\frac{1}{2}} \equiv \langle x\rangle$. The most important property of these spaces is that $\mathcal{S}^{\alpha',\beta'} \subset \mathcal{S}^{\alpha,\beta}$ for $\alpha' \geq \alpha$, $\beta' \geq \beta$ and that this embedding is compact if and only if both inequalities are strict. In particular, $\Lambda^{\alpha}\circ\overline{\Lambda}^{\beta}$ has compact resolvent in all of these spaces, as soon as $\alpha > 0$ and $\beta > 0$.

Proposition
If L satisfies the poly-Hörmander condition, then for each $\varepsilon > 0$ there exists $\delta > 0$ and a constant $C \in (0,\infty)$ such that

$$\|u\|_{\delta,\delta} \leq C\left(\|Lu\| + \|u\|_{0,\varepsilon}\right) ,$$

for each $u \in C_0^{\infty}(\mathbb{R}^n)$

Proof:
We divide the whole space $\mathbb{R}^d$ into cubes centred around $\mathbb{Z}^d$ and we consider a partition of unity $\{\phi_x\}_{x\in\mathbb{Z}^d}$. In this way, we can write an arbitrary smooth function u as

$$u = \sum_{x\in\mathbb{Z}^d}\varphi_x u = \sum_{x\in\mathbb{Z}^d}\tilde{u}_x , \tag{5.2}$$

where $\tilde{u}_x \in C_0^{\infty}(\mathbf{B}(x,1))$. The crucial step is to prove the inequality

$$\|\tilde{u}\|_{\delta,\delta} \leq c\left(\|L\tilde{u}\| + \|\tilde{u}\|\right) + C\langle x\rangle^{\varepsilon}\|\tilde{u}\| \tag{5.3}$$

for all $\tilde{u} \in C_0^{\infty}\left(\mathbf{B}(x,1)\right)$. Given this inequality, it is relatively easy exercise to conclude the proof of the proposition.

We can show (5.3) starting with (5.1), knowing that roughly $\|\tilde{u}\|_{\delta,\delta} \approx \langle x\rangle^{\delta}\|\tilde{u}\|_{\delta,0}$. Using the Jensen inequality (applied to the first two factors) one can write

$$\langle x\rangle^{N+\delta}\|\Lambda^{\delta}\tilde{u}\| = \langle x\rangle^{N+\delta}\frac{\|\Lambda^{\delta}\tilde{u}\|}{\|\tilde{u}\|}\|\tilde{u}\| \leq C\frac{\|\Lambda^{\delta}\tilde{u}\|^J}{\|\tilde{u}\|^{J-1}} + \tilde{c}\langle x\rangle^{(N+\delta)(1+\frac{1}{J-1})}\|\tilde{u}\|.$$

On the other hand for any self-adjoint positive operator A and any $J > 0$ one has

$$\|A\tilde{u}\|^J \leq \|A^J\tilde{u}\|\|\tilde{u}\|^{J-1}.$$

Inserting this into our estimate, we have

$$\begin{aligned}\langle x\rangle^{N+\delta}\|\Lambda^{\delta}\tilde{u}\| &\leq \|\Lambda^{\delta J}\tilde{u}\| + \tilde{c}\langle x\rangle^{(N+\delta)(1+\frac{1}{J-1})}\|\tilde{u}\| \\ &\leq C\left\{\left(\|L\tilde{u}\| + \|\tilde{u}\|\right) + \tilde{c}\langle x\rangle^{(N+\delta)(1+\frac{1}{J-1})}\|\tilde{u}\|\right\}\end{aligned}$$

For any fixed J, we can choose δ small enough such that $J\delta$ is smaller than the ε from (5.1). This leads to

$$\langle x\rangle^{\delta}\|\Lambda^{\delta}\tilde{u}\| \le C\left(\|L\tilde{u}\| + \|\tilde{u}\|\right) + C\langle x\rangle^{\delta\left(1+\frac{1}{J-1}\right)+\frac{N}{J-1}}\|\tilde{u}\| \; .$$

Since J can be made arbitrarily large (at the expense of making δ very small), the exponent $\delta\left(1+\frac{1}{J-1}\right)+\frac{N}{J-1}$ can be made arbitrarily small and hence (5.3) follows. ∎

