Quantum mechanics for probabilists

Sabine Jansen

January 15, 2021

Abstract

Many topics of modern probability have counterparts in mathematical physics and quantum mechanics. For example, the study of the parabolic Anderson model is related to Anderson localization; interacting particle systems and spin systems are related to quantum spin systems and quantum many-body theory; and the Gaussian free field as well as Malliavin calculus connect to Euclidean quantum field theory. The aim of these notes is to give an introduction to quantum mechanics for mathematicians with a probability background, providing basic intuition and a dictionary facilitating access to the mathematical physics literature. The focus is on connections with probability, notably Markov processes, rather than partial differential equations and spectral theory.

Contents

1 Waves and particles

1.1 Particles

Newton's second law of motion

In classical mechanics, the motion of a particle of mass m in an external force field $\boldsymbol{F}(\boldsymbol{x})$ is described with ordinary differential equations. Let

$$\mathbb{R} \to \mathbb{R}^3, \quad t \mapsto \boldsymbol{x}(t)$$

be a smooth curve, the *trajectory* of the particle. Time-derivatives are denoted with dots. The *velocity* and *acceleration* at time t are given by the first and second time-derivatives

$$\boldsymbol{v}(t) = \dot{\boldsymbol{x}}(t) = \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}(t), \quad \boldsymbol{a}(t) = \ddot{\boldsymbol{x}}(t) = \frac{\mathrm{d}^2\boldsymbol{x}}{\mathrm{d}t^2}(t).$$

Newton's second law of motion says that force is mass times acceleration, $\boldsymbol{F} = m\boldsymbol{a}$. This translates into an ordinary differential equation

$$m\ddot{\boldsymbol{x}}(t) = \boldsymbol{F}\big(\boldsymbol{x}(t)\big). \tag{1.1}$$

The force field $\boldsymbol{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is *conservative* if there exists a function $V : \mathbb{R}^3 \to \mathbb{R}$, such that $\boldsymbol{F} = -\nabla V$. The function V is called *potential*. If $\boldsymbol{x}(t)$ solves the ODE (1.1) and the force is conservative, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}m|\dot{\boldsymbol{x}}(t)|^2 + V\big(\boldsymbol{x}(t)\big)\Big) = 0,$$

i.e., the *energy* $\frac{1}{2}m|\dot{\boldsymbol{x}}|^2 + V(\boldsymbol{x})$ is conserved. The energy is the sum of the *kinetic energy* $\frac{1}{2}m|\dot{\boldsymbol{x}}|^2$ and the *potential energy* $V(\boldsymbol{x})$.

Hamilton function. Momentum

The energy is closely related to the *Hamilton function* $H : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ given by

$$H(\boldsymbol{p}, \boldsymbol{x}) = \frac{1}{2m}|\boldsymbol{p}|^2 + V(\boldsymbol{x}). \tag{1.2}$$

The variable $\boldsymbol{p}$ is called *momentum*. Hamilton functions are important in classical mechanics because they encode the dynamics. Indeed, the differential equation (1.1) is equivalent to the system

$$\dot{\boldsymbol{p}}(t) = -\nabla_{\boldsymbol{x}} H\big(\boldsymbol{p}(t), \boldsymbol{x}(t)\big), \quad \dot{\boldsymbol{x}}(t) = \nabla_{\boldsymbol{p}} H\big(\boldsymbol{p}(t), \boldsymbol{x}(t)\big). \tag{1.3}$$

The second equation reads $\dot{\boldsymbol{x}} = \frac{1}{m}\boldsymbol{p}$, which gives the relation

$$\boldsymbol{p} = m\boldsymbol{v}.$$

For the Hamiltonian (1.2), momentum is mass times velocity.

Warning. The relation $\boldsymbol{p} = m\boldsymbol{v}$ is correct for the Hamilton function (1.2) which encodes the ODE (1.1) with conservative force field $\boldsymbol{F} = -\nabla V$. For other Hamilton functions, encoding different ODEs, the relation between momentum and velocity may change. For example, the Hamilton function for a relativistic particle with rest mass m without external force is

$$H(\boldsymbol{p}, \boldsymbol{x}) = \sqrt{m^2c^4 + c^2|\boldsymbol{p}|^2}$$

with $c > 0$ the speed of light (3×10^8 meter per second). The reader may note the similarity with Einstein's famous formula

$$E = mc^2.$$

The system of ODEs (1.3) becomes $\dot{\boldsymbol{p}} = 0$ and

$$\dot{\boldsymbol{x}} = \nabla_{\boldsymbol{p}} H = \frac{c^2}{\sqrt{m^2c^4 + c^2|\boldsymbol{p}|^2}}\,\boldsymbol{p} = \frac{1}{m\sqrt{1 + |\boldsymbol{p}|^2/(m^2c^2)}}\,\boldsymbol{p}.$$

instead of $\dot{\boldsymbol{x}} = \frac{1}{m}\boldsymbol{p}$. For a massless particle ($m = 0$), the relations read

$$H(\boldsymbol{p}, \boldsymbol{x}) = c\,|\boldsymbol{p}|, \quad \dot{\boldsymbol{x}} = c\,\frac{\boldsymbol{p}}{|\boldsymbol{p}|}.$$

In particular, the speed is equal to the speed of light, as the velocity has norm $|\dot{\boldsymbol{x}}| = c$. We call these particles *photons*.

1.2 Waves

Plane wave

Let $\boldsymbol{k} \in \mathbb{R}^3$ and $\omega > 0$. We call the function

$$\varphi(\boldsymbol{x}, t) = \exp\big(\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)\big) \tag{1.4}$$

as well as its real and imaginary parts $\cos(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)$ and $\sin(\boldsymbol{k} \cdot x - \omega t)$ *monochromatic plane waves* with *wave vector* $\boldsymbol{k}$ and *angular frequency* ω. It function represents an oscillating signal that propagates in the direction of the vector $\boldsymbol{k}$ at speed $\omega/|\boldsymbol{k}|$ (for $\boldsymbol{k} \neq 0$). The signal could be, for example, light, sound, or an electromagnetic wave.

In the direction of $\boldsymbol{k}$, the signal is periodic in $\boldsymbol{x}$ with period

$$\lambda = \frac{2\pi}{|\boldsymbol{k}|},$$

called *wave length*. For light, different wavelengths λ correspond to different colors, from violet (400 nanometer) to red (700 nanometer). The signal is also periodic in time with period

$$T = \frac{2\pi}{\omega}.$$

The inverse

$$\nu = \frac{1}{T} = \frac{\omega}{2\pi}$$

is the *frequency* of the wave.

The angular frequency is considered to be a function of $\boldsymbol{k}$; functions are called *dispersion relations*. For example, light in the vacuum propagates at constant speed $\omega/|\boldsymbol{k}| = c = 3 \times 10^8$ meter per second, hence

$$\omega = \omega(\boldsymbol{k}) = c\,|\boldsymbol{k}|. \tag{1.5}$$

Other signals or light in a medium may have different speeds of propagation, hence different dispersion relations $\omega(\boldsymbol{k})$. The dispersion relation is non-linear when the speed of propagation depends on the wavelength—as is the case for light in air or water, a phenomenon at the origin of rainbows.

Wave equation

The plane wave (1.4) satisfies

$$\frac{\partial^2\varphi}{\partial t^2}(\boldsymbol{x},t) = -\omega^2\varphi(\boldsymbol{x},t), \quad \Delta\varphi(\boldsymbol{x},t) = -|\boldsymbol{k}|^2\varphi(\boldsymbol{x},t).$$

For the dispersion relation (1.5), we have $\omega^2 = c^2|\boldsymbol{k}|^2$ hence

$$\frac{\partial^2\varphi}{\partial t^2} - c^2\Delta\varphi = 0, \tag{1.6}$$

plane waves solve the *wave equation* (1.6). Other solutions are obtained as superpositions of plane waves

$$\varphi(\boldsymbol{x},t) = \int A(\boldsymbol{k})\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega(\boldsymbol{k})t)}\,\mathrm{d}\boldsymbol{k}. \tag{1.7}$$

1.3 Wave packet

Let us have a closer look at the superposition (1.7) of plane waves when the amplitude $A(\boldsymbol{k})$ is proportional to the indicator of a small interval. To simplify matters, we work in $\mathbb{R}$ instead of $\mathbb{R}^3$.

Suppose we are given a dispersion relation $\omega(k)$, not necessarily of the form $\omega(k) = c|k|$. Pick $k_0 \in \mathbb{R}$, $\varepsilon > 0$ and consider the function

$$\varphi(x,t) = \frac{1}{2\varepsilon}\int_{k_0-\varepsilon}^{k_0+\varepsilon} \exp\bigl(\mathrm{i}(kx - \omega(k)t)\bigr)\,\mathrm{d}k.$$

In general the integral cannot be computed explicitly, but for small ε there is a classical heuristic computation. Define

$$\operatorname{sinc}(x) = \begin{cases} \frac{\sin x}{x}, & x \neq 0, \\ 1, & x = 0, \end{cases}$$

the *sine cardinal*, and let

$$v_0 = \frac{\mathrm{d}\omega}{\mathrm{d}k}(k_0).$$

On $[k_0 - \varepsilon, k_0 + \varepsilon]$ we approximate $\omega(k) \approx \omega_0 + v_0(k - k_0)$ and

$$\begin{aligned} \varphi(x,t) &\approx \frac{1}{2\varepsilon} \int_{k_0-\varepsilon}^{k_0+\varepsilon} \exp\Big(\mathrm{i}\big(kx - [\omega_0 + v_0(k-k_0)]t\big)\Big)\,\mathrm{d}k \\ &= \mathrm{e}^{\mathrm{i}(k_0 x - \omega_0 t)}\, \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \mathrm{e}^{\mathrm{i}q(x - v_0 t)}\,\mathrm{d}q \\ &= \mathrm{e}^{\mathrm{i}(k_0 x - \omega_0 t)} \operatorname{sinc}\Big(\varepsilon(x - v_0 t)\Big). \end{aligned}$$

The approximate is the product of the *envelope* $\operatorname{sinc}[\varepsilon(x - v_0 t)]$ and the oscillating signal $\exp(\mathrm{i}(k_0 x - \omega_0 t))$. The envelope is maximal at $x = x_0(t) = v_0 t$ and goes to zero as $|x - x_0(t)| \to \infty$. The peak $x_0(t)$ of the envelope moves at speed v_0, called *group velocity*.

The envelope spreads out more when ε is small. As a measure for the width of the envelope, we can take the distance π/ε between the peak $v_0 t$ and the zeros $\varepsilon(x - v_0 t) = \pi$. Notice that the width π/ε is large when the window $[k_0 - \varepsilon, k_0 + \varepsilon]$ is small:

> The spread of the wave packet in x-space and the window in k-space cannot both be small. The smaller the window in k-space, the larger the spread in x-space.

Often $2\pi/k_0$ is small compared to the envelope's width π/ε, then the real part of $\varphi(\boldsymbol{x}, t)$ is drawn as a rapidly oscillating signal varying within the borders prescribed by the envelope; the reader is is encouraged to do a quick internet search for the keyword "wave packet" and look for pictures and animations.

1.4 Wave-particle duality

Light

Is light a wave or a stream of particles? At the end of the 19th century, the answer was: a wave. Around 1900 however, in order to explain blackbody radiation, Max Planck postulated that energy in electromagnetic waves is exchanged with matter in discrete units or *quanta* of energy. The energy quantum E is proportional to the wave's frequency ν, the proportionality constant $h > 0$ is *Planck's constant*. Thus

$$E = h\nu.$$

The relation can also be written with the angular frequency ω and the *reduced Planck constant*

$$\hbar = \frac{h}{2\pi}$$

(latex command `\hbar`) as

$$E = \hbar\omega. \tag{1.8}$$

It looks as if light of angular frequency ω is made of particles of energy $E = \hbar\omega$. Because of the formula $H(\boldsymbol{p}, \boldsymbol{x}) = c|\boldsymbol{p}|$ for the Hamiltonian of a photon and the dispersion relation $\omega(\boldsymbol{k}) = c|\boldsymbol{k}|$, it is natural to identify

$$\boldsymbol{p} = \hbar\boldsymbol{k}. \tag{1.9}$$

Later developments confirmed that both points of view are useful: Some phenomena are best explained by viewing light as a wave, others are best explained by viewing light as a stream of particles (*photons*).

Matter

In the 1920s De Broglie suggested corpuscular matter can be viewn as a wave too. If we keep the relations (1.8) and (1.9) and combine them with the expression $H(\boldsymbol{p}, \boldsymbol{x}) = |\boldsymbol{p}|^2/(2m)$ for the Hamiltonian of a free particle (no external force), we arrive at the formula

$$E = \frac{|\boldsymbol{p}|^2}{2m} = \frac{\hbar^2}{2m}|\boldsymbol{k}|^2$$

for the energy of a particle and at the dispersion relation

$$\omega(\boldsymbol{k}) = \frac{1}{\hbar}E = \frac{\hbar}{2m}|\boldsymbol{k}|^2.$$

With this dispersion relation, the plane wave

$$e_{\boldsymbol{k}}(\boldsymbol{x}, t) = \exp\bigl(\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - \omega(\boldsymbol{k})t\bigr)$$

satisfies

$$\mathrm{i}\hbar\frac{\partial}{\partial t}e_{\boldsymbol{k}} = \frac{\hbar^2}{2m}|\boldsymbol{k}|^2 e_{\boldsymbol{k}} = -\frac{\hbar^2}{2m}\Delta e_{\boldsymbol{k}}.$$

Superpositions of plane waves as in (1.7) satisfy

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\varphi = -\frac{\hbar^2}{2m}\Delta\varphi, \tag{1.10}$$

which is the *Schrödinger equation* for a particle of mass m in the absence of an external potential V.

