Chapter 1

From Stochastic Process to Quantum Mechanics: A Sketch

1.1. How to Read this Note?

Tips for reading this note:

- **Definitions** are in bold font.
- *Important* statements are in italic font.
- Only important equations are numbered.
- Questions are in red color.
- Conclusions are in green color.
- You can omit all the nonsense about stochastic process, just focusing on quantum mechanics.
- Footnotes are also negligible (except for those containing **definition**s).

1.2. Superposition Principle and Time Evolution

We mimic the deriving of stochastic process. Unlike stochastic process, where the system is described by probability, a quantum system is represented by wave-function. Given a set $\mathcal{X}$ that contains all possible states of the quantum system, a **wave-function** is a map $f: \mathcal{X} \to \mathbb{C}$. Especially when $\mathcal{X}$ is a manifold, we suppose that f is smooth, that is $f \in C^\infty(\mathcal{X})$. In addition, when $\mathcal{X}$ is Euclidean, namely $\mathcal{X} = \mathbb{R}^d$, we further suppose that f is an element of Schwartz space $\mathcal{S}(\mathbb{R}^d)$.[1.1]

First consider the analogy of master equation in stochastic process. The key feature of master equation is its linearity. The first axiom of quantum mechanics, superposition principle, claims the same linearity.

Axiom 1.1. *[Superposition Principle] Physical laws that operate on quantum states shall be linear.*

An implication of superposition principle is how quantum states (precisely, their wave-functions) evolve with time. To give raise to time evolution, we have to add time-dependence to wave-function, which makes a wave-function $f: \mathcal{X} \times \mathbb{R} \to \mathbb{C}$, where $\mathbb{R}$ characterizes the time-axis. Then, axiom 1.1 claims that the equation of time evolution (as a physical law that operates on a quantum state) shall be linear: $\partial f / \partial t = L(f)$ where the operation L is linear.[1.2] Mathematically, linearity imports a kernel $r: \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ such that (if $\mathcal{X}$ is discrete, then the integral is defined to be $\int_{\mathcal{X}} \mathrm{d}x \to \sum_{x \in \mathcal{X}}$)

$$\mathrm{i}\frac{\partial f}{\partial t}(y,t) = \int_{\mathcal{X}} \mathrm{d}x\, f(x,t) r(x,y). \tag{1.1}$$

The right hand side can be seen as a generalization of linear transformation in $\mathbb{R}^n$ like $\sum_{i=1}^n f_i r_{ij}$. The imaginary i is employed for convenience. You can compare equation 1.1 with the master equation of continuous time Markovian process. The only difference is the imaginary number in the front.

1.1. A **Schwartz space** $\mathcal{S}(\mathbb{R}^d)$ contains smooth function $f: \mathbb{R}^d \to \mathbb{C}$ that decays "exponentially fast" at infinity. Precisely, for any m-order polynomial P_m and any m-order partial derivative D^n, with integers $m, n \geqslant 0$, we have $|P_m(x) D^n f(x)| \to 0$ as $\|x\| \to \infty$. Functions in Schwartz space are usually termed by **Schwartz function**s.

1.2. Why not $\partial^2 f / \partial t^2 = L(f)$? This may give rise to another axiom that we will know the whole history and future of a wave-function if we know it (or its norm, namely the distribution of particles) at any time t. In short, the evolutionary equation of wave-function is first order on time.

1.3. Probability Interpretation Implies Hermitianity

Then, the probability interpretation add restriction to the transition rate. In stochastic process, the restriction is about the normalization of transition density. But in quantum mechanics, it is about unitary.

AXIOM 1.2. *[Probability Interpretation (Part 1)] Given a wave-function f of quantum system, the probabilistic density that the system at state x is given by $|f(x)|^2 = f^*(x)f(x)$.*

Since probabilistic density shall be normalized, axiom 1.2 implies that, for any wave-function f,

$$\int_{\mathcal{X}} \mathrm{d}x\, f^*(x) f(x) = 1. \tag{1.2}$$

Since this equation holds for all wave-functions, time evolution sustains it too. Namely, for a time-dependent wave-function $f(x,t)$,

$$\int_{\mathcal{X}} \mathrm{d}x\, f^*(x,t) f(x,t) = \int_{\mathcal{X}} \mathrm{d}x\, f^*(x,t') f(x,t') = 1$$

holds for any $t, t' \in \mathbb{R}$. We are to show what this results in the transition rate r (which determines the time evolution). Equation 1.2 gives

$$f(y, t+\Delta t) = f(y,t) - \mathrm{i}\Delta t \int \mathrm{d}x\, f(x,t)\, r(x,y) + o(\Delta t).$$

Plugging into equation 1.2 for wave-function $f(\cdot, t+\Delta t)$ gives

$$\int_{\mathcal{X}} \mathrm{d}y \left[f^*(y,t) + \mathrm{i}\Delta t \int_{\mathcal{X}} \mathrm{d}x\, f^*(x,t)\, r^*(x,y) \right] \left[f(y,t) - \mathrm{i}\Delta t \int_{\mathcal{X}} \mathrm{d}x'\, f(x',t)\, r(x',y) \right] = 1.$$