Now we go back to our physical model of a chain of oscillators coupled to two heat baths at different temperatures. Since the above results apply (under mild growth conditions on the potentials V_1 and V_2), it is sufficient in order to show compactness of the resolvent of L, to obtain a bound of the type

$$\|u\|_{0,\varepsilon} \le C(\|Lu\| + \|u\|) \; . \tag{5.4}$$

We should note that the operator $\overline{\Lambda}$ is not a good choice - the technique just does not work. A more convenient choice is the operator of multiplication by the Hamiltonian H of the system, instead of the previous $\overline{\Lambda}$. The useful property is that the energy operator H commutes with Liouville operator which makes the analysis possible, i.e. $[X_H, H] = 0$. We are therefore going to argue that for some choices of the pinning and the coupling potentials, one can obtain a bound of the type

$$\|H^{\varepsilon}u\| \le C(\|Lu\| + \|u\|) \; , \tag{5.5}$$

which then implies (5.4) because H grows at least polynomially in all directions.

We introduce the spaces

$$\mathcal{F}_{\alpha}^{k} = \left\{ f \;:\; |D^i f| \le C_i\langle x\rangle^{\alpha-\min\{|i|,k\}} \right\} .$$

The main result is as follows

Theorem

If $V_1 \in \mathcal{F}_{\alpha}^{2}$ and $V_2 \in \mathcal{F}_{\beta}^{2}$ with $\beta > \alpha > 2$ and moreover $V_1 \ge c\langle x\rangle^{\alpha}$, $V_2 \ge c\langle x\rangle^{\beta}$, $xV_1' \ge c\langle x\rangle^{\alpha} - \tilde{c}$, $xV_2' \ge c\langle x\rangle^{\beta} - \tilde{c}$ and finally $(V_2'')^{-1} \in \mathcal{F}_0$, then the inequality (5.4) holds true.

Some Comments

- If the conditions $V_1 \geq c\langle x\rangle^\alpha$, $V_2 \geq c\langle x\rangle^\beta$ are not satisfied then the pinning potential dominates and there is no sufficient transport of energy and the bound (5.4) does not hold in general.
- The condition $(V_2'')^{-1} \in \mathcal{F}_0$ is not necessary.
- The condition $\beta > \alpha > 2$ can be relaxed and it is sufficient to assume $\beta \geq \alpha \geq 2$.

Idea of the Proof:
Our generator has the form

$$L = X_H + \mathcal{L}_{OU} + cp_0^2 + cp_N^2$$

(where $\mathcal{L}_{OU}$ is an Ornstein-Uhlenbeck operator). As before in the Step 1, it is straightforward to get the bound

$$\|p_0 u\| \leq c\left(\|Lu\| + \|u\|\right) .$$

We would like to mimic Step 4 of the proof of the Hörmander theorem to get bounds on

$$\langle H^{\varepsilon-1}\left[X_0, f\right], gu\rangle ,$$

where g is some function with well-controlled growth (use a priori bounds), and f is a function on which we obtained bounds from a previous step (induction). Starting with $f = p_0$ and $g = q_1 - q_0$, we get

$$\begin{aligned}\langle H^{\varepsilon-1}\nabla V_2(q_1 - q_0)u, (q_1 - q_0)u\rangle &= \langle H^{\varepsilon-1}\left[X_0, p_0\right]u, (q_1 - q_0)u\rangle + \\ &+ \langle H^{\varepsilon-1}\nabla V_1(q_0)u, (q_1 - q_0)u\rangle ,\end{aligned} \tag{5.6}$$

where

$$H = \sum \frac{p_i^2}{2} + \sum V_1(q_i) + \sum V_2(q_i - q_{i-1}) .$$

The term involving X_0 can be bounded in pretty much the same way as before. Concerning the other term, if we would like to bound

$$\langle H^{\varepsilon-1}\nabla V_1(q_0)u, q_1 u\rangle$$

and

$$\langle H^{\varepsilon-1}\nabla V_1(q_0)u, q_0 u\rangle$$

separately, we would need

$$H^{\varepsilon-1}\nabla V_1(q_0)q_0$$

bounded for some ε. This is not realistic since, if V_1 behaves like a polynomial at infinity, $\nabla V_1(q_0)q_0 \sim V_1(q_0)$ at infinity and at the same time $H \sim V_1(q_0)$ in the direction q_0.