2 Quantum mechanics

2.1 Wave function

The basic object of the Schrödinger picture is a measurable function $\psi : \mathbb{R}^d \to \mathbb{C}$ that is square-integrable with L^2-norm 1,

$$||\psi||^2 = \int_{\mathbb{R}^d} |\psi(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} = 1. \tag{2.1}$$

The Fourier transform is denoted $\widehat{\psi}$. It is given by

$$\widehat{\psi}(\boldsymbol{k}) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \psi(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$

if the integral is absolutely convergent, which is the case for a dense class of functions $\psi \in L^2(\mathbb{R}^d)$. The formula for the Fourier transform is inverted as

$$\psi(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \widehat{\psi}(\boldsymbol{k}) \, \mathrm{d}\boldsymbol{k}. \tag{2.2}$$

The *Plancherel theorem* says that a function ψ and its Fourier transform have the same L^2-norm, hence

$$||\widehat{\psi}||^2 = \int_{\mathbb{R}^d} |\widehat{\psi}(\boldsymbol{k})|^2 \, \mathrm{d}\boldsymbol{k} = 1. \tag{2.3}$$

In view of the normalizations (2.1) and (2.3), it is tempting to interpret $|\psi(\boldsymbol{x})|^2$ and $|\widehat{\psi}(\boldsymbol{k})|^2$ as probability densities, and this is indeed what we shall do:

> $|\psi(\boldsymbol{x})|^2$ is the probability density function for the position of the particle.

while

> $|\widehat{\psi}(\boldsymbol{k})|^2$ is the probability density function for $\frac{1}{\hbar}\boldsymbol{p}$, with $\boldsymbol{p}$ the momentum and $\hbar$ the reduced Planck constant.

Let us emphasize one fact:

> *One* function ψ encodes simultaneously *two* probability distributions!

As a consequence, fixing the probability distribution of the particle position imposes restrictions on the momentum distribution and vice-versa.

2.2 Uncertainty principle

Let us fix a wave function $\psi \in L^2(\mathbb{R}^d)$. It is customary in quantum mechanics to use angular brackets for expectations. For $j = 1, \dots, d$, define

$$\langle X_j \rangle := \int_{\mathbb{R}^d} x_j \, |\psi(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x}, \qquad \langle X_j^2 \rangle := \int_{\mathbb{R}^d} x_j^2 \, |\psi(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x}$$
$$\langle P_j \rangle := \int_{\mathbb{R}^d} \hbar k_j \, |\widehat{\psi}(\boldsymbol{k})|^2 \, \mathrm{d}\boldsymbol{k}, \qquad \langle P_j^2 \rangle := \int_{\mathbb{R}^d} \hbar^2 k_j^2 \, |\widehat{\psi}(\boldsymbol{k})|^2 \, \mathrm{d}\boldsymbol{k}.$$

We may think of $\boldsymbol{X}$ as a random position vector with probability distribution function $|\psi(\boldsymbol{x})|^2$, then $\langle X_j \rangle$ and $\langle X_j^2 \rangle$ represent the expected value and second moment of the component X_j. Expected values of other polynomials are defined in a similar manner, e.g., the variance of X_j is

$$\Big\langle \big(X_j - \langle X_j \rangle\big)^2 \Big\rangle = \int_{\mathbb{R}^d} \big(x_j - \langle X_j \rangle\big)^2 |\psi(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} = \langle X_j^2 \rangle - \langle X_j \rangle^2.$$

Finally define deviations

$$\sigma_{\boldsymbol{X}} := \left(\Big\langle \sum_{j=1}^{d} \big(X_j - \langle X_j \rangle\big)^2 \Big\rangle \right)^{1/2}, \quad \sigma_{\boldsymbol{P}} := \left(\Big\langle \sum_{j=1}^{d} \big(P_j - \langle P_j \rangle\big)^2 \Big\rangle \right)^{1/2}$$

In general the integrals considered above may diverge, but we will only consider wave functions ψ for which

$$\langle \, |\boldsymbol{X}|^2 \rangle = \Big\langle \sum_{j=1}^{d} X_j^2 \Big\rangle, \quad \langle \, |\boldsymbol{P}|^2 \rangle = \Big\langle \sum_{j=1}^{d} P_j^2 \Big\rangle$$

are both finite. For such wave functions ψ, all integrals considered above are absolutely convergent.

Theorem 2.1. *Let $\psi \in L^2(\mathbb{R}^d)$ be such that $\langle |\boldsymbol{X}|^2 \rangle < \infty$ and $\langle |\boldsymbol{P}|^2 \rangle < \infty$. Then*

$$\sigma_{\boldsymbol{X}} \, \sigma_{\boldsymbol{P}} \geq \frac{\hbar}{2}.$$

Example 2.2 (Gaussian wave function). It is instructive to work out an example in dimension $d = 1$. Pick $s > 0$ and set

$$\psi(x) = \frac{1}{(2\pi s^2)^{1/4}} \, \exp\Big(-\frac{x^2}{4s^2}\Big).$$

Then $\int |\psi(x)|^2 \, \mathrm{d}x = 1$ and the Fourier transform is

$$\widehat{\psi}(k) = \frac{1}{(2\pi/[4s^2])^{1/4}} \, \exp\Big(-s^2 k^2\Big).$$

Noticing that

$$|\psi(x)|^2 = \frac{1}{\sqrt{2\pi s^2}} \exp\Big(-\frac{x^2}{2s^2}\Big), \quad |\widehat{\psi}(k)|^2 = \frac{1}{\sqrt{2\pi/[4s^2]}} \exp\Big(-\frac{k^2}{2/[4s^2]}\Big),$$

we recognize the probability density functions of normal random variables with variances s^2 and $(4s^2)^{-1}$. It follows that

$$\sigma_X = s, \quad \sigma_P = \frac{\hbar}{2s}$$

hence $\sigma_X \, \sigma_P = \hbar/2$: for Gaussian wave functions, the inequality from Theorem 2.1 holds true and is in fact an equality.

The uncertainty principle shows that $\sigma_{\boldsymbol{X}}$ and $\sigma_{\boldsymbol{P}}$ cannot both be equal to zero. A probabilist might say:

> There is no way to make the position and the momentum both deterministic,

though a physicist may object to the use of the word "deterministic".

More generally, the smaller the variance of the momentum, the larger the variance of the position, and vice-versa, which should remind the reader of the considerations on wave packets in Section 1.3.

2.3 Observables as operators

How do we express expected values of functions of position and momentum? For functions that depend on position or momentum but not both, we simply use integrals involving the probability densities $|\psi(\boldsymbol{x})|^2$ and $|\widehat{\psi}(\boldsymbol{k})|^2$. For functions that depend on the position and the momentum, the probabilist's first instinct is to look for a joint distribution of the vector $(\boldsymbol{P}, \boldsymbol{X})$, however quantum mechanics proceeds differently.

Hamiltonian

Consider for example the Hamiltonian of a particle of mass m in an external potential $V : \mathbb{R}^d \to \mathbb{R}$,

$$H = \frac{1}{2m}|\boldsymbol{P}|^2 + V(\boldsymbol{X}).$$

It is reasonable to define the expected value as

$$\begin{aligned}\langle H \rangle &= \Big\langle \frac{1}{2m}|\boldsymbol{P}|^2 \Big\rangle + \big\langle V(\boldsymbol{X}) \big\rangle \\ &= \int_{\mathbb{R}^d} \frac{\hbar^2}{2m}|\boldsymbol{k}|^2 |\widehat{\psi}(k)|^2 \, \mathrm{d}\boldsymbol{k} + \int_{\mathbb{R}^d} V(\boldsymbol{x}) \, |\psi(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x}.\end{aligned}$$

Mathematical properties of the Fourier transform allow us to replace the integral over $\boldsymbol{k}$ by an integral over $\boldsymbol{x}$. The equality

$$\int_{\mathbb{R}^d} |\boldsymbol{k}|^2 |\widehat{\psi}(k)|^2 \,\mathrm{d}\boldsymbol{k} = \int_{\mathbb{R}^d} \overline{\psi(\boldsymbol{x})} \big(-\Delta\psi\big)(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}$$

yields

$$\langle H \rangle = \int_{\mathbb{R}^d} \overline{\psi(\boldsymbol{x})} \Big(-\frac{\hbar^2}{2m} \Delta\psi(\boldsymbol{x}) + V(\boldsymbol{x})\psi(\boldsymbol{x})\Big) \,\mathrm{d}\boldsymbol{x}. \tag{2.4}$$

The expression is rewritten more compactly if we view H as an *operator* in $L^2(\mathbb{R}^d)$,

$$\big(H\psi\big)(\boldsymbol{x}) = -\frac{\hbar^2}{2m} \Delta\psi(\boldsymbol{x}) + V(\boldsymbol{x})\psi(\boldsymbol{x}) \tag{2.5}$$

and use the scalar product

$$\langle \psi, \varphi \rangle = \int_{\mathbb{R}^d} \overline{\psi(\boldsymbol{x})}\, \varphi(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}.$$

Eq. (2.4) becomes

$$\langle H \rangle = \langle \psi, H\psi \rangle.$$

Remark 2.3. Eq. (2.5) does not define the operator H completely because the domain $\mathcal{D}(H) \subset L^2(\mathbb{R}^d)$ is not specified. The domain must be such that the operator H is self-adjoint. Finding domains and proving self-adjointness is an important topic in functional analysis and mathematical quantum mechanics.

Differential operators and multiplication operators

The considerations on the Hamiltonian motivate the point of view adopted in quantum mechanics. Physical quantities, called *observables*, are represented by operators A. Expected values are evaluated by scalar product $\langle \psi, A\psi \rangle$ that involve the normalized wave function ψ.

Functions of position are associated with multiplication operators. For example, the potential $V(\boldsymbol{X})$ is associated with the operator

$$M_V : \ \mathcal{D}(M_V) \to L^2(\mathbb{R}^3), \ \psi \mapsto M_V\psi$$

with domain

$$\mathcal{D}(M_V) = \left\{ \psi \in L^2(\mathbb{R}^d) \ : \ \int_{\mathbb{R}^d} |V(\boldsymbol{x})|^2 \, |\psi(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} < \infty \right\}$$

given by

$$\big(M_V\psi\big)(\boldsymbol{x}) = V(\boldsymbol{x})\psi(\boldsymbol{x}).$$

Functions of momentum are associated with differential operators, as summarized in the formula

$$\boldsymbol{P} = \frac{\hbar}{\mathrm{i}} \nabla.$$

Differential operators in $\boldsymbol{x}$-space become multiplication operators in $\boldsymbol{k}$-space, i.e., after Fourier transform. The functions in the operator domains are not necessarily pointwise differentiable, instead derivatives are understood as weak derivatives, which the reader may know from the theory of Sobolev spaces.

For example, the j-th component P_j of the momentum is identified with the operator

$$P_j : \mathcal{D}(P_j) \to L^2(\mathbb{R}^3), \quad \psi \mapsto P_j\psi = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x_j}\psi$$

with domain

$$\mathcal{D}(P_j) = \left\{ \psi \in L^2(\mathbb{R}^d) \ : \ \int_{\mathbb{R}^d} |k_j|^2 \, |\widehat{\psi}(\boldsymbol{k})|^2 \, \mathrm{d}\boldsymbol{k} < \infty \right\}. \tag{2.6}$$

The choice of domain is motivated by the relations

$$\frac{\hbar}{\mathrm{i}}\frac{\partial\psi}{\partial x_j}(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hbar k_j \widehat{\psi}(\boldsymbol{k}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{k}$$

(remember (2.2)) and

$$\widehat{\frac{\hbar}{i}\frac{\partial\psi}{\partial x_j}}(\boldsymbol{k}) = \hbar k_j \widehat{\psi}(\boldsymbol{k}).$$

2.4 Hilbert space

The level of abstraction can be pushed a bit further. Instead of working with the concrete space $L^2(\mathbb{R}^d)$, we may work with a general Hilbert space $\mathcal{H}$. The scalar product $\langle \cdot, \cdot \rangle$ is taken to be conjugate-linear in the first entry and linear in the second entry. The normalized wave function is replaced by a unit vector

$$\psi \in \mathcal{H}, \quad ||\psi||^2 = \langle \psi, \psi \rangle = 1.$$

Observables are self-adjoint operators A in the Hilbert space $\mathcal{H}$. The expected value of an observable relative to ψ is

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle \quad \big(\psi \in \mathcal{D}(A)\big).$$

There are states ψ for which a given observable A has no well-defined expected value, much in the same way as some random variables have infinite variance or no expected value. The expected value $\langle A \rangle_\psi$ does not change if ψ is multiplied by a scalar $u \in \mathbb{C}$ with modulus $|u| = 1$:

$$\langle u\psi, A(u\psi) \rangle = \overline{u}u \langle \psi, A\psi \rangle = |u|^2 \, \langle \psi, A\psi \rangle = \langle \psi, A\psi \rangle. \tag{2.7}$$

Therefore ψ and $u\psi$ represent the same physical state of a system:

A state in quantum mechanics is a ray in Hilbert space.