By inserting equation 1.2 for wave-function $f(\cdot,t)$ and taking derivative on Δt, it implies

$$\int_{\mathcal{X}^2} \mathrm{d}x\, \mathrm{d}y\, f^*(x,t) f(y,t) \left[r^*(x,y) - r(y,x) \right] = 0$$

holds for any f in $\mathcal{S}(\mathbb{R}^d)$, thus

$$r^*(x,y) = r(y,x). \tag{1.3}$$

Namely, *r is Hermitian*.[1.3]

1.4. In Euclidean Space: Path Integral

In this section and the followings, we focus on the situation where $\mathcal{X} = \mathbb{R}^d$, on which Fourier transform can be performed.[1.4] Thus, given t, we have wave-function $f(\cdot,t) \in \mathcal{S}(\mathbb{R}^d)$.

We are trying to derive a generic path integral formalism. Given a small $\Delta t > 0$, equation 1.1 gives

$$f(x', t+\Delta t) = \int_{\mathbb{R}^d} \mathrm{d}x\, f(x,t) [\delta(x'-x) - \mathrm{i}\, r(x,x')\Delta t] + o(\Delta t)$$

We are to convert the $[\cdots]$ part into exponential. To do so, we use the inverse Fourier transformations

$$\delta(x'-x) = \int_{\mathbb{R}^d} \frac{\mathrm{d}k}{(2\pi)^d} \exp(\mathrm{i}k_\alpha (x'^\alpha - x^\alpha)),$$

and

$$r(x,x') = \int_{\mathbb{R}^d} \frac{\mathrm{d}k}{(2\pi)^d} \exp(\mathrm{i}k_\alpha (x'^\alpha - x^\alpha))\, \hat{r}(x,k),$$

1.3. A function $f: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is **Hermitian** if $f^*(x,y) = f(y,x)$ for each $x, y \in \mathbb{R}^d$. In the traditional approach of quantum mechanics, we have $r(x,y) = \langle y|\hat{H}|x\rangle$, where $\hat{H}$ is the Hamiltonian operator.

1.4. This is critical for our analysis. But I think this can be generalized to any manifold on which the Green's function of Laplacian equation can be solved.

in which

$$\hat{r}(x,k) := \int_{\mathbb{R}^d} \mathrm{d}\epsilon \exp(-\mathrm{i}k_\alpha \epsilon^\alpha)\, r(x, x+\epsilon). \tag{1.4}$$

Then, the $[\cdots]$ part is converted into exponential by

$$\begin{aligned}[\cdots] &= \int_{\mathbb{R}^d} \frac{\mathrm{d}k}{(2\pi)^d} \exp(\mathrm{i}k_\alpha(x'^\alpha - x^\alpha))[1 - \mathrm{i}\hat{r}(x,k)\Delta t] \\ &= \int_{\mathbb{R}^d} \frac{\mathrm{d}k}{(2\pi)^d} \exp(\mathrm{i}k_\alpha(x'^\alpha - x^\alpha) - \mathrm{i}\hat{r}(x,k)\Delta t) + o(\Delta t)\end{aligned}$$

Plugging back to $f(x', t+\Delta t)$ results in

$$f(x', t+\Delta t) = \int_{\mathbb{R}^d} \mathrm{d}x \int_{\mathbb{R}^d} \frac{\mathrm{d}k}{(2\pi)^d} f(x,t) \exp(\mathrm{i}k_\alpha(x'^\alpha - x^\alpha) - \mathrm{i}\hat{r}(x,k)\Delta t) + o(\Delta t)$$

Re-denoting $x_0 := x$, $x_1 := x'$, $k^0 := k$ (for "balancing" indices, we put the "temporal" index of k as superscript, thus the α-component of the k at the 0-th "time-slice" is written by k^0_α),

$$f(x_1, t+\Delta t) = \int_{\mathbb{R}^d} \mathrm{d}x_0 \int_{\mathbb{R}^d} \frac{\mathrm{d}k^0}{(2\pi)^d} \exp(\mathrm{i}k^0_\alpha(x_1^\alpha - x_0^\alpha) - \mathrm{i}\hat{r}(x_0, k^0)\Delta t) f(x_0, t) + o(\Delta t)$$

Repeating this process N times, we arrive at[1.5]

$$f(x_N, t+N\Delta t) = \int D(x,k)\, f(x_0, t) \exp(\mathrm{i}S(x,k)) + o(N\Delta t), \tag{1.5}$$

where the integral

$$\int D(x,k) := \int_{\mathbb{R}^d} \mathrm{d}x_0 \int_{\mathbb{R}^d} \frac{\mathrm{d}k^0}{(2\pi)^d} \cdots \int_{\mathbb{R}^d} \mathrm{d}x_{N-1} \int_{\mathbb{R}^d} \frac{\mathrm{d}k^{N-1}}{(2\pi)^d} \tag{1.6}$$

and

$$S(x,k) := \sum_{i=0}^{N-1} \Delta t \left[\left(\frac{x^\alpha_{i+1} - x^\alpha_i}{\Delta t} \right) k^i_\alpha - \hat{r}(x_i, k^i) \right]. \tag{1.7}$$

If we recognize $(x_{i+1} - x_i)/\Delta t$ as the velocity $\dot{x}_i$, then the $S(x,k)$ can be seen as the Legendre transformation $\dot{x}\,p - H(x,p)$, in which *k is proportional to momentum p and $\hat{r}(x,k)$ plays the role of Hamiltonian $H(x,p)$*.