On the other hand, we have

$$\nabla V_1(q_0)(q_1 - q_0) \leq \frac{|q_1 - q_0|^p}{p} + \frac{|\nabla V_1(q_0)|^{p'}}{p'}$$

with $\frac{1}{p} + \frac{1}{p'} = 1$. Now, if we can choose p and p' in such way that

$$\frac{|q_1 - q_0|^p}{p} \ll V_2(q_1 - q_0) \tag{5.7}$$

and

$$\frac{|\nabla V_1(q_0)|^{p'}}{p'} \ll V_1(q_0) \tag{5.8}$$

are both satisfied, then we get a bound on the second term in (5.6).

Supposing that $V_1(q) \sim q^n$ and $V_2(q) \sim q^m$, then using (5.7) and (5.8), we have to take

$$p < m \qquad \text{and} \qquad p' < \frac{n}{n-1}$$

i.e.

$$\frac{1}{p} > \frac{1}{m} \qquad \text{and} \qquad \frac{1}{p'} > 1 - \frac{1}{n}.$$

Hence

$$\frac{1}{p} + \frac{1}{p'} > 1 + \frac{1}{m} - \frac{1}{n}$$

which implies $n < m$. ∎

Bibliography

[EH00] J.-P. ECKMANN and M. HAIRER. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. Comm. Math. Phys. 212, no. 1, (2000), 105–164.

[EH03] J.-P. ECKMANN and M. HAIRER. Spectral properties of hypoelliptic operators. Comm. Math. Phys. 235, no. 2, (2003), 233–253.

[EPR99a] J.-P. ECKMANN, C.-A. PILLET, and L. REY-BELLET. Entropy production in nonlinear, thermally driven Hamiltonian systems. J. Statist. Phys. 95, no. 1-2, (1999), 305–331.

[EPR99b] J.-P. ECKMANN, C.-A. PILLET, and L. REY-BELLET. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. Comm. Math. Phys. 201, no. 3, (1999), 657–697.

[FK87] G. W. FORD and M. KAC. On the quantum Langevin equation. J. Statist. Phys. 46, no. 5-6, (1987), 803–810.

[FKM65] G. W. FORD, M. KAC, and P. MAZUR. Statistical mechanics of assemblies of coupled oscillators. J. Mathematical Phys. 6, (1965), 504–515.

[HM88] B. HELFFER and A. MOHAMED. Characterisation du spectre essentiel de loperateur de Schrödinger avec un champ magnetique. Ann. Inst. Fourier (Grenoble) 38, no. 2, (1988), 951–12.

[HN85] B. HELFFER and J. NOURRIGAT. Hypoellipticite maximale pour des operateurs polynomes de champs de vecteurs, Progress in Mathematics vol. 58 , Birkhauser Boston Inc., Boston, MA, 1985.

[HN04] F. HERAU and F. NIER. Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. Arch. Ration. Mech. Anal. 171, no. 2, (2004), 151–218.

[HN05] B. HELFFER and F. NIER. Hypoelliptic estimates and spectral theory for Fokker- Planck operators and Witten Laplacians, Lecture Notes in Mathematics vol. 1862, Springer-Verlag, Berlin, 2005.

[Hor67] L. HÖRMANDER. Hypoelliptic second order differential equations. Acta Math. 119, (1967), 147–171.

[Hor85] L. HÖRMANDER. The Analysis of Linear Partial Differential Operators I–IV. Springer, New York, 1985.

[Koh69] J. J. KOHN. Pseudo-differential operators and non-elliptic problems. In Pseudo-Diff. Operators (C.I.M.E., Stresa, 1968), 157–165. Edizioni Cremonese, Rome, 1969.

[Koh78] J. J. KOHN. Lectures on degenerate elliptic problems. In Pseudodifferential operator with applications (Bressanone, 1977), 89151. Liguori, Naples, 1978.

[OR73] O. A. OLEJNIK & E. V. RADKEVIC, Second order equations with non-negative characteristic form, Providence Amer. Math. Soc. (1973).

[RBT02] L. REY-BELLET and L. E. THOMAS. Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. Comm. Math. Phys. 225, no. 2, (2002), 305–329.

[RS80] M. REED and B. SIMON. Methods of Modern Mathematical Physics IIV. Academic Press, San Diego, California, 1980.

[Vil06] C. VILLANI. Hypocoercivity, 2006. Unpublished preprint.

[Yos80] K. YOSIDA. Functional Analysis. Springer, New York, 6th ed., 1980.