Probability	Quantum mechanics
Probability distribution $\mathbb{P}$	Ray in Hilbert space, unit vector $\psi \in \mathcal{H}$
Real-valued random variable X	Self-adjoint operator A
Expected value $\mathbb{E}[X]$	$\langle \psi, A\psi \rangle$
$X = \lambda$, $\mathbb{P}$-almost surely	$A\psi = \lambda\psi$
$\lvert X \rvert \leq M$ everywhere	Bounded operator norm $\lvert\lvert A \rvert\rvert \leq M$

Table 2.1: Probability vs. quantum mechanics

The self-adjointness of A guarantees that the expected value is real because

$$\langle \psi, A\psi \rangle = \langle A^*\psi, \psi \rangle = \langle A\psi, \psi \rangle = \overline{\langle \psi, A\psi \rangle}$$

and $z = \overline{z}$ implies $\operatorname{Im} z = 0$. Eigenvectors play a special role: if

$$A\psi = \lambda\psi,$$

then the variance of A relative to ψ vanishes,

$$\langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 = \langle \psi, \lambda^2 \psi \rangle - \langle \psi, \lambda\psi \rangle^2 = 0.$$

For later purpose we introduce the set $B(\mathcal{H})$ of bounded operators on $\mathcal{H}$. An operator $A : \mathcal{D}(A) \to \mathcal{H}$ belongs to $B(\mathcal{H})$ if $\mathcal{D}(A) = \mathcal{H}$ and

$$||A\varphi|| \leq M||\varphi||$$

for all $\varphi \in \mathcal{H}$ and some constant $M \in [0, \infty)$. The smallest constant M is the *operator norm* of A,

$$||A|| = \sup_{\substack{\psi \in \mathcal{H}, \\ \psi \neq 0}} \frac{||A\psi||}{||\psi||}.$$

If A is a bounded observable, i.e., $A = A^*$ and $A \in B(\mathcal{H})$, then for all $\psi \in \mathcal{H}$,

$$\big|\langle \psi, A\psi \rangle\big| \leq ||A||\, ||\psi||^2 = ||A||$$

The expected value of A is bounded by a constant that does not depend on the state ψ. This is similar to bounded random variables: if a measurable map $X : \Omega \to \mathbb{R}$ on some measurable space $(\Omega, \mathcal{F})$ satisfies $|X(\omega)| \leq M$ for all $\omega \in \Omega$, then $|\mathbb{E}[X]| \leq M$ for every probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$.

2.5 Time evolution: Schrödinger picture

Time-dependent systems are modeled by time-dependent wave functions. For a particle with Hamilton operator (2.5), the time evolution is governed by a partial differential equation, the *Schrödinger equation*

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{x},t) = \Big(-\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{x})\Big)\psi(\boldsymbol{x},t) \tag{2.8}$$

compare Eq. (1.10). In the Hilbert space set-up, the time-dependent state of the system is given by a time-dependent unit vector $\psi(t) \in \mathcal{H}$ that solves

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi(t) = H\psi(t) \tag{2.9}$$

with H a self-adjoint operator, the *Hamilton operator*. It can be shown that if $\psi(0) \equiv \psi_0$ is in the domain of H, then there exists a uniquely defined map $t \mapsto \psi(t)$ from $\mathbb{R}$ to $\mathcal{H}$ that is norm-differentiable, stays in $\mathcal{D}(H)$, and satisfies (2.9) for all $t \in \mathbb{R}$. The solution is given by

$$\psi(t) = \exp\Big(-\mathrm{i}\frac{t}{\hbar}H\Big)\psi_0. \tag{2.10}$$

The operator

$$U_t = \exp\Big(-\mathrm{i}\frac{t}{\hbar}H\Big)$$

is defined using spectral calculus of self-adjoint operators. The right-hand side of (2.10) is well-defined for all time-zero vectors $\psi_0 \in \mathcal{H}$.

The family $(U_t)_{t\in\mathbb{R}}$ is a *strongly continuous unitary group*, which means the following.

(i) Each $U_t : \mathcal{H} \to \mathcal{H}$ is *unitary*, i.e., it is norm-preserving and bijective.

(ii) The family satisfies the group property

$$U_{t+s} = U_t U_s \quad (s,t \in \mathbb{R}).$$

(iii) For every $\varphi \in \mathcal{H}$ and all $t \in \mathbb{R}$,

$$\lim_{s\to t} ||U_s\varphi - U_t\varphi|| = 0.$$

Stone's theorem establishes a one-to-one correspondence between strongly continuous unitary groups and self-adjoint operators. The theorem is similar to the Hille-Yosida theorem for strongly continuous contraction semi-groups in Banach spaces, which enters the theory of a special class of continous-time Markov processes, the Feller-Dynkin processes. The unitary group $(U_t)_{t\in\mathbb{R}}$ plays a role analogous to the Markov semi-groups $(P_t)_{t\geq 0}$ and the Hamilton operator is similar to the infinitesimal generator L of the process. The analogies are summarized in Table 2.2.

Markov processes (Feller-Dynkin)	Schrödinger equation
Banach space	Hilbert space
Contraction semi-group $(P_t)_{t\geq 0}$	Unitary group $(U_t)_{t\in\mathbb{R}}$
Hille-Yosida theorem	Stone theorem
Infinitesimal generator L	Hamilton operator H
$P_t = \exp(-tL)$	$U_t = \exp(-\mathrm{i}tH/\hbar)$

Table 2.2: Markov processes vs. Schrödinger time evolution.

Eigenfunctions play again a special role. If

$$H\psi = E\psi$$

with $E \in \mathbb{R}$, then

$$\psi(t) = \mathrm{e}^{-\mathrm{i}tE/\hbar}\psi(0)$$

and the only effect of time evolution is to multiply the unit vector $\psi(0)$ with a complex number of modulus 1. As noted in Eq. (2.7), such a multiplication does not affect expected values and the resulting physical states are considered identical:

> Eigenfunctions of the Hamilton operator correspond to stationary states.

2.6 Time evolution: Heisenberg picture

Instead of letting the time evolution act on the state ψ, we can also let it act on observables. We introduce a family of maps on the set of bounded operators

$$\tau_t : \ B(\mathcal{H}) \to B(\mathcal{H}), \quad \tau_t(A) = U_t^* A U_t.$$

Notice that for all $\psi \in \mathcal{H}$,

$$\langle U_t\psi, AU_t\psi\rangle = \langle \psi, \tau_t(A)\psi\rangle.$$

For Markov processes, a loose analogue of the Schrödinger and Heisenberg picture is the action of the semi-group $(P_t)_{t\geq 0}$ on measures or functions. If $(X_t)_{t\geq 0}$ is a real-valued Markov process with initial law μ_0, defined on some underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then for every bounded measurable test function $f : \mathbb{R} \to \mathbb{R}$, the expected value of $f(X_t)$ can be expressed in two ways: we have

$$\mathbb{E}\big[f(X_t)\big] = \int_{\mathbb{R}} (P_t f)\,\mathrm{d}\mu_0 = \int_{\mathbb{R}} f\,\mathrm{d}\mu_t, \quad \mu_t = \mu_0 P_t.$$

See Table 2.3.

Markov processes	Quantum mechanics
Semi-group acts on functions	Heisenberg picture
$P_t f$	$\tau_t(A) = U_t^* A U_t$
Semi-group acts on measures	Schrödinger picture
$\mu_0 P_t$	$U_t \psi$

Table 2.3: Schrödinger vs. Heisenberg picture

The family $(\tau_t)_{t\in\mathbb{R}}$ satisfies the group property

$$\tau_{s+t}(A) = \tau_s\big(\tau_t(A)\big) \quad (s,t \in \mathbb{R},\ A \in B(\mathcal{H}))$$

or more succinctly, $\tau_{s+t} = \tau_s \circ \tau_t$. If the Hamilton operator is bounded, then for each $A \in B(\mathcal{H})$, the mapping $t \mapsto \tau_t(A)$ is differentiable in operator norm with derivative

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\tau_t(A) &= \Big(-\frac{\mathrm{i}}{\hbar}HU_t\Big)^* AU_t + U_t^* A\Big(-\frac{\mathrm{i}}{\hbar}HU_t\Big) \\ &= \frac{\mathrm{i}}{\hbar}(HU_t^* AU_t - U_t^* AU_t H) \\ &= \frac{\mathrm{i}}{\hbar}\big(H\tau_t(A) - \tau_t(A)H\big).\end{aligned}$$

The relation is rewritten in a compact form with the *commutator*

$$[A,B] = AB - BA$$

as

$$\frac{\mathrm{d}}{\mathrm{d}t}\tau_t(A) = \frac{\mathrm{i}}{\hbar}[H, \tau_t(A)].$$

If the Hamilton operator H is unbounded, then complications may arise because of domains.

2.7 Commutation relations

Canonical commutation relations

Let $\mathcal{H}$ be the Hilbert space of square-integrable, measurable functions $\psi : \mathbb{R}^d \to \mathbb{C}$. Let X_j be the multiplication operator

$$(X_j\psi)(\boldsymbol{x}) = x_j\psi(\boldsymbol{x})$$

with domain

$$\mathcal{D}(X_j) = \left\{\psi \in \mathcal{H} :\ \int_{\mathbb{R}^d} x_j^2 |\psi(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} < \infty\right\}.$$

The momentum operators are given by $P_j = \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x_j}$, $j = 1, \dots, d$ with domain (2.6). Then, for every smooth wave function ψ and all $\ell, j \in \{1, \dots, d\}$, we have

$$\frac{\partial}{\partial x_j}\big(x_\ell \psi(\boldsymbol{x})\big) = \delta_{j,\ell}\, \psi(\boldsymbol{x}) + x_\ell \frac{\partial}{\partial x_j} \psi(\boldsymbol{x})$$

with $\delta_{j,\ell}$ the Kronecker delta, hence $[P_j, X_\ell]\psi = \frac{\hbar}{\mathrm{i}} \delta_{j,\ell} \psi$ or or equivalently,

$$[X_\ell, P_j]\psi = \mathrm{i}\hbar \delta_{\ell,j} \psi$$

This holds true for all smooth wave functions ψ. Physicists write

$$[X_\ell, P_j] = \mathrm{i}\hbar \delta_{\ell,j} \tag{2.11}$$

where the right-hand side should be read as $\mathrm{i}\hbar\delta_{\ell,j}$ times the identity operator. The relations (2.11) with $j, \ell = 1, \dots, d$ are called *canonical commutation relations*.

From a mathematical point of view, commutators $[A, B]$ of unbounded operators $A, B \notin B(\mathcal{H})$ are dangerous objects. Section VIII.5 in the book [RS75] describes in detail some pitfalls. Unitary groups provide a way out.

Weyl relations

For $\boldsymbol{a}$, $\boldsymbol{b} \in \mathbb{R}^d$, define unitary operators $U(\boldsymbol{a})$ and $V(\boldsymbol{b})$ by

$$\big(U(\boldsymbol{a})\psi\big)(\boldsymbol{x}) = \exp\big(\mathrm{i}\boldsymbol{a} \cdot \boldsymbol{x}\big)\psi(\boldsymbol{x}), \quad \big(V(\boldsymbol{b})\psi\big)(\boldsymbol{x}) = \psi\big(\boldsymbol{x} + \hbar\boldsymbol{b}\big).$$

Notice that for smooth ψ,

$$\frac{\partial}{\partial t}\big(V(t\boldsymbol{b})\psi\big)(\boldsymbol{x})\bigg|_{t=0} = \hbar\big(\boldsymbol{b} \cdot \nabla\psi\big)(\boldsymbol{x}) = \mathrm{i}\big(\boldsymbol{b} \cdot \boldsymbol{P}\big)\psi(\boldsymbol{x}).$$

In fact the operators $\boldsymbol{b} \cdot \boldsymbol{P}$ and $\boldsymbol{a} \cdot \boldsymbol{X}$, with suitable domains, are self-adjoint and

$$U(\boldsymbol{a}) = \exp(\mathrm{i}\boldsymbol{a} \cdot \boldsymbol{X}), \quad V(\boldsymbol{b}) = \exp(\mathrm{i}\boldsymbol{b} \cdot \boldsymbol{P}).$$

We compute

$$\begin{aligned}\big(V(\boldsymbol{b})U(\boldsymbol{a})\psi\big)(\boldsymbol{x}) &= \exp\big(\mathrm{i}\boldsymbol{a} \cdot (\boldsymbol{x} + \hbar\boldsymbol{b})\big)\psi(\boldsymbol{x} + \hbar\boldsymbol{b}) \\ &= \exp(\mathrm{i}\hbar\boldsymbol{a} \cdot \boldsymbol{b})\big(U(\boldsymbol{a})V(\boldsymbol{b})\psi\big)(\boldsymbol{x})\end{aligned}$$

which yields the *Weyl relations*

$$U(\boldsymbol{a})V(\boldsymbol{b}) = \exp(-\mathrm{i}\hbar\boldsymbol{a} \cdot \boldsymbol{b})V(\boldsymbol{b})U(\boldsymbol{a}). \tag{2.12}$$

The canonical commutation relations (2.11) are recovered by taking partial derivatives with respect to a_ℓ and b_j. The reader may think of the Weyl relations as an exponential form of the canonical commutation relations. A notable advantage of the Weyl relations is that they involve only bounded operators.

Heisenberg group

Physicists often look for *representations* of commutation relations, i.e., families of matrices or operators that satisfy a given set of relations but may differ from the operators $\boldsymbol{a} \cdot \boldsymbol{X}$ and $\boldsymbol{b} \cdot \boldsymbol{P}$ with which we started.