1.5. An Useful Expansion

In this section, we claim an mathematical theorem that is critical for continuing the journey. It extends a function in Schwartz space (ensuring Fourier transformation), to a generalized function. With this extension, the function can be expanded by a series of generalized functions.

Define the n-th order **moment** (with $n \geqslant 0$) of a function $M_n : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ by

$$M_n^{\alpha_1 \cdots \alpha_n}(f) := \int_{\mathbb{R}^d} \mathrm{d}x\, f(x)\, (x^{\alpha_1} \cdots x^{\alpha_n}). \tag{1.8}$$

Since $f \in \mathcal{S}(\mathbb{R}^d)$, the moment $M_n(f)$ is finite for any n. Denote $\mathcal{S}_A$ the subspace of Schwartz space in which every function is analytic, termed as **analytic Schwartz space**. Then, for an arbitrary function $\varphi \in \mathcal{S}_A(\mathbb{R}^d)$, Taylor expanding at origin gives

$$\begin{aligned}\int_{\mathbb{R}^d} \mathrm{d}x\, f(x)\, \varphi(x) &= \sum_{n=0}^{+\infty} \frac{1}{n!} \left[\int_{\mathbb{R}^d} \mathrm{d}x\, f(x)\, (x^{\alpha_1} \cdots x^{\alpha_n}) \right] (\partial_{\alpha_1} \cdots \partial_{\alpha_n} \varphi)(0) \\ &= \sum_{n=0}^{+\infty} \frac{1}{n!} M_n^{\alpha_1 \cdots \alpha_n}(f)\, (\partial_{\alpha_1} \cdots \partial_{\alpha_n} \varphi)(0).\end{aligned}$$

1.5. We have to show that the residue is an $o(N\Delta t)$, but this is far from trivial.

On the other hand, because of the identity

$$(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\varphi)(0)=\int_{\mathbb{R}^d}\mathrm{d}x\,\delta(x)\,(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\varphi)(x),$$

integration by parts on the right hand side gives

$$(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\varphi)(0)=(-1)^n\int_{\mathbb{R}^d}\mathrm{d}x\,(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\delta)(x)\varphi(x),$$

where we have omitted the boundary terms since φ is vanishing at boundary. Then, plugging this back, we find

$$\int_{\mathbb{R}^d}\mathrm{d}x\,f(x)\varphi(x)=\int_{\mathbb{R}^d}\mathrm{d}x\left[\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}M_n^{\alpha_1\cdots\alpha_n}(f)(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\delta)(x)\right]\varphi(x).$$

Since φ is arbitrary, we finall arrive at

$$f(x)=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}M_n^{\alpha_1\cdots\alpha_n}(f)\,(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\delta)(x). \tag{1.9}$$

With this expansion, the left hand side is not a function anymore, but shall be treated as a generalized function (as the same in the right hand side).

Moments also relate to Fourier transformation. We have the Fourier transformation

$$\hat{f}(k)=\int_{\mathbb{R}^d}\mathrm{d}x\exp(-\mathrm{i}k_\alpha x^\alpha)f(x).$$

Taking derivatives on k gives

$$\begin{aligned}(\partial^{\alpha_1}\cdots\partial^{\alpha_n}\hat{f})(0)&=\lim_{k\to 0}\int_{\mathbb{R}^d}\mathrm{d}x\exp(-\mathrm{i}k_\alpha x^\alpha)f(x)(-\mathrm{i})^n(x^{\alpha_1}\cdots x^{\alpha_n})\\&=(-\mathrm{i})^n\int_{\mathbb{R}^d}\mathrm{d}x\,f(x)(x^{\alpha_1}\cdots x^{\alpha_n}).\end{aligned}$$

Namely, the Taylor expansion coefficient of $\hat{f}$ is $(-\mathrm{i})^nM_n(f)$, and

$$\hat{f}(k)=\sum_{n=0}^{\infty}\frac{(-\mathrm{i})^n}{n!}M_n^{\alpha_1\cdots\alpha_n}(f)(k_{\alpha_1}\cdots k_{\alpha_n}). \tag{1.10}$$

This relates the moments and the Fourier transformation of a function in analytic Schwartz space. *We can construct the f out of a sequence of moments.* For ensuring convergence, we further demand that $\hat{f}$ is analytic, namely $\hat{f}\in\mathcal{S}_A(\mathbb{R}^d)$.