The three-dimensional real *Heisenberg group* consists of the set of upper triangular 3×3 matrices with diagonal entries equal to 1 and real entries above the diagonal,

$$M(a,b,c) = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad a,\, b,\, c \in \mathbb{R}.$$

equipped with the usual matrix product. Define $U_a = M(a,0,0)$ and $V_b = M(0,b,0)$, and let $I = E_{1,3}$ be the matrix that has zeros everywhere except in the upper right corner. The matrix $E_{1,3}$ is not in the Heisenberg group but its exponential is. It is easily checked that $I^2 = 0$ and $\exp(cI) = M(0,0,c)$. In addition,

$$U_a V_b = \exp(abI) V_b U_a,$$

which bears a strong resemblance to the Weyl relations (2.12).

3 Path integrals

3.1 Laplacian, Brownian motion, Wiener measure

The Hamilton operator of a free particle in $L^2(\mathbb{R}^d)$ is given by

$$H_0 = -\frac{\hbar^2}{2m}\Delta$$

with domain

$$\mathcal{D}(H_0) = \left\{\psi \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d} |\boldsymbol{k}|^4 |\widehat{\psi}(\boldsymbol{k})|^2 \,\mathrm{d}\boldsymbol{k} < \infty\right\}. \tag{3.1}$$

We note

$$\widehat{H_0\psi}(\boldsymbol{k}) = \frac{\hbar^2}{2m}|\boldsymbol{k}|^2\widehat{\psi}(\boldsymbol{k}).$$

To simplify formulas, we choose units in such a way that

$$\hbar = 1, \quad m = 1$$

so that $H_0 = -\frac{1}{2}\Delta$. The reader familiar with Brownian motion may remember that $\frac{1}{2}\Delta$, with suitable domain, is precisely the infinitesimal generator of Brownian motion. The relation is best understood by looking at the family of operators

$$P_t = \exp\bigl(-tH_0\bigr), \quad t \geq 0$$

instead of $U_t = \exp(-\mathrm{i}tH_0)$. For $t > 0$ the operator P_t is an integral operator

$$\bigl(P_t f\bigr)(\boldsymbol{x}) = \int_{\mathbb{R}^d} p_t(\boldsymbol{x},\boldsymbol{y}) f(\boldsymbol{y}) \,\mathrm{d}\boldsymbol{y}$$

with kernel

$$p_t(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{(2\pi t)^{d/2}} \exp\Bigl(-\frac{|\boldsymbol{x}-\boldsymbol{y}|^2}{2t}\Bigr).$$

The kernel is non-negative $p_t \geq 0$ and satisfies

$$\int_{\mathbb{R}^d} p_t(\boldsymbol{x},\boldsymbol{y}) \,\mathrm{d}\boldsymbol{y} = 1$$

as well as the semi-group property

$$\int_{\mathbb{R}^d} p_t(\boldsymbol{x},\boldsymbol{z}) p_s(\boldsymbol{z},\boldsymbol{y}) \,\mathrm{d}\boldsymbol{z} = p_{s+t}(\boldsymbol{x},\boldsymbol{y}).$$

The probabilist will recognize the transition function of Brownian motion (or more precisely, the Radon-Nikodym derivative of the transition function with respect to Lebesgue measure). Mathematical physicists and analysts often prefer to phrase things with probability measures on spaces of paths rather than stochastic processes and random variables, and it is often enough to work on finite time horizons $[0, \beta]$. Let $C([0,\beta];\mathbb{R}^d)$ be the space of continous paths $\gamma : [0,\beta] \to \mathbb{R}^d$. The space is equipped with the topology of uniform convergence and associated Borel σ-algebra.

Theorem 3.1. *Fix $\beta > 0$. Then, for all $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^d$, there exists a uniquely defined measure $\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}$ on $C([0,\beta];\mathbb{R}^d)$ such that*

$$\begin{aligned}\int f\big(\gamma(0),\gamma(t_1),\dots,\gamma(t_n),\gamma(\beta)\big)\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}(\,\mathrm{d}\gamma)\\ = \int_{(\mathbb{R}^d)^n} f(\boldsymbol{x},\boldsymbol{q}_1,\dots,\boldsymbol{q}_n,\boldsymbol{y}) p_{t_1}(\boldsymbol{x},\boldsymbol{q}_1)p_{t_2-t_1}(\boldsymbol{q}_1,\boldsymbol{q}_2)\cdots p_{t_n-t_{n-1}}(\boldsymbol{q}_{n-1},\boldsymbol{q}_n)\\ \times p_{\beta-t_{n-1}}(\boldsymbol{q}_n,\boldsymbol{y})\,\mathrm{d}\boldsymbol{q}_1\cdots\mathrm{d}\boldsymbol{q}_n\end{aligned}$$

for all $n \in \mathbb{N}$, $0 < t_1 < \cdots < t_n < \beta$, and bounded measurable $f : (\mathbb{R}^d)^n \to \mathbb{R}$.

The measure has total mass

$$\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}\Big(C([0,\beta];\mathbb{R}^d)\Big) = p_\beta(\boldsymbol{x},\boldsymbol{y})$$

and charges only paths starting in $\boldsymbol{x}$ and ending in $\boldsymbol{y}$,

$$\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}\Big(\{\gamma :\, \gamma(0) \neq \boldsymbol{x} \text{ or } \gamma(\beta) \neq \boldsymbol{y}\}\Big) = 0.$$

The probability measure $(p_\beta(\boldsymbol{x},\boldsymbol{y}))^{-1}\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}$ is the distribution of a Brownian bridge; roughly, d-dimensional Brownian motion $(\boldsymbol{B}_t)_{t\in[0,\beta]}$ started in $\boldsymbol{x}$ and conditioned on $\boldsymbol{B}_\beta = \boldsymbol{y}$.

3.2 Feynman-Kac formula

The Feynman-Kac formula provides a representation of the integral kernel of $\exp(-t(H_0+V))$ for some external potentials V, in terms of integrals over paths space against the Brownian bridge measure $\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}$ from Theorem 3.1.

Let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable function and $H_0 = -\frac{1}{2}\Delta$ the operator with domain (3.1). By some abuse of notation we use the same letter V for the multiplication operator M_V. If V is bounded, then $H = H_0 + V$ with domain $\mathcal{D}(H) = \mathcal{D}(H_0)$ is self-adjoint.

Theorem 3.2. *If V is bounded, then for all $t > 0$, the operator $\exp(-t(H_0+V))$ is an integral operator with kernel*

$$k_t(\boldsymbol{x},\boldsymbol{y}) = \int \exp\Big(-\int_0^t V(\gamma(s))\,\mathrm{d}s\Big)\mu^{t}_{\boldsymbol{x},\boldsymbol{y}}(\,\mathrm{d}\gamma), \qquad (\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d).$$

The outer integral is over continuous paths $\gamma : [0,t] \to \mathbb{R}^d$ starting in $\gamma(0) = \boldsymbol{x}$ and ending in $\gamma(t) = \boldsymbol{y}$.

Probabilists might be more familiar with another form of the Feynman-Kac formula. Let

$$\big(\Omega, \mathscr{F}, (\mathbb{P}_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{R}^d}, (\boldsymbol{B}_s)_{s\geq 0}\big)$$

be a d-dimensional Brownian family, by which we mean the following:

(i) $(\Omega, \mathscr{F})$ is a measurable space.
(ii) Each $\mathbb{P}_{\boldsymbol{x}}$, $\boldsymbol{x} \in \mathbb{R}^d$, is a probability measure on $(\Omega, \mathscr{F})$.
(iii) Each $\boldsymbol{B}_s$, $s \geq 0$, is a measurable map from Ω to $\mathbb{R}^d$.
(iv) For every $\boldsymbol{x} \in \mathbb{R}^d$: $\mathbb{P}_{\boldsymbol{x}}(\boldsymbol{B}_0 = \boldsymbol{x}) = 1$ and for all $n \in \mathbb{N}$ and $0 < t_1 < \cdots < t_n$, the distribution of $(\boldsymbol{B}_{t_1}, \ldots, \boldsymbol{B}_{t_n})$ under $\mathbb{P}_{\boldsymbol{x}}$ has probability density function
$$p_{t_1}(\boldsymbol{x}; \boldsymbol{q}_1) p_{t_2-t_1}(\boldsymbol{q}_1, \boldsymbol{q}_2) \cdots p_{t_n - t_{n-1}}(\boldsymbol{q}_{n-1}, \boldsymbol{q}_n).$$
(v) The sample paths $s \mapsto \boldsymbol{B}_s(\omega)$ are continuous, $\mathbb{P}_{\boldsymbol{x}}$-almost surely, for every $\boldsymbol{x}$.

Let

$$u(\boldsymbol{x}, t) = \mathbb{E}_x\Big[\exp\Big(-\int_0^t V(\boldsymbol{B}_s)\,\mathrm{d}s\Big) u_0(\boldsymbol{B}_t)\Big].$$

Notice $u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x})$.

Theorem 3.3. *Under suitable assumptions on u_0 and V, the function $u(\boldsymbol{x}, t)$ satisfies the partial differential equation*

$$\frac{\partial}{\partial t} u = \frac{1}{2}\Delta u - Vu.$$

3.3 Discrete Laplacian. Continuous-time random walk

Feynman-Kac formulas exist as well for quantum particles on a lattice. Let $\ell \in \mathbb{N}$. The Hilbert space for a free quantum particle hopping on the lattice $\{1, \ldots, \ell\}$ is $\mathbb{C}^\ell$. Column vectors in $\mathbb{C}^\ell$ are identified with functions from $\{1, \ldots, \ell\}$ to $\mathbb{C}$. The discrete Laplacian with periodic boundary conditions acts on vectors $\varphi \in \mathbb{C}^\ell$ as

$$(L_0\varphi)(x) = \varphi(x+1) - 2\varphi(x) + \varphi(x-1). \tag{3.2}$$

with $\varphi(0) = \varphi(\ell)$, $\varphi(\ell+1) = \varphi(1)$. The linear map L_0 is given by the Hermitian matrix

$$L_0 = \big(L_0(x,y)\big)_{x,y=1,\ldots,\ell} = \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 & 1 \\ 1 & -2 & 1 & \cdots & 0 & 0 \\ & & \ddots & & & \\ 1 & 0 & 0 & \cdots & 1 & -2. \end{pmatrix}$$

The Hamiltonian of a free quantum particle, with periodic boundary conditions, is $H_0 = -\frac{1}{2}L_0$.

The matrix L_0 has non-negative off-diagonal matrix elements and row sums equal to zero. Such matrices are called *Q-matrices* in probability theory. Q-matrices are important because they generate semi-groups of stochastic matrices: if Q is a Q-matrix, then for each $t \geq 0$, the matrix exponential $P_t = \exp(tQ)$ is a *stochastic matrix*, i.e., it has non-negative matrix elements and row sums equal to zero. The semi-group $(P_t)_{t\geq 0}$ has the additional property

$$\lim_{t\to 0} P_t = \mathrm{id} = P_0,$$

sometimes called *standardness*. Such semi-groups $(P_t)_{t\geq 0}$ are associated with *Markov families* with *state space* $E = \{1, \ldots, \ell\}$ and *càdlàg sample paths*, i.e., families

$$\big(\Omega, \mathscr{F}, (\mathbb{P}_x)_{x\in E}, (X_t)_{t\geq 0}\big)$$

such that

(i) $(\Omega, \mathscr{F})$ is a measurable space.
(ii) Each $\mathbb{P}_x$, $x \in E$, is a probability measure on $(\Omega, \mathscr{F})$.
(iii) Each X_t, $t \geq 0$, is a measurable map from Ω to E.
(iv) For every $x \in E$: $\mathbb{P}_x(X_0 = x) = 1$ and

$$\mathbb{P}_x\big(X_{t_1} = x_1, \ldots, X_{t_n} = x_n\big) = P_{t_1}(x, x_1)P_{t_2-t_1}(x_1, x_2)\cdots P_{t_n-t_{n-1}}(x_{n-1}, x_n).$$

for all $n \in \mathbb{N}$, $0 < t_1 < \cdots < t_n$, and $x_1, \ldots, x_n \in E$.
(v) For every $x \in E$: the sample paths $t \mapsto X_t(\omega)$ are right-continuous and have left limits, $\mathbb{P}_x$-almost surely.

The process $(X_t)_{t\geq 0}$ associated with the Q-matrix $\frac{1}{2}L_0$ is a *continuous-time random walk* on $\{1, \ldots, \ell\}$ with periodic boundary conditions.

3.4 Lie-Trotter product formula

The exponential of complex numbers satisfies $\exp(a + b) = \exp(a)\exp(b)$. For matrices A and B that do not commute, it is no longer true that $\exp(A + B)$ is equal to $\exp(A)\exp(B)$, however the following is true.

Theorem 3.4 (Lie product formula)**.** *Let A and B be two $d \times d$ matrices. Then*

$$\exp(A + B) = \lim_{n\to\infty} \Big(\exp\big(\frac{1}{n}A\big)\exp\big(\frac{1}{n}B\big)\Big)^n.$$

Extensions to unbounded operators in Hilbert spaces are available, however conditions on the operators need to be imposed in order to be able to define an exponential, and the type of convergence has to be specified (convergence

in operator norm, in strong topology...). Infinite-dimensional versions of the Lie product formula are called *Trotter product formulas*.

Let us apply the theorem to $A = t\frac{1}{2}L_0$ with $t > 0$ and L_0 the discrete Laplacian with periodic boundary conditions from (3.2). The matrix B is chosen as $B = -tV$ with V a diagonal matrix

$$V = \begin{pmatrix} V(1) & 0 & \cdots & 0 \\ 0 & V(2) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & V(\ell) \end{pmatrix}, \quad V(1), \ldots, V(\ell) \in \mathbb{R}.$$

Diagonal matrices are the analogue in $\mathbb{C}^\ell$ of multiplication operators in $L^2(\mathbb{R}^d)$. Define

$$H = -\frac{1}{2}L_0 + V.$$

and

$$P_t^0 = \exp\Big(-\frac{t}{2}L_0\Big).$$

Then $-tH = A + B$ and

$$\mathrm{e}^{-tH} = \lim_{n\to\infty} \big(P_{t/n}^0 \mathrm{e}^{-tV/n}\big)^n.$$

The right-hand side has matrix elements

$$\big(P_{t/n}^0 \mathrm{e}^{-tV/n}\big)^n(x,y) = \sum_{x_1,\ldots,x_{n-1}} \prod_{k=1}^n P_{t/n}^0(x_{k-1}, x_k) \mathrm{e}^{-tV(x_k)/n}$$

with $x_0 = x$ and $x_n = y$. In terms of the associated Markov family,

$$\big(P_{t/n}^0 \mathrm{e}^{-tV/n}\big)^n(x,y) = \mathbb{E}_x\Big[\exp\Big(-\sum_{k=1}^n \frac{t}{n} V(X_{tk/n})\Big) \mathbb{1}_{\{X_t = y\}}\Big].$$

It follows that

$$\mathrm{e}^{-tH}(x,y) = \mathbb{E}_x\Big[\exp\Big(-\int_0^t V(X_s)\,\mathrm{d}s\Big) \mathbb{1}_{\{X_t = y\}}\Big],$$

which is again a Feynman-Kac formula. The Feynman-Kac formula and the irreducibility of the continuous-time random walk imply the strict positivity of matrix elements,

$$\mathrm{e}^{-tH}(x,y) > 0 \tag{3.3}$$

for all $t > 0$ and $x, y \in \{1, \ldots, \ell\}$.

3.5 Perron-Frobenius theorem. Ground states

Physicists are often interested in minimizing the expected energy $\langle \psi, H\psi \rangle$ for a given Hamilton operator H,

$$E_0 = \inf\big\{ \langle \psi, H\psi \rangle : \psi \in \mathcal{H},\, ||\psi|| = 1 \big\}.$$

In finite-dimensional Hilbert spaces, the infimum on the right-hand side is always a minimum and E_0 is the smallest eigenvalue, and every minimizer is necessarily an eigenvector. The minimum E_0 is the smallest eigenvector,

$$E_0 = \min\big\{ \lambda \in \mathbb{R} :\ \lambda \text{ is an eigenvalue of } H \big\},$$

and every minimizer ψ_0 is an eigenvector, $H\psi_0 = E_0\psi_0$. The minimum E_0 and minimizers ψ_0 are called *ground state energy* and *ground states*.

It is natural to ask whether the ground state is unique. The Feynman-Kac formula helps answer this question by bringing in the *Perron-Frobenius theorem*, a theorem from linear algebra about matrices with positive entries.

Perron-Frobenius theorem

We recall the definition of the *spectral radius* of a matrix A: it is given by

$$\rho(A) = \max\big\{ |\lambda| :\ \lambda \text{ is an eigenvalue of } A \big\}.$$

The adjoint of A is the matrix A^* with entries $A^*(x,y) = \overline{A(y,x)}$. For a matrix with real entries, the adjoint is the same as the transpose A^{T}.

Theorem 3.5. *Let $\ell \in \mathbb{N}$ and $A = (A(x,y))_{x,y=1,\dots,\ell}$ a matrix with strictly positive entries, $A(x,y) > 0$. Then the spectral radius $\rho(A)$ is an eigenvalue with algebraic multiplicity 1 of both A and A^*, and the associated eigenvectors can be chosen to have strictly positive entries.*

Thus if we write $\lambda_0 = \rho(A)$, then λ_0 is a simple eigenvalue of A and every other eigenvalue λ satisfies $|\lambda| < \lambda_0$. Moreover there exist strictly positive mappings $w, v : \{1, \dots, \ell\} \to (0, \infty)$ such that

$$\sum_{x=1}^{\ell} v(x)A(x,y) = \lambda_0 v(y), \quad \sum_{y=1}^{\ell} A(x,y)w(y) = \lambda_0 w(x).$$

We may associate with v a *row vector* $\boldsymbol{v}$ and with the mapping w a *column vector*, then we have the matrix equalities

$$\boldsymbol{v}A = \lambda_0 \boldsymbol{v}, \quad A\boldsymbol{w} = \lambda_0 \boldsymbol{w}.$$

The scalar λ_0 is the *principal eigenvalue*, the vectors $\boldsymbol{w}$ and $\boldsymbol{v}$ are called *right* and *left principal eigenvectors*.

In the context of Markov chains, the Perron-Frobenius theorem provides sufficient conditions for the existence of a unique invariant measure. Indeed if $A = P$ is a stochastic matrix with strictly positive entries, then the principal eigenvalue is $\lambda_0 = 1$, the right principal eigenvector is the vector $\boldsymbol{v} = \mathbf{1}$ that has all its entries equal to 1, and the left principal eigenvector $\boldsymbol{w} = \boldsymbol{\mu}$ is an invariant measure.

Non-degeneracy of ground states

Now we turn to ground states of Hermitian matrices H for which the semigroup $(\exp(-tH))_{t\geq 0}$ satisfies the positivity condition (3.3). This covers in particular the Hamiltonian $H = -\frac{1}{2}L_0 + V$ treated in Section 3.4.

Theorem 3.6. *Let H be a Hermitian $\ell\times\ell$ matrix ($H = H^*$). Suppose that the matrix $\exp(-tH)$ has strictly positive entries, for all $t > 0$. Then the smallest eigenvalue E_0 is a simple eigenvalue and there is an associated eigenvector with strictly positive entries.*

Put differently, the ground state ψ_0 is unique (up to multiplication by a scalar, $\psi_0 \to c\,\psi_0$) and strictly positive, $\psi_0(x) > 0$ for all x. The spirit of the argument carries over to infinite-dimensional spaces, however existence of ground states may fail and statements become more involved. See Chapter XIII.12 about non-degeneracy of ground states in [RS78].

Proof of Theorem 3.6. It is a standard result from linear algebra that every Hermitian matrix H is diagonalizable with an orthonormal basis of eigenvectors and it has only real eigenvalues. Let ψ_0 be an eigenvector of H for the eigenvalue E_0. Then

$$\mathrm{e}^{-tH}\psi_0 = \mathrm{e}^{-tE_0}\psi_0$$

for all $t > 0$ and ψ_0 is an eigenvector of $\exp(-tH)$.

On the other hand, the eigenvalues of $\exp(-tH)$ are precisely the exponentials $\exp(-t\lambda)$ of the eigenvalues $\lambda \in \mathbb{R}$ of H. Because of $\lambda \geq E_0$ for all eigenvalues, we have

$$0 \leq \exp(-t\lambda) \leq \exp(-tE_0)$$

and we see that $\exp(-tE_0)$ is equal to the spectral radius of $\exp(-tH)$. Let us fix some arbitrary $t > 0$, e.g., $t = 1$. Because of the strict positivity of the entries of $\exp(-H)$, we can apply the Perron-Frobenius theorem and deduce that the spectral radius $\exp(-E_0)$ is a simple eigenvalue and there exists an eigenvector vector $\boldsymbol{v}$ with strictly positive entries. The vector $\boldsymbol{v}$ spans the eigenspace so there exists a constant $c \in \mathbb{R}$ such that $\psi_0 = c\,\boldsymbol{v}$. This proves that the eigenspace $\mathcal{N}(H - E_0\,\mathrm{id})$ is one-dimensional and spanned by the vector $\boldsymbol{v}$, which has strictly positive entries. □

Ground state transformation

Theorem 3.6 allows us to transform the semi-group $(\exp(-tH))_{t\geq 0}$ of matrices with strictly positive matrix elements into a proper Markov semi-group. Let ψ_0 be the unique eigenvector with strictly positive entries that is normalized, $\sum_{x=1}^{\ell} \psi_0(x)^2 = 1$. Write $H - E_0$ instead of $H - E_0\,\mathrm{id}$, where "id" is the identity matrix. Set

$$P_t(x,y) = \frac{1}{\psi_0(x)}\,\mathrm{e}^{-t(H-E_0)}(x,y)\psi_0(y), \quad \mu(x) = \psi_0(x)^2. \tag{3.4}$$

Then

$$P_t(x,y) > 0, \quad \sum_{y=1}^{\ell} P_t(x,y) = 1$$

and

$$\sum_{x=1}^{\ell} \mu(x) = 1, \quad \sum_{x=1}^{\ell} \mu(x) P_t(x,y) = \mu(y). \tag{3.5}$$

Thus $(P_t)_{t\geq 0}$ is a semi-group of stochastic matrices with invariant measure μ. Moreover μ and P_t satisfy the *detailed balance equation*

$$\mu(x)P_t(x,y) = \psi_0(x)\mathrm{e}^{-t(H-E_0)}(x,y)\psi_0(y) = \mu(y)P_t(y,x).$$

and the measure μ is *reversible*. The ground state transformation is discussed by Bakry, Gentil and Ledoux [BGL14, Section 1.15.6], it is closely related to the h-transform of Markov processes.

3.6 Harmonic oscillator and Ornstein-Uhlenbeck process

The quantum harmonic oscillator is an important reference model in quantum mechanics. The Ornstein-Uhlenbeck process is a Markov process and stochastic diffusion that models systems with Brownian noise and a linear drift towards equilibrium. The ground state transformation explained for discrete systems in the previous section shows that the harmonic oscillator and Ornstein-Uhlenbeck process are deeply related.

In this section we bring back in the mass $m > 0$ of a particle and the reduced Planck constant $\hbar > 0$.

Harmonic oscillator

The Hamilton operator of the *harmonic oscillator* with angular frequency $\omega_0 > 0$ acts on smooth functions $\psi : \mathbb{R} \to \mathbb{C}$ as

$$H\psi(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}(x) + \frac{1}{2}m\omega_0^2 x^2\psi(x). \tag{3.6}$$

The operator H with suitable domain $\mathcal{D}(H) \subset L^2(\mathbb{R})$ is self-adjoint. The name *harmonic oscillator* reflects the behavior of the associated classical system: For the Hamilton function $H(p,x) = \frac{1}{2m}p^2 + \frac{1}{2}\omega_0^2 x^2$, the differential equation (1.3) reads $\dot{p} = -\partial_x H = -\omega_0^2 x$, $\dot{x} = \partial_p H = p$, which gives

$$m\ddot{x}(t) + m\omega_0^2 x(t) = 0.$$

The solutions are given by linear combinations of $\cos(\omega_0 t)$ and $\sin(\omega_0 t)$. The differential equation models the behavior of oscillating systems (think of a pendulum).

The spectrum of the Hamilton operator H consists of a discrete countably infinite set of simple eigenvalues, the eigenvectors form a *complete orthonormal system*. Remember that a family $(e_n)_{n\in\mathbb{N}_0}$ is a complete orthonormal system if $\langle e_n, e_m\rangle = \delta_{n,m}$ and the set of finite linear combinations $\sum_k x_k e_{n_k}$ is dense in $\mathcal{H}$.

Theorem 3.7. *The Hilbert space has a complete orthonormal system $(\psi_n)_{n\in\mathbb{N}_0}$ that consists of eigenfunctions of H, with*

$$H\psi_n = \hbar\omega_0\Big(n + \frac{1}{2}\Big)\psi_n \qquad (n \in \mathbb{N}_0).$$

An explicit computation shows that

$$\psi_0(x) = \frac{1}{(2\pi/[2m\omega_0/\hbar])^{1/4}} \exp\Big(-\frac{1}{2}\frac{m\omega_0}{\hbar}x^2\Big) \tag{3.7}$$

is a strictly positive, normalized eigenfunction for the smallest eigenvalue $E_0 = \hbar\omega_0/2$. More generally, ψ_n is proportional to a polynomial of degree n times ψ_0, see for example [Far10, Section 2.10].

Ornstein-Uhlenbeck process

From now on we go back to $\hbar = 1$ and $m = 1$. The following computation is motivated by strict positivity of the eigenfunction (3.7) and the ground state transformation (3.4). Let $f \in C_c^\infty(\mathbb{R})$, i.e., $f : \mathbb{R} \to \mathbb{C}$ is smooth with compact support. Then

$$\begin{aligned}
&\frac{1}{\psi_0(x)}\Big((H - E_0)(\psi_0 f)\Big)(x) \\
&\quad= \frac{1}{\psi_0(x)}\Big\{\Big((H - E_0)\psi_0(x)\Big)f(x) - \frac{\hbar^2}{m}\psi_0'(x)f'(x) - \frac{\hbar^2}{2m}\psi_0(x)f''(x)\Big\} \\
&\quad= -\frac{\hbar^2}{2m}f''(x) + \hbar\omega_0 x f'(x) = -Lf(x)
\end{aligned} \tag{3.8}$$

with

$$Lf(x) = \frac{1}{2}\sigma^2 f''(x) - \theta x f'(x), \quad \sigma^2 = \frac{\hbar^2}{m}, \quad \theta = \hbar\omega_0. \tag{3.9}$$

The operator L with suitable domain $\mathcal{D}(L)$, as an operator in the Banach space $C_0(\mathbb{R})$ of continuous functions vanishing at infinity, is the generator of a continuous-time Markov process $(X_t)_{t\geq 0}$, the *Ornstein-Uhlenbeck process*. The Ornstein-Uhlenbeck process satisfies a stochastic differential equation with linear drift and Brownian noise,

$$\mathrm{d}X_t = -\theta X_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t.$$

The Ornstein-Uhlenbeck process has a unique stationary distribution μ, the normal law $\mathcal{N}(0, \sigma^2/[2\theta])$ with mean zero and variance $\sigma^2/(2\theta)$. For our choice $\sigma = 1$ and $\theta = \omega_0$, the variance becomes

$$\frac{\sigma^2}{2\theta} = \frac{\hbar}{2m\omega_0}$$

and we recognize that $\psi_0(x)^2$ with $\psi_0(x)$ given in (3.7) is precisely the probability density of the invariant measure μ. This is similar to the discrete relations (3.4) and (3.5).