1.6. Expansion of Transition Rate

Now we apply the expansion derived in the previous section to transition rate $r(x,y)$. Denote $R_n(x)$ as the n-th order moment of the map $\epsilon\mapsto r(x,x+\epsilon)$, that is (moment is defined by equation 1.8)

$$R_n^{\alpha_1\cdots\alpha_n}(x):=\int_{\mathbb{R}^d}\mathrm{d}\epsilon\,(\epsilon^{\alpha_1}\cdots\epsilon^{\alpha_n})r(x,x+\epsilon). \tag{1.11}$$

Using equation 1.9, we directly get

$$r(x,x+\epsilon)=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}R_n^{\alpha_1\cdots\alpha_n}(x)\,(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\delta)(\epsilon). \tag{1.12}$$

It claims that transition rate r, thus the time evolution of wave-function (equation 1.1), is completely determined by the moments R_ns.

Plugging equation 1.12 back to equation 1.1 gives

$$\mathrm{i}\frac{\partial f}{\partial t}(x,t)=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_{\mathbb{R}^d}\mathrm{d}w\, f(w,t)R_n^{\alpha_1\cdots\alpha_n}(w)\,(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\delta)(x-w).$$

Using the parity of $\partial^n\delta$, we change $x-w$ to $w-x$ in $(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\delta)(x-w)$. Then, after integration by parts, we integrate over w, which results in

$$\mathrm{i}\frac{\partial f}{\partial t}(x,t)=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}(\partial_{\alpha_1}\cdots\partial_{\alpha_n})[R_n^{\alpha_1\cdots\alpha_n}(x)f(x,t)]. \tag{1.13}$$

This is a quantum analogy to the Kramers-Moyal expansion in stochastic process.

Interestingly, the Taylor expansion of the "Hamiltonian" $\hat{r}(x,k)$ also relates to the moments R_ns (the $\hat{r}(x,k)$ is defined by equation 1.4). Indeed, if we Taylor expand $\hat{r}(x,k)$ by k at origin, then the coefficient is

$$\lim_{k\to 0}\frac{\partial}{\partial k_{\alpha_1}}\cdots\frac{\partial}{\partial k_{\alpha_n}}\int_{\mathbb{R}^d}\mathrm{d}\epsilon\exp(-\mathrm{i}k_\alpha\epsilon^\alpha)r(x,x+\epsilon)=(-\mathrm{i})^n\left[\int_{\mathbb{R}^d}\mathrm{d}\epsilon\,(\epsilon^{\alpha_1}\cdots\epsilon^{\alpha_n})r(x,x+\epsilon)\right].$$

The $[\cdots]$ is recognized as R_n. So, $\hat{r}(x,k)$ can be Taylor expanded as

$$\hat{r}(x,k)=\sum_{n=0}^{\infty}\frac{(-\mathrm{i})^n}{n!}R_n^{\alpha_1\cdots\alpha_n}(x)(k_{\alpha_1}\cdots k_{\alpha_n}). \tag{1.14}$$

Again, the details of $S(x,k)$ can be completely determined by the moments R_ns.

Comparing with the traditional Hamiltonian $\hat{r}(x,k)=k^2/(2m)+V(x)$, we have all $R_n(x)=0$ except for $R_0(x)=V(x)$ and $R_2(x)=-1/m$. Equation 1.12 implies

$$r(x,y)=V(x)\delta(y-x)-\frac{1}{2m}\nabla^2\delta(y-x).$$

Then equation 1.1 becomes (after integration by parts and omitting the boundary term)

$$\mathrm{i}\frac{\partial f}{\partial t}(x,t)=\int_{\mathbb{R}}\mathrm{d}y f(y,t)\left[V(x)\delta(y-x)-\frac{1}{2m}\nabla^2\delta(y-x)\right]=-\frac{1}{2m}\nabla^2 f(x,t)+V(x)f(x,t),$$

which is exactly the Schrödinger equation.

1.7. On Analyticity

We have set wave-functions in the analytic Schwartz space in which every function is analytic. At the first sight, analyticity seems too strict. But practically, when we solve Schrödinger equation on some explicit instances such as harmonic oscillator or hydrogen atom (in both angular and radical equations), we does restrict ourselves to analytic functions.

For example, when solving the stationary Schrödinger equation of one-dimensional harmonic oscillator, we suppose that wave-function has the form

$$f(x)=\exp(-x^2)\left[\sum_{n=0}^{\infty}a_n x^n\right],$$

where a_ns are to be determined. The factor $\exp(-x^2)$ is employed for an exponentially fast decay as $|x|\to\infty$ (thus faster than the inverse of any polynomial). And the factor $[\cdots]$ is a Taylor series. Namely, f is in the analytic Schwartz space $\mathcal{S}_A(\mathbb{R})$. The energy quantization emerges for ensuring the convergence of the series $\sum_n a_n x^n$.

In the traditional textures of quantum mechanics, wave-functions are regarded as elements in $L^2(\mathbb{R}^d)$, such that $\int\mathrm{d}x|f(x)|^2=1$ (remark that $|\cdots|$ denotes the norm of complex number). In addition, some order of smoothness are appended so as to make the Schrödinger equation (involving a Laplacian) well-defined. Apparently, we have $\mathcal{S}_A(\mathbb{R}^d)\subset L^2(\mathbb{R}^d)$. But, while solving a Schrödinger equation, we only consider the functions in $\mathcal{S}_A(\mathbb{R}^d)$. So, as I think, the space $\mathcal{S}_A(\mathbb{R}^d)$ has been sufficient.