Unitary equivalence

The somewhat computational relations sketched above can be recast as a more structural statement using unitary equivalence of Hilbert spaces. Let ψ_0 be the normalized positive ground state from (3.7) and $\mu = \mathcal{N}(0, 1/(2\omega_0))$ the normal distribution with mean zero and variance $1/(2\omega_0)$.

In addition to the Hilbert space $\mathcal{H} = L^2(\mathbb{R}) \equiv L^2(\mathbb{R}, \mathrm{Leb})$ of complex-valued measurable functions with norm

$$||\psi||_{\mathcal{H}}^2 = \int_{-\infty}^{\infty} |\psi(x)|^2 \,\mathrm{d}x,$$

we introduced the Hilbert space $\mathfrak{h} = L^2(\mathbb{R}, \mu)$ with norm

$$||f||_{\mathfrak{h}}^2 = \int_{-\infty}^{\infty} |f(x)|^2 \mu(\,\mathrm{d}x) = \int_{-\infty}^{\infty} |f(x)|^2 \psi_0(x)^2 \,\mathrm{d}x.$$

The linear map

$$U : \mathfrak{h} \to \mathcal{H}, \quad f \mapsto \psi_0 f$$

is norm-preserving

$$||Uf||_{\mathcal{H}}^2 = ||\psi_0 f||_{\mathcal{H}}^2 = ||f||_{\mathfrak{h}}^2$$

and bijective with norm-preserving inverse

$$U^{-1} : \mathcal{H} \to \mathfrak{h}, \ \psi \mapsto \frac{1}{\psi_0}\psi.$$

Thus U is a unitary isomorphism between $\mathcal{H}$ and $\mathfrak{h}$. Disregarding issues with domains etc. we have

$$U^{-1}(H - E_0)U = -L.$$

3.7 Action functional

We conclude with some additional vocabulary and complementary points of view. They are helpful for some aspects of quantum field theory but can be skipped on first reading.

Path integrals: heuristic formulas

Let us go back to the Wiener measure or more precisely, the Brownian bridge measures $\mu^{\beta}_{\boldsymbol{x},\boldsymbol{y}}$ from Theorem 3.1. Physicists tend to think of the measure $\mu^{t}_{\boldsymbol{x},\boldsymbol{y}}$ as a measure on the set $\Gamma^{t}_{\boldsymbol{x},\boldsymbol{y}}$ of continuous paths $\gamma : [0,t] \to \mathbb{R}^d$ with $\gamma(0) = \boldsymbol{x}$ and $\gamma(t) = 0$ that is absolutely continuous with respect to some fictitious analogue of Lebesgue measure,

$$\mu^{t}_{\boldsymbol{x},\boldsymbol{y}}(A) \propto \int_{\Gamma^{t}_{\boldsymbol{x},\boldsymbol{y}}} \mathbb{1}_A(\gamma) \exp\left(-\frac{1}{2}\int_0^t |\dot\gamma(s)|^2 \,\mathrm{d}s\right) \mathcal{D}[\gamma]. \tag{3.10}$$

Mathematicians view this expression as highly problematic because the infinite-dimensional Lebesgue measure $\mathcal{D}[\gamma]$ is not defined and sample paths of Brownian motion are not differentiable. The heuristic formula (3.10) exponential is motivated by a correct formula on the kernel $p_t(\boldsymbol{x},\boldsymbol{y})$

$$p_{t/n}(x_0,x_1)p_{t/n}(x_1,x_2)\cdots p_{t/n}(x_{n-1},x_n) \\ = \frac{1}{(2\pi t/n)^{d/2}} \exp\left(-\sum_{i=1}^{n} \frac{|x_i - x_{i-1}|^2}{2t/n}\right).$$

The exponent looks like a Riemann sum for the integral $\int_0^t |\dot\gamma(s)|^2 \,\mathrm{d}s$ with $\gamma(jt/n) = x_j$,

$$\sum_{j=1}^{n} \frac{|x_j - x_{j-1}|^2}{2t/n} = \frac{1}{2}\sum_{j=1}^{n} \frac{t}{n}\Big|\frac{x_j - x_{j-1}}{t/n}\Big|^2 \approx \frac{1}{2}\int_0^t |\dot\gamma(s)|^2 \,\mathrm{d}s.$$

Let's brush our concerns aside and insert the problematic formula (3.10) into the Feynman-Kac formula. Then we get another ill-defined formula,

$$k_t(\boldsymbol{x},\boldsymbol{y}) \propto \int_{\Gamma^{t}_{\boldsymbol{x},\boldsymbol{y}}} \exp\left(-\int_0^t \Big\{\frac{1}{2}m|\dot\gamma(s)|^2 + V(\gamma(s))\Big\}\,\mathrm{d}s\right)\mathcal{D}[\gamma]$$

for the integral kernel of $\exp(-tH)$.

A similar ill-defined formula is derived heuristically for the integral kernel of the unitary $U_t = \exp(-\mathrm{i}tH/\hbar)$, with the help of the Lie-Trotter product formula and an explicit formula for the integral kernel of $\exp(-\mathrm{i}H_0 t/m)$ where $H_0 = -\frac{\hbar^2}{2m}\Delta$. The heuristic formula is

$$\mathrm{e}^{-\mathrm{i}tH/\hbar}(\boldsymbol{x},\boldsymbol{y}) \propto \int_{\Gamma^{t}_{\boldsymbol{x},\boldsymbol{y}}} \exp\left(\frac{\mathrm{i}}{\hbar}\int_0^t \Big\{\frac{1}{2}m|\dot\gamma(s)|^2 - V(\gamma(s))\Big\}\,\mathrm{d}s\right)\mathcal{D}[\gamma].$$

This is the *Feynman path integral* representation of the *propagator* in physics.

Action functional and Lagrange formulation of classical mechanics

The function

$$L(\boldsymbol{x}, \boldsymbol{v}) = \frac{1}{2} m |\boldsymbol{v}|^2 - V(\boldsymbol{x})$$

is the *Lagrange function*, the functional $\gamma \mapsto S(\gamma)$ given by

$$S(\gamma) = \int_0^t \Big\{ \frac{1}{2} m |\dot\gamma(s)|^2 - V(\gamma(s)) \Big\} \, \mathrm{d}s = \int_0^t L\big(\gamma(s), \dot\gamma(s)\big) \, \mathrm{d}s$$

is the *action functional*. The action functional enters the *Lagrangian formulation* of classical mechanics: The equation of motion corresponds to critical points of the action functional, i.e., points for which

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} S(\gamma + \varepsilon h)\big|_{\varepsilon=0} = 0$$

for all smooth variations $h : [0,t] \to \mathbb{R}^d$ with $h(0) = h(t) = 0$. Assuming that integrals and derivatives can be exchanged, we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_0^t L\big(\gamma(s) + \varepsilon h(s), \dot\gamma(s) + \varepsilon \dot h(s)\big) \, \mathrm{d}s \Big|_{\varepsilon=0} \\
&\quad = \int_0^t \Big\{ h(s) \cdot \big(\nabla_{\boldsymbol{x}} L\big)(\gamma(s), \dot\gamma(s)) + \dot h(s) \cdot \big(\nabla_{\boldsymbol{v}} L\big)(\gamma(s), \dot\gamma(s)) \Big\} \, \mathrm{d}s \\
&\quad = \int_0^t h(s) \cdot \Big\{ \big(\nabla_{\boldsymbol{x}} L\big)(\gamma(s), \dot\gamma(s)) - \frac{\mathrm{d}}{\mathrm{d}s} \big(\nabla_{\boldsymbol{v}} L\big)(\gamma(s), \dot\gamma(s)) \Big\} \, \mathrm{d}s
\end{aligned}$$

The term in braces is equal to

$$-\nabla V(\gamma(s)) - \frac{\mathrm{d}}{\mathrm{d}s} m \dot\gamma(s) = -\nabla V(\gamma(s)) - m \ddot\gamma(s),$$

it vanishes for solutions of $m\ddot\gamma = -\nabla V(\gamma)$. Action functionals on path space play a role in probability as well, the natural context is large deviations of stochastic processes [FW84, FK06].

Propagators, imaginary time, Wick rotation

The operator $U_t = \exp(-\mathrm{i}tH/\hbar)$ and its integral kernel are called *propagator* in physics. The operator $P_t = \exp(-tH/\hbar)$ from the semi-group $(P_t)_{t \geq 0}$ is called *imaginary-time propagator* because of the formal relation $P_t = U_{t/\mathrm{i}}$, or $U_t = P_{\mathrm{i}t}$. Switching back and forth between t and $\mathrm{i}t$ is a useful device, especially when aiming for a probabilistic interpretation of various formulas. The device sometimes goes under the name of *Wick rotation*—notice that the linear map $z \mapsto \mathrm{i}z$ in the complex plane is a rotation with angle $\pi/2$.

4 Harmonic oscillator

Setting all parameters $m, \hbar, \omega_0, \sigma^2, \theta$ to 1, the Hamilton operator (3.6) and the generator (3.9) of the Ornstein-Uhlenbeck process become

$$H = -\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}x^2, \quad L = \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - x\frac{\mathrm{d}}{\mathrm{d}x}.$$

The eigenfunction from (3.7) becomes $\psi_0(x) = \pi^{-1/4}\exp(-x^2/2)$ and the relation (3.8) reads

$$-Lf = \frac{1}{\psi_0}\Big(H - \frac{1}{2}\Big)(f\psi_0). \tag{4.1}$$

We view H as an operator in $L^2(\mathbb{R}) = L^2(\mathbb{R}, \mathrm{Leb})$ and L as an operator in $L^2(\mathbb{R}, \mu_0)$ with $\mu_0(\mathrm{d}x) = \pi^{-1/2}\exp(-x^2)\mathrm{d}x$. The precise choice of domain is left open. For now we only ask that $\mathcal{D}(L)$ contains all polynomials and $\mathcal{D}(H)$ contains all functions $f : \mathbb{R} \to \mathbb{R}$ with

$$f \in C^\infty(\mathbb{R}), \quad \forall m, k \in \mathbb{N}_0 : \sup_{x\in\mathbb{R}}\Big|x^k\frac{\mathrm{d}^m f}{\mathrm{d}x^m}(x)\Big| < \infty.$$

The space of such functions is the *Schwartz space* $\mathcal{S}(\mathbb{R})$.

4.1 Hermite polynomials, Hermite functions

The Ornstein-Uhlenbeck operator L and the Hamilton operator H can be diagonalized explicitly, the eigenfunctions are given by *Hermite polynomials* and *Hermite functions*. The n-th physicist's *Hermite polynomial* is the polynomial

$$H_n(x) = (-1)^n \mathrm{e}^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathrm{e}^{-x^2}.$$

(The probabilist's preferred choice is $\exp(\pm x^2/2)$ instead of $\exp(\pm x^2)$.) The Hermite polynomials satisfy the recurrence relation

$$H_0(x) = 1, \quad H_n(x) = 2xH_{n-1}(x) - H'_{n-1}(x).$$

We note without proof that

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} H_n(x)H_m(x)\mathrm{e}^{-x^2}\,\mathrm{d}x = 2^n n!\delta_{n,m},$$

in particular the polynomials are *orthogonal* with respect to the Gaussian measure $\mu_0(\mathrm{d}x)$. The n-th *Hermite function* is

$$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x) \mathrm{e}^{-x^2/2} = \frac{1}{\sqrt{2^n n!}} H_n(x) \psi_0(x). \tag{4.2}$$

The Hermite functions form a complete orthonormal system in $L^2(\mathbb{R})$.

Proposition 4.1. *We have $LH_n = -nH_n$, for all $n \in \mathbb{N}_0$.*

Proof. We show first, by induction over $n \in \mathbb{N}_0$, that

$$H_0' = 0, \quad H_n' = 2nH_{n-1} \ (n \geq 1).$$

For $n = 0$, this follows from $H_0 \equiv 1$. For the induction step, we use the recurrence relation and the induction hypothesis, which yields

$$\begin{aligned} H_n'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\big(2xH_{n-1}(x) - H_{n-1}'(x)\big) \\ &= 2H_{n-1}(x) + 2xH_{n-1}'(x) - 2(n-1)H_{n-2}'(x) \\ &= 2H_{n-1}(x) + 2(n-1)\big(2xH_{n-2}(x) - H_{n-2}'(x)\big) \\ &= 2nH_{n-1}(x). \end{aligned}$$

The induction is complete. Next we note that for all smooth f,

$$Lf(x) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}\big(f'(x) - 2xf(x)\big)$$

which gives

$$LH_n = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}\big(-H_{n+1}\big) = -nH_n. \qquad \square$$

Combining the proposition with the similarity relation (4.1), we obtain right away

$$H\psi_n = \big(n + \frac{1}{2}\big)\psi_n \quad (n \in \mathbb{N}_0) \tag{4.3}$$

and the Hermite functions form a complete orthonormal system of eigenfunctions of H. Physicists arrive at this equation slightly differently, with the help of *creation and annihilation* or *raising and lowering* operators.