Mathematically, if $\mathcal{S}_A(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, then analyticity is fine. It means, we can approximate any function in $L^2(\mathbb{R}^d)$ by using a series of analytic Schwartz functions. This condition is much easier to be held yet quite powerful. The key point is that a function has negligible measure on the "boundary" of $\mathbb{R}^d$. Precisely, for any $\varepsilon > 0$, there is an $X > 0$ such that

$$\int_{\partial\mathbb{R}^d(X)} \mathrm{d}x\, |f(x)|^2 < \varepsilon$$

where the set $\partial\mathbb{R}^d(X) := \{x \in \mathbb{R}^d \colon \|x\| > X\}$. So, we can simply omit the region $\partial\mathbb{R}^d(X)$ and consider the f on the closed region $U_X(0) := \{x \in \mathbb{R}^d \colon \|x\| \leqslant X\}$. Within this compact region, Weierstrass approximation theorem claims that f can be approximate by a polynomial at any precision. But a polynomial cannot be negligible on the "boundary" $\partial\mathbb{R}^d(X)$. Instead, it will increase as $\|x\|^n$ where n is the order of the polynomial. Inspired by the solution of harmonic oscillator, we weight it by $\exp(-\|x\|^2)$. Namely, we use $p(x)\exp(-\|x\|^2)$ to approximate f on $\mathbb{R}^d$, which means the polynomial $p(x)$ approximates $f(x)\exp(\|x\|^2)$ in the compact region $U_X(0)$. In fact, any function $w \in \mathcal{S}_A(\mathbb{R}^d)$ that has a continuous inverse (namely, $1/w(x)$ is continuous on x) can serve as the weight function, and $\exp(-\|x\|^2)$ is just an instance. We thus conclude, *for any function $f \in L^2(\mathbb{R}^d)$ and any $\varepsilon > 0$, there are a polynomial p and an analytic Schwartz function w such that*

$$\int_{\mathbb{R}^d} \mathrm{d}x\, |f(x) - p(x)w(x)| < \varepsilon. \tag{1.15}$$

The $p(x)w(x)$ is an analytic Schwartz function.

1.8. Locality Imports a Cut-Off on Moments

We then introduce the third axiom about locality, and discuss what it will induce.

Axiom 1.3. *[Locality] Interaction shall be local.*

This implies that the time evolution (equation 1.1) is local. To make this clear, we consider an example, in which $R_n(x) = c^n$ for some constant c, and set the dimension $d = 1$. Then, the time evolution (equation 1.13) at $x = 0$ becomes

$$\mathrm{i}\frac{\partial f}{\partial t}(0,t) = \sum_{n=0}^{\infty} \frac{(-c)^n}{n!} \frac{\partial^n f}{\partial x^n}(0,t).$$

The last expression happens to be the Taylor series of $f(x,t)$ at $x = -c$, namely $f(-c,t)$. So, we conclude that $R_n(x) = c^n$ for some constant c implies

$$\mathrm{i}\frac{\partial f}{\partial t}(0,t) = f(-c,t).$$

If we change the value of f at $x = -c$, then the time evolution at $x = 0$ changes accordingly. It means non-locality.

In physics, a local equation generally refers to an operation on f which contains f itself and *finite* number of partial derivatives of f, such as

$$\mathrm{i}\frac{\partial f}{\partial t}(x,t) = \mathcal{L}\left(f(x,t), \frac{\partial f}{\partial x}(x,t), \frac{\partial^2 f}{\partial x^2}(x,t), \ldots, \frac{\partial^n f}{\partial x^n}(x,t) \right),$$

where $\mathcal{L}$ is an analytic function. This is easy to understand because to compute $(\partial^n f/\partial x^n)(0,t)$ using numerical method with difference Δx, only $f(x,t)$ with $x \in \{0, \Delta x, \ldots, n\Delta x\}$ are employed. So, $(\partial f/\partial t)(0,t)$ cannot "perceive" the $f(x,t)$ outside the neighborhood $\{x \colon |x| \leqslant n\Delta x\}$. Since Δx can be arbitrarily small (but not vanishing), the equation is local.

In the previous discussion, we have shown that, with a cut-off N_{cut} on R_ns such that $R_n = 0$ for any $n > N_{\mathrm{cut}}$, the time evolution is local. And without such a cut-off, we can construct a sequence of R_ns such that the time evolution is not local. Now, we are to prove that, generally, without a cut-off, any sequence of R_ns (such that for any $N > 0$, there are infinite many R_ns that are not vanishing), the time evolution is non-local. This then imports a cut-off on moments.

Before proving, we first declare a property of analytic function. Consider the Taylor expansion of $f \in \mathcal{S}_A(\mathbb{R}, \mathbb{C})$ at origin

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots.$$

The information of f is completely encoded in the infinite sequence of a_ns. This is the result of a theorem which claims that two analytic functions are equal if they agree in any interval (hence we can obtain the Taylor series of the function within the interval). What if we only know a portion of the infinite sequence of a_ns? For example, if we only know the a_0, then only the value of $f(x)$ at $x = 0$ is determined. Further, if we also know the a_1, then we can give a good approximation to $f(x)$ in a very tiny neighborhood of $x = 0$, since $f(x) = a_0 + a_1 x + o(x^2)$. Then, if we additionally know the a_2, then the approximation becomes as good as before in a larger neighborhood of $x = 0$, since $f(x) = a_0 + a_1 x + a_2 x^2 + o(x^3)$ and for keeping the scale of residue, the size of neighborhood can increase a little. This analysis indicates that the more a_ns we know, in a larger neighborhood of $x = 0$ can we properly approximate $f(x)$. Our vision becomes border and border if we know more and more a_ns. Until knowing the whole sequence of a_ns, we realize the complete picture of $f(x)$ (based on the previous theorem about analytic function).

It also indicates that, for a sufficient large N, we can keep $f(x)$ (approximately) invariant when we tune the a_ns with $n > N$ while keep the other a_ns invariant. So, based on the equation

$$\mathrm{i}\frac{\partial f}{\partial t}(0,t) = \sum_{n=0}^{\infty} \frac{1}{n!} R_n(0) a_n,$$

if there is not a cut-off to the infinite sequence of $R_n(0)$s, we can modify the $(\partial f / \partial t)(0,t)$ by tuning an a_n where n can be arbitrarily large. This, however, will leave the $f(x)$ around $x = 0$ invariant. It means that the value of f with x far away from origin can affect the time evolution of f at origin. That is, time evolution is non-local. Hence, there must be a cut-off on the sequence of $R_n(0)$s. This discussion also holds for any value of x other than $x = 0$. We conclude that *time evolution is local if and only if there is a positive integer N_{cut} such that $R_n = 0$ for any $n > N_{\mathrm{cut}}$.*

1.9. Hermitianity on Moments

Now we study the R_ns. We first investigate what Hermitianity implies for our generic quantum mechanics. Hermitianity implies

$$r^*(x, x+\epsilon) = r(x+\epsilon, x).$$

For simplicity, we temporally abbreviate the notations by $\alpha := (\alpha_1, \ldots, \alpha_n)$ and $\partial_\alpha^n := \partial_{\alpha_1} \cdots \partial_{\alpha_n}$.[1.6] Then, the left hand side is expanded as

$$r^*(x, x+\epsilon) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} (R_n^*)^\alpha(x) (\partial_\alpha^n \delta)(\epsilon),$$

and the right hand side

$$r(x+\epsilon, x) = r(x+\epsilon, (x+\epsilon)-\epsilon) = \sum_{n=0}^{+\infty} \frac{1}{n!} R_n^\alpha(x+\epsilon) (\partial_\alpha^n \delta)(\epsilon),$$

where we have used the relation $(-1)^n (\partial_\alpha^n \delta)(-\epsilon) = (\partial_\alpha^n \delta)(\epsilon)$. Put these together (left hand side minus the right hand side),

$$\sum_{n=0}^{+\infty} \frac{1}{n!} [(-1)^n (R_n^*)^\alpha(x) - R_n^\alpha(x+\epsilon)] (\partial_\alpha^n \delta)(\epsilon) = 0.$$

Applying on an arbitrary function $\varphi \in \mathcal{S}(\mathbb{R}^d, \mathbb{C})$ (integrating over ϵ), we get (after integration by parts)

$$\sum_{n=0}^{+\infty} \frac{1}{n!} [(R_n^*)^\alpha(x) \partial_\alpha^n \varphi(0) + (-1)^{n+1} \partial_\alpha^n [R_n^\alpha(x) \varphi(0)]] = 0,$$

1.6. It turns out that this is not a convenient abbreviation. Call for optimizations.

or using the original notations,

$$\sum_{n=0}^{+\infty} \frac{1}{n!}[(R_n^*)^{\alpha_1\cdots\alpha_n}(x)(\partial_{\alpha_1}\cdots\partial_{\alpha_n}\varphi)(0) + (-1)^{n+1}(\partial_{\alpha_1}\cdots\partial_{\alpha_n})[R_n^{\alpha_1\cdots\alpha_n}(x)\varphi(0)]] = 0.$$

This is what the Hermitianity of transition rate r induces on the relation between moments R_ns and their complex conjugations R_n^*s.

Expanding this summation term by term, the left hand side becomes

$$\begin{aligned}
&[\![n=0]\!] = (R_0^*)(x)\varphi(0) - R_0^\alpha(x)\varphi(0)\\
&[\![n=1]\!] + (R_1^*)^\alpha(x)\partial_\alpha\varphi(0) + \partial_\alpha R_1^\alpha(x)\varphi(x) + R_1^\alpha(x)\partial_\alpha\varphi(0)\\
&[\![n=2]\!] + \frac{1}{2}(R_2^*)^{\alpha\beta}(x)\partial_\alpha\partial_\beta\varphi(0) - \frac{1}{2}\partial_\alpha\partial_\beta R_2^{\alpha\beta}(x)\varphi(0) - \partial_\alpha R_2^{\alpha\beta}(x)\partial_\beta\varphi(0) - \frac{1}{2}R_2^{\alpha\beta}(x)\partial_\alpha\partial_\beta\varphi(0)\\
&[\![n>2]\!] + \cdots
\end{aligned}$$