4.2 Creation and annihilation operators

Let X and $P = \frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}$ be the position and momentum operators in $L^2(\mathbb{R})$. Then for all $\psi \in C_c^\infty(\mathbb{R})$,

$$H\psi = \frac{1}{2}(P^2 + X^2)\psi.$$

If P and X were real numbers rather than operators, the operator H would look like the modulus of a complex number:

$$x^2 + y^2 = |x + \mathrm{i}y|^2 = (x - \mathrm{i}y)(x + \mathrm{i}y) \quad (x, y \in \mathbb{R}). \tag{4.4}$$

Define the formal operators

$$A = \frac{1}{\sqrt{2}}(X + \mathrm{i}P), \quad A^\dagger = \frac{1}{\sqrt{2}}(X - \mathrm{i}P)$$

i.e.

$$A = \frac{1}{\sqrt{2}}\Big(x + \frac{\mathrm{d}}{\mathrm{d}x}\Big), \quad A^\dagger = \frac{1}{\sqrt{2}}\Big(x - \frac{\mathrm{d}}{\mathrm{d}x}\Big).$$

The operators are merely formal because we haven't specified the domains. We carry on with formal computations disregarding questions of domains. For operators, the analogue of the identity (4.4) is no longer true. Instead we have

$$(X - \mathrm{i}P)(X + \mathrm{i}P) = X^2 + P^2 + \mathrm{i}[X, P].$$

Because of the canonical commutation relation $[X, P] = \mathrm{i} = \mathrm{i}\,\mathrm{id}$, this yields

$$X^2 + P^2 = (X - \mathrm{i}P)(X + \mathrm{i}P) + 1$$

and

$$H = A^\dagger A + \frac{1}{2}.$$

We also note the relation

$$[A, A^\dagger]\psi = \psi,$$

valid for all $\psi \in \mathcal{S}(\mathbb{R})$.

Proposition 4.2. *We have $\langle \psi, A\varphi\rangle = \langle A^\dagger\psi, \varphi\rangle$ for all $\varphi, \psi \in \mathcal{S}(\mathbb{R})$, and*

$$A(A^\dagger)^n\psi_0 = n(A^\dagger)^{n-1}\psi_0, \quad H(A^\dagger)^n\psi_0 = \Big(n + \frac{1}{2}\Big)(A^\dagger)^n\psi_0 \tag{4.5}$$

as well as

$$\langle (A^\dagger)^m\psi_0, (A^\dagger)^n\psi_0\rangle = n!\delta_{n,m}.$$

Proof sketch. A straightforward integration by parts shows $\langle \psi, \varphi'\rangle = -\langle \psi', \varphi\rangle$ and the first identity follows since A and $A^\dagger$ differ only by the sign in front of the derivative. Next,

$$A\psi_0 = \frac{1}{\sqrt{2}}\big(x + \frac{\mathrm{d}}{\mathrm{d}x}\big)\pi^{-1/4}\mathrm{e}^{-x^2/2} = 0. \tag{4.6}$$

The rest of the proof is mainly algebraic and based on the relation $(AA^\dagger - A^\dagger A)\psi = \psi$. For example, the first equality in (4.5) is proven by induction. For $n = 0$, it is precisely Eq. (4.6). For the induction step, we compute

$$\begin{aligned} A(A^\dagger)^n\psi_0 &= (A^\dagger A)(A^\dagger)^{n-1}\psi_0 + (A^\dagger)^{n-1}\psi_0 \\ &= A^\dagger(n-1)(A^\dagger)^{n-2}\psi_0 + (A^\dagger)^{n-1}\psi_0 = n(A^\dagger)^{n-1}\psi_0. \end{aligned}$$ □

The eigenfunction equation (4.3) is recovered with the following proposition.

Proposition 4.3. *The Hermite functions from* (4.2) *satisfy*

$$\psi_n = \frac{1}{\sqrt{n!}}(A^\dagger)^n \psi_0 \qquad (n \in \mathbb{N}_0).$$

Proof. For $n = 0$, there is nothing to prove. For $n \geq 1$, we note

$$\begin{aligned}\frac{1}{\sqrt{n!}}(A^\dagger)^n \psi_0 &= \frac{(-1)^n}{\sqrt{2^n n!}}\, \mathrm{e}^{x^2/2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2/2} \psi_0 \\ &= \frac{(-1)^n}{\sqrt{2^n n! \sqrt{\pi}}}\, \mathrm{e}^{-x^2/2} \Big(\mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \Big) \\ &= \frac{1}{\sqrt{2^n n!}}\, \psi_0(x) H_n(x) = \psi_n(x). \end{aligned}$$

□

Combining the previous two propositions, we obtain

$$A^\dagger \psi_n = \sqrt{n+1}\, \psi_{n+1}, \quad A\psi_n = \sqrt{n}\, \psi_{n-1}.$$

The operator A lowers the index of the basis function ψ_n, much in the same way as a derivative maps a polynomial of degree n to a polynomial of degree $n-1$. The operators $A^\dagger$ and A are called *raising* and *lowering* operators, also *creation* and *annihilation* operators.

4.3 Birth and death process. Poisson-Charlier polynomials

Creation and annihilation sounds a bit like birth and death. Can we cash in on the analogy in a probabilistically meaningful way? The short answer is: yes! The precise answer involves a birth and death process. Gaussian measures are replaced with Poisson distributions and Hermite polynomials with Poisson-Charlier polynomials.

Birth and death process

A *birth and death process* is a continous-time Markov process $(N_t)_{t\geq 0}$ with state space $\mathbb{N}_0$ and transitions $n \to n \pm 1$. We consider processes with linear death rates and constant birth rate $\lambda > 0$. The Q-matrix acts on functions $f : \mathbb{N}_0 \to \mathbb{C}$ as

$$(Qf)(n) = \lambda\big(f(n+1) - f(n)\big) + n\big(f(n-1) - f(n)\big).$$

The Poisson distribution

$$p_\lambda(n) = \frac{\lambda^n}{n!} \mathrm{e}^{-\lambda}$$

is a reversible measure for the process, hence Q is a symmetric operator in the weighted ℓ^2-space $\ell^2(\mathbb{N}_0, p_\lambda)$. The scalar product in the space is

$$\langle f, g\rangle_\lambda = \sum_{n\in\mathbb{N}_0} \overline{f(n)}g(n)\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}.$$

The symmetry relation for Q reads

$$\langle f, Qg\rangle_\lambda = \langle Qf, g\rangle_\lambda.$$

It is valid for all $f, g : \mathbb{N}_0 \to \mathbb{C}$ that go to zero sufficiently fast at infinity.

The following theorem says that Q is unitarily equivalent to the Ornstein-Uhlenbeck generator L.

Theorem 4.4. *There exists a unitary operator*

$$U : \ L^2\left(\mathbb{R}, \frac{1}{\sqrt{\pi}}\,\mathrm{e}^{-x^2}\,\mathrm{d}x\right) \to \ell^2(\mathbb{N}_0, p_\lambda),$$

such that L and Q, with suitable domains, satisfy

$$Q = U^{-1}LU.$$

The proof builds on an explicit diagonalization of Q.

Remark 4.5. By Theorem 4.4 and the unitary equivalence of $-L$ and $H - \frac{1}{2}$, the operator $-Q$ is also unitarily equivalent to $H - \frac{1}{2}$. Thus we may view the Ornstein-Uhlenbeck process and the birth-and-death process as two different probabilistic incarnations of one and the same quantum object, the harmonic oscillator.

Poisson-Charlier polynomials

Set

$$(\mathsf{D}f)(n) = f(n+1) - f(n), \quad (\delta f)(n) = \frac{n}{\lambda}f(n-1) - f(n).$$

A straightforward computation yields

$$\langle f, Dg\rangle_\lambda = \langle \delta f, g\rangle_\lambda$$

for all $f, g : \mathbb{N}_0 \to \mathbb{C}$ that go to zero sufficiently fast at infinity. In addition,

$$\frac{1}{\lambda}(Qf)(n) = -(\delta\mathsf{D}f)(n).$$

We define a sequence $(\mathscr{C}_k)_{k\in\mathbb{N}_0}$ of functions from $\mathbb{N}_0$ to $\mathbb{R}$ by $\mathscr{C}_k = \delta^k\mathbf{1}$. Equivalently, $\mathscr{C}_0(n) \equiv 1$ and

$$\mathscr{C}_k(n) = \frac{n}{\lambda}\mathscr{C}_{k-1}(n-1) - \mathscr{C}_{k-1}(n) \quad (n \in \mathbb{N}_0).$$

An induction over k shows that $\mathscr{C}_k$ is a polynomial of degree k with highest-order term n^k/λ^k. The functions $\mathscr{C}_k(n)$ are called *Charlier* or *Poisson-Charlier* polynomials. They satisfy the orthogonality relation

$$\langle \mathscr{C}_k, \mathscr{C}_j \rangle_\lambda = \sum_{n=0}^{\infty} \mathscr{C}_k(n)\mathscr{C}_j(n)\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} = \frac{k!}{\lambda^k}\,\delta_{k,j} \tag{4.7}$$

see [PT11, Chapter 10] for details and further references.

Proposition 4.6. *We have* $Q\mathscr{C}_k = -k\mathscr{C}_k$.

Proof. We show first

$$\mathsf{D}\mathscr{C}_k = \frac{k}{\lambda}\mathscr{C}_{k-1}.$$

The proof is by induction over k. For $k = 0$, we have $\mathsf{D}\mathscr{C}_0 = \mathsf{D}\mathbf{1} = 0$. For the induction step, we apply the recurrence relation and the induction hypothesis and obtain

$$\begin{aligned}\mathscr{C}_k(n+1) - \mathscr{C}_k(n) &= \frac{1}{\lambda}\mathscr{C}_{k-1}(n) + \frac{n}{\lambda}\big(\mathscr{C}_{k-1}(n) - \mathscr{C}_{k-1}(n-1)\big) \\ &\qquad - \big(\mathscr{C}_{k-1}(n+1) - \mathscr{C}_{k-1}(n)\big) \\ &= \frac{1}{\lambda}\mathscr{C}_{k-1}(n) + \frac{k-1}{\lambda}(\delta\mathscr{C}_{k-2})(n) \\ &= \frac{k}{\lambda}\mathscr{C}_{k-1}(n).\end{aligned}$$

This completes the induction. To conclude we observe

$$Q\mathscr{C}_k = -\lambda\delta\mathsf{D}\mathscr{C}_k = -k\delta\mathscr{C}_{k-1} = -k\mathscr{C}_k. \qquad \square$$

The unitary U in Theorem 4.4 maps normalized Hermite polynomials to normalized Charlier polynomials. Operators that map a complete orthonormal system to another complete orthonormal system are automatically unitary. The unitary equivalence of the operators follows from the eigenfunction relations given in Propositions 4.1 and 4.6.

Remark 4.7. There is a treasure-trove of relations between orthogonal polynomials and stochastic processes, beyond the Gaussian and Poisson world. See the book by Schoutens [Sch00] and Section 2.7 in the book by Bakry, Gentil and Ledoux [BGL14].

5 Spin

Probabilists are most likely to encounter spins in the context of interacting particle systems (spin flip dynamics on $\{-1,1\}^{\mathbb{Z}^d}$) and statistical mechanics (Ising model). Another context in which spin and spherical harmonics appear is the study of random fields and statistics on spheres [MP11].

Here we provide some quantum background for the concept of spin and additional probabilistic connections. The reader most interested in discrete systems should jump right ahead to Section 5.4.

5.1 Rotations in $\mathbb{R}^3$

Geometry

Let $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$ be the canonical basis vectors in $\mathbb{R}^3$. Consider the matrix

$$R(\theta;\boldsymbol{e}_z)=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}. \tag{5.1}$$

The matrix describes a rotation in $\mathbb{R}^3$ of angle θ around the axis $\boldsymbol{e}_z$. Notice the group property

$$R(\theta+\alpha;\boldsymbol{e}_z)=R(\theta;\boldsymbol{e}_z)R(\alpha;\boldsymbol{e}_z)\qquad(\alpha,\theta\in\mathbb{R}).$$

For every $\boldsymbol{v}=(x,y,z)\in\mathbb{R}^3$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\theta}R(\theta;\boldsymbol{e}_z)\boldsymbol{v}\Big|_{\theta=0}=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}-y\\x\\0\end{pmatrix}=\boldsymbol{e}_z\times\boldsymbol{v}$$

and if we call the antisymmetric matrix appearing in the previous equation A_z, then

$$R(\theta;\boldsymbol{e}_z)=\exp(\theta A_z).$$

More generally, let $\boldsymbol{n}=(p,q,r)$ be a unit vector in $\mathbb{R}^3$. Then for all $\boldsymbol{v}\in\mathbb{R}^3$,

$$\boldsymbol{n}\times\boldsymbol{v}=\begin{pmatrix}qz-ry\\rx-pz\\py-qx\end{pmatrix}=\begin{pmatrix}0&-r&q\\r&0&-p\\-q&p&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}.$$

Let $A(\boldsymbol{n})$ be the antisymmetric matrix appearing in the previous equation. Then

$$R(\theta;\boldsymbol{n})=\exp\big(\theta A(\boldsymbol{n})\big)$$

represents a rotation of angle θ around $\boldsymbol{n}$.