Let us examine this temporal result. For simplicity, we set $R_n = 0$ for any $n > 2$. If R_2 is constant, then it implies $R_2 \in \mathbb{R}$ (hint: collecting the term proportional to $\partial_\alpha\partial_\beta\varphi(0)$). It in turn leads to $R_1^*(x) = -R_1(x)$, namely $R_1(x)$ is purely imaginary (this is a surprise). And then, $R_0(x) = R_0^*(x) + \partial_\alpha R_1^\alpha(x)$. If further take R_1 as constant, then we get $R_0 \colon \mathbb{R}^d \to \mathbb{R}$. This is the situation in the traditional Hamiltonian with dissipation.[1.7] This is the quantum mechanical analogy of Langevin process, where R_n is cut-off at $n = 2$ (such that $R_n = 0$ for any $n > 2$) and R_2 is constant.

Because the sequence of R_ns is cut-off by N_{cut}, we find $R_{N_{\text{cut}}}(x) = (R_{N_{\text{cut}}}^*)(x)$, thus $R_{N_{\text{cut}}}$ is a real function. Then, by induction, we can obtain a recursive equation that gives the relation between R_ns and their complex conjugations.

1.10. Uncertainty Principle

An **observable** is a real function on the collection of states of a quantum system. The second part of probability interpretation is about the expectation of an observable.

AXIOM 1.4. *[Probability Interpretation (Part 2)] Given an observable O and a wave-function f of quantum system, the expected value of the observable is given by*

$$\mathbb{E}_f[O(t)] = \int_{\mathcal{X}} \mathrm{d}x\, f^*(x,t) f(x,t) O(x).$$

For a single particle system, we use $x(t)$ to denote the position of the particle at time t. Velocity is defined as usual,

$$v(t) := \lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t}.$$

Both position and velocity are observables. Further, we can compute the variance of an observable O as an expectation $\mathrm{Var}_f[O(t)] := \mathbb{E}_f[(O - \mathbb{E}_f[O(t)])^2]$. Then, uncertainty principle claims:

AXIOM 1.5. *[Uncertainty Principle] In a single particle quantum system, given a wave-function f, the variances of position and of velocity have the relation*

$$\Delta x \Delta v := \sqrt{\mathrm{Var}_f[x(t)]\,\mathrm{Var}_f[v(t)]} \sim \hbar/(2m),$$

where m is the mass of the particle, and $\hbar$ is the reduced Planck's constant.

In this section, we are to show how uncertainty principle restricts further to the transition rate. For simplicity, we set $d = 1$ throughout this section. Since axiom 1.5 holds for all wave-functions of a single particle, we choose f to be Gaussian at $t = 0$, that is

$$f(x,0) = (2\pi\sigma^2)^{-1/4} \exp\left(-\frac{x^2}{4\sigma^2}\right).$$

1.7. If $R_1 \neq 0$, such that $\hat{r}(x,k) = k^2/(2m) + \mu k + V(x)$ for some constant μ, then the time-reversal symmetry breaks. Indeed, μk serves as a friction term. But we usually treat systems with dissipation as incomplete.

The coefficient guarantees the normalization of f. By axiom 1.4, we have

$$\mathbb{E}_f[x(0)] := \int_{\mathbb{R}} \mathrm{d}x\, f^*(x,0) f(x,0) x = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}^d} \mathrm{d}x \exp\left(-\frac{x^2}{2\sigma^2}\right) x = 0,$$

then

$$\mathrm{Var}_f[x(0)] := \int_{\mathbb{R}} \mathrm{d}x\, f^*(x,0) f(x,0) x^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}^d} \mathrm{d}x \exp\left(-\frac{x^2}{2\sigma^2}\right) x^2 = \sigma^2.$$

To evaluate $\mathrm{Var}_f[v(0)]$, we have to use the time evolution to calculate $x(\Delta t)$. By equation 1.1, we have (for brevity, we omit the subscript cut in N_{cut})

$$f(x,\Delta t) = f(x,0) - \mathrm{i}\Delta t \sum_{n=0}^{N} \frac{(-1)^n}{n!} \partial^n [R_n(x) f(x,0)] - \frac{\mathrm{i}}{2} \Delta t^2 \sum_{n=0}^{N} \sum_{n'=0}^{N} \cdots.$$

We have to evaluate up to $o(\Delta t^2)$, so as to give the variance of velocity which scales as $x^2(\Delta t)/\Delta t^2$. But this would be too complicated (the double summation). For simplification, we consider a "free particle" where all R_ns but the R_N are vanishing. This is the situation when σ tends to zero, because there are more σ factors in the denominator if there are more derivatives on f, and as σ tends to zero, the term proportional to $R_N(x)\partial^N[f(x,0)]$ surpasses all the other terms.[1.8] In this situation, equation 1.1 reduces to

$$\mathrm{i}\frac{\partial f}{\partial t}(x,t) = \frac{(-1)^N}{N!} R_N(x) \partial^N f(x,t).$$

Hence,

$$\begin{aligned} f(x,\Delta t) &= f(x,0) + \frac{\partial f}{\partial t}(x,0)\Delta t + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(x,0)\Delta t^2 + o(\Delta t^2) \\ &= f(x,0) - \mathrm{i}\frac{(-1)^N}{N!} R_N(x)\partial^N f(x,0)\Delta t - \frac{\mathrm{i}}{2}\frac{(-1)^N}{N!} R_N(x)\partial^N \frac{\partial f}{\partial t}(x,0)\Delta t^2 + o(\Delta t^2). \end{aligned}$$

Inserting $(\partial f/\partial t)$ again gives

$$f(x,\Delta t) = f(x,0) - \mathrm{i}\frac{(-1)^N}{N!} R_N(x)\partial^N f(x,0)\Delta t - \frac{1}{2}\left[\frac{(-1)^N}{N!} R_N(x)\right]^2 \partial^{2N} f(x,0)\Delta t^2 + o(\Delta t^2).$$

Now we are to evaluate $\mathrm{Var}_f[v(0)]$. First, we consider $\mathbb{E}_f[v(0)] = \mathbb{E}_f[x(\Delta t) - x(0)]/\Delta t$. The f used for the expectation is $f(x,\Delta t)$ instead of $f(x,0)$. So, it cannot be expanded as $(\mathbb{E}_f[x(\Delta t)] - \mathbb{E}_f[x(0)])/\Delta t$ except for treating $x(0)$ as a constant, the constant $\mathbb{E}_f[x(0)]$ which has been evaluated as zero. Then, we have

$$\mathbb{E}_f[v(0)] = \frac{\mathbb{E}_f[x(\Delta t)]}{\Delta t} = \int_{\mathbb{R}} \mathrm{d}x\, f^*(x,\Delta t) f(x,\Delta t) \frac{x}{\Delta t}$$

Inserting $f(x,\Delta t)$ up to $o(\Delta t)$, and knowing that R_N is real (see section 1.9), we find $\mathbb{E}_f[v(0)] = o(1)$. Then we have (using the famous formula of variance)

$$\mathrm{Var}_f[v(0)] = \mathbb{E}_f[v^2(0)] - \mathbb{E}_f^2[v(0)] = \mathbb{E}_f[v^2(0)] + o(1).$$

Now focus on $\mathbb{E}_f[v^2(0)]$. Recalling that $x(0)$ is treated as the constant $\mathbb{E}_f[x(0)] = 0$,

$$\mathbb{E}_f[v^2(0)] = \frac{\mathbb{E}_f[(x(\Delta t) - \mathbb{E}_f[x(0)])^2]}{\Delta t^2} = \frac{\mathbb{E}_f[x^2(\Delta t)]}{\Delta t^2}.$$

Hence,

$$\begin{aligned} \mathbb{E}_f[v^2(0)] &= \int_{\mathbb{R}} \mathrm{d}x\, f^*(x,\Delta t) f(x,\Delta t) \frac{x^2}{\Delta t^2} \\ &= \left(\frac{R_N(x)}{N!}\right)^2 \left[\int_{\mathbb{R}} \mathrm{d}x \partial^N f(x,0) \partial^N f(x,0) x^2 - \int_{\mathbb{R}} \mathrm{d}x \partial^{2N} f(x,0) f(x,0) x^2\right] + o(1). \end{aligned}$$

1.8. In the traditional approach of quantum mechanics, $\sigma \to 0$ indicates that the momentum is large (since momentum is proportional to $\partial/\partial x$), so the term with the highest order of momentum will dominates the time evolution. When the momentum is large enough, the kinetic term dominates Hamiltonian, the potential becomes omittable, making the particle free (namely, unconstrained by potential).

Using maxima,we get the table 1.1. We find that only when $N=2$, and when R_2 is a constant, can $\Delta x\Delta v$ satisfy the uncertainty principle 1.5. Any other situation depends on σ, which is either vanishing or diverging when σ tends to zero. We then conclude that *uncertainty principle forces that $N=2$, and together with Hermitianity, R_2 is a real constant*.

N	$\Delta x\Delta v$
1	$\frac{R_1(x)\sigma}{\sqrt{2}}$
2	$\frac{R_2(x)}{2}$
3	$\mathrm{i}\sqrt{\frac{7}{6}}\frac{R_3(x)}{8\sigma}$
4	$\sqrt{\frac{5}{3}}\frac{R_4(x)}{16\sigma^2}$
5	$\mathrm{i}\sqrt{\frac{217}{30}}\frac{R_5(x)}{128\sigma^3}$
6	$\sqrt{\frac{21}{5}}\frac{R_6(x)}{256\sigma^4}$

Table 1.1. These are computed by Maxima. Define `f: (2*%pi*σ^2)^(-1/4)*exp(-x^2/(4*σ^2));`, `N: 5;` (for instance), and `g: diff(f, x, N)*diff(f, x, N)*x^2-diff(f, x, 2*N)*f*x^2;` which is the $[\cdots]$ part in the expression of $\langle v^2(0)\rangle$. Then $\Delta x\Delta v$ is given by `sqrt((R/N!)^2*integrate(g, x, -inf, inf)*σ^2);`.