For later purpose we note the relation

$$\boldsymbol{e}_x \times (\boldsymbol{e}_y \times \boldsymbol{r}) - \boldsymbol{e}_y \times (\boldsymbol{e}_x \times \boldsymbol{r}) = \boldsymbol{e}_z \times \boldsymbol{r}. \tag{5.2}$$

Linear algebra. Matrix group $SO(3)$

The *special orthogonal group* in $\mathbb{R}^3$, denoted $SO(3)$, is the group of orthogonal 3×3-matrices with determinant 1. A classical result from linear algebra that for every $R \in SO(3)$ can be brought into the form (5.1) after suitable change of orthonormal basis. Thus every $R \in SO(3)$ represents a rotation as discussed above. Furthermore, every matrix $R \in SO(3)$ can be written as the exponential of some antisymmetric matrix A (however the matrix need not be unique). Conversely, for every antisymmetric matrix A, the exponential $R = \exp(A)$ is in $SO(3)$. Notice that every antisymmetric matrix with real matrix elements has diagonal elements zero, hence trace equal to zero and $\det \exp(A) = \exp(\operatorname{tr} A) = 1$.

5.2 Angular momentum

In classical mechanics

Let $\boldsymbol{r}(t) = (x(t), y(t), z(t))$ be the trajectory of a particle in $\mathbb{R}^3$. Suppose that the motion is a rotation of angular speed ω around the z-axis,

$$\boldsymbol{r}(t) = R(\omega t; \boldsymbol{e}_z)\boldsymbol{r}(0).$$

Then

$$\dot{\boldsymbol{r}}(t) = \omega \boldsymbol{e}_z \times \boldsymbol{r}(t), \quad \boldsymbol{r}(t) \times \dot{\boldsymbol{r}}(t) = (x^2 + y^2)\omega \boldsymbol{e}_z.$$

Multiplying with the mass $m > 0$ of the particle, we get

$$\boldsymbol{r}(t) \times m\dot{\boldsymbol{r}}(t) = (x^2 + y^2)\omega \boldsymbol{e}_z = I_z \omega \boldsymbol{e}_z, \quad I_z = m(x^2 + y^2).$$

In the cases that interest us, momentum is mass times velocity. The *angular momentum* is the vector

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}.$$

For rotations as above, the vector $\boldsymbol{L}$ is parallel to the axis of rotation and its magnitude is the product of the speed of rotation and the *moment of inertia* of the mass around the rotational axis.

In quantum mechanics

In quantum mechanics, the angular momentum becomes an operator-valued vector, $\boldsymbol{L} = \boldsymbol{r} \times \frac{\hbar}{\mathrm{i}}\nabla$. Thus for $\psi : \mathbb{R}^3 \to \mathbb{C}$, leaving aside domain issues,

$$(L_z\psi)(x, y, z) = \Big(x\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial y} - y\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\Big)\psi(x, y, z).$$

Notice

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi\big(R(t;\boldsymbol{e}_z)\boldsymbol{r}\big)\bigg|_{t=0} = (\nabla\psi(\boldsymbol{r}))\cdot(\boldsymbol{e}_z\times\boldsymbol{r}) = \boldsymbol{e}_z\cdot(\boldsymbol{r}\times\nabla\psi(\boldsymbol{r})),$$

where we have used

$$\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = \det[\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}] = \det[\boldsymbol{b},\boldsymbol{c},\boldsymbol{a}] = \boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a}).$$

As a consequence,

$$\frac{\hbar}{\mathrm{i}}\,\frac{\mathrm{d}}{\mathrm{d}t}\psi\big(R(t;\boldsymbol{e}_z)\boldsymbol{r}\big)\bigg|_{t=0} = (L_z\psi)(\boldsymbol{r}).$$

In this sense L_z generates a unitary group of rotations of the wave function.

Commutation relations

A straightforward but somewhat tedious computation yields

$$[L_x,L_y] = \mathrm{i}\hbar L_z,\quad [L_y,L_z] = \mathrm{i}\hbar L_x,\quad [L_z,L_x] = \mathrm{i}\hbar L_y, \tag{5.3}$$

The first equality is a sibling of the relation (5.2) for cross products. The second and third equalities are obtained from the first equality by a cyclic permutation of (L_x,L_y,L_z).

The commutation relations are often rewritten with raising and lowering operators. Set $L^{\pm} = L_x \pm \mathrm{i}L_y$. Then

$$[L^+,L^-] = 2\hbar L_z\quad [L_z,L^-] = -\hbar L^-,\quad [L_z,L^+] = \hbar L^+. \tag{5.4}$$

Suppose that ψ is an eigenfunction of L_z, i.e., $L_z\psi = \ell\hbar\psi$. Then—assuming there is no issue with domains—we have

$$L_zL^+\psi = L^+L_z\psi + [L_z,L^+]\psi = (\ell+1)\hbar L^+\psi.$$

Similarly, $L_zL^-\psi = (\ell-1)\hbar L^+\psi$. Hence $L^{\pm}$ map eigenfunctions of L_z to eigenfunctions of L_z, all the while raising or lowering the eigenvalue.

5.3 Spherical harmonics. Brownian motion on a sphere

Laplace operator on the sphere

Angular momentum is all about rotations and it does not care about the radial component of a function. That is, suppose that

$$\psi(\boldsymbol{r}) = f\big(|\boldsymbol{r}|\big)g\Big(\frac{\boldsymbol{r}}{|\boldsymbol{r}|}\Big)$$

for all $\boldsymbol{r} = (x, y, z) \in \mathbb{R}^3 \setminus \{0\}$ and suitable functions $f : (0, \infty) \to \mathbb{C}$,

$$g : \mathbb{S}^2 \to \mathbb{C}, \quad \mathbb{S}^2 = \{\boldsymbol{r} \in \mathbb{R}^3 : \ |\boldsymbol{r}| = 1\}.$$

Then

$$(L_x^2 + L_y^2 + L_z^2)\psi(\boldsymbol{r}) = f\big(|\boldsymbol{r}|\big)\,(-\hbar^2 \Delta_{\mathbb{S}^2} g)\Big(\frac{\boldsymbol{r}}{|\boldsymbol{r}|}\Big)$$

with $\Delta_{\mathbb{S}}$ the Laplace operator on the sphere. It can be defined as follows [Shu01, Section 22.2]. Set $\tilde{g}(\boldsymbol{r}) = g(\boldsymbol{r}/|\boldsymbol{r}|)$. Then $\tilde{g}(t\boldsymbol{r}) = \tilde{g}(\boldsymbol{r})$ for all t and therefore

$$t^2(\Delta\tilde{g})(t\boldsymbol{r}) = \Delta\tilde{g}(\boldsymbol{r}).$$

Then $\Delta_{\mathbb{S}^2} g$ is the uniquely defined function with

$$\Delta\tilde{g}(\boldsymbol{r}) = \frac{1}{|\boldsymbol{r}|^2}\big(\Delta_{\mathbb{S}^2} g\big)\Big(\frac{\boldsymbol{r}}{|\boldsymbol{r}|}\Big).$$

The operator $\frac{1}{2}\Delta_{\mathbb{S}^2}$ is the infinitesimal generator of Brownian motion on the sphere [Yos49, Hsu02]. The spherical heat semi-group is also presented in [BGL14, Section 2.2].

Spherical coordinates

The operators L_x, L_y, and L_z are associated with differential operators on $\mathbb{S}^2$ in a similar way. The expression of L_z is particularly simple in spherical coordinates

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta$$

where $r \geq 0$, $\theta \in [0, \pi]$, and $\varphi \in [0, 2\pi)$. On the unit sphere $r = 1$. The equations above determine a map $(r, \theta, \varphi) \mapsto \boldsymbol{r}(r, \theta, \varphi)$. Let $f : \mathbb{R}^3 \to \mathbb{R}$ and g be the function f in polar coordinates, i.e.,

$$f\big(\boldsymbol{r}(r, \theta, \varphi)\big) = g(r, \theta, \varphi).$$

Then

$$(L_z f)\big(\boldsymbol{r}(r, \theta, \varphi)\big) = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial\varphi} g(r, \theta, \varphi).$$

By some abuse of notation we write

$$L_z = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial\varphi}$$

and we also use the same letter L_z for the differential operator on $\mathbb{S}^2$.

Spherical harmonics

Let $L^2(\mathbb{S}^2)$ be the space of complex-valued measurable functions that are square-integrable with respect to the unique rotationally invariant measure on $\mathbb{S}^2$ with total mass 4π (in spherical coordinates, $\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$).

Theorem 5.1. *The space $L^2(\mathbb{S}^2)$ has a complete orthonormal system Y_ℓ^m indexed by*

$$\ell \in \mathbb{N}_0, \quad m \in \{-\ell, -\ell+1, \dots, \ell-1, \ell\}$$

such that

$$\Delta_{\mathbb{S}^2} Y_\ell^m = -\ell(\ell+1) Y_\ell^m, \quad \frac{1}{\hbar} L_z Y_\ell^m = m Y_\ell^m.$$

The eigenfunctions Y_ℓ^m are called *spherical harmonics*. In spherical coordinates,

$$Y_\ell^m(\theta, \varphi) \propto \mathrm{e}^{\mathrm{i}m\varphi} P_\ell^m(\cos\theta)$$

with P_ℓ^m given by yet another family of orthogonal polynomials, the *Legendre polynomials*. In probability theory, the spherical harmonics enter the eigenfunction expansion of the transition function of Brownian motion on the sphere [Yos49]. They also help analyze random fields on a sphere [MP11].

Theorem 5.1 also implies that the full space $L^2(\mathbb{R}^3)$ has a complete orthonormal system of joint eigenvectors of $\boldsymbol{L}^2$ and L_z—just multiply spherical harmonics with an orthonormal system for the radial coordinate. The spectrum of $\boldsymbol{L}^2$ is

$$\sigma(\boldsymbol{L}^2) = \big\{\ell(\ell+1)\hbar^2 : \ \ell \in \mathbb{N}_0\big\},$$

the spectrum of L_z consists of integer multiples of $\hbar$. Basis functions satisfy

$$\boldsymbol{L}^2 \psi = \ell(\ell+1)\hbar^2 \psi, \quad L_z \psi = m\hbar\psi$$

with ℓ, m as in Theorem 5.1. The existence of a basis of joint eigenvectors reflects that $\boldsymbol{L}^2$ and L_z commute—remember from linear algebra that two Hermitian matrices that commute can be codiagonalized.

5.4 Pauli matrices. Spin $1/2$

Pauli matrices

The Pauli matrices are three 2×2 matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

They form a basis of the three-dimensional real vector space of Hermitian 2×2-matrices with trace zero and satisfy the relations $\sigma_x\sigma_y = \mathrm{i}\,\sigma_z = -\sigma_x\sigma_y$ and cyclic permutations thereof, as well as $\sigma_x^2 = \sigma_y^2 = \sigma_z^2$. Define

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y, \quad S_z = \frac{\hbar}{2}\sigma_z$$

then

$$[S_x, S_y] = \mathrm{i}\hbar S_z, \quad [S_y, S_z] = \mathrm{i}\hbar S_x, \quad [S_z, S_x] = \mathrm{i}\hbar S_y.$$

Thus we have found Hermitian 2×2 matrices that satisfy the same set of commutation relations as the components L_x, L_y, L_z of the angular momentum. The raising and lowering operators become

$$S^+ = S_x + \mathrm{i}S_y = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S^- = S_x - \mathrm{i}S_y = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Further notice that the eigenvalues of S_z are $\pm\hbar/2$ and

$$\boldsymbol{S}^2 = S_x^2 + S_y^2 + S_z^2 = \frac{3}{4}\hbar^2 = \ell(\ell+1)\hbar^2, \quad \ell = \frac{1}{2}.$$

This should remind us of Theorem 5.1: the action of the matrices S_x, S_y, S_z in $\mathbb{C}^2$ is similar to the action of L_x, L_y, L_z in a $(2\ell+1)$-dimensional eigenspace of $-\Delta_{\mathbb{S}^2}$, except now $\ell = 1/2$.

Particles with spin $1/2$

The question arises whether the matrices S_x, S_y, S_z have any physical relevance at all even though $\ell = 1/2$ does not appear in Theorem 5.1. The answer is yes: Some physics experiments (e.g., the *Stern-Gerlach experiment*) are best explained by viewing particles like the electron not as a point, but instead as an object with some internal degree of freedom. They are often visualized as a little sphere that can rotate around its own axis. The rotation around the axis brings its own angular momentum, called *spin*.

The visualization is then often complemented by somewhat mysterious-sounding phrases like "when you rotate the object by 360 degrees you haven't come full circle, instead you have to rotate twice". This sentence reflects the following: let us set $\hbar = 1$, then

$$\exp(\mathrm{i}2\pi S_z) = \exp(\mathrm{i}2\pi\sigma_z/2) = \begin{pmatrix} \exp(\mathrm{i}\pi) & 0 \\ 0 & \exp(-\mathrm{i}\pi) \end{pmatrix} = -\mathrm{id}.$$

(The relation between $\exp(\mathrm{i}\theta\sigma_z/2)$ and rotations is explained in more detail in the next subsection.)

The mathematical implementation of the picture is simply to choose a different Hilbert space. The Hilbert space of a spin $1/2$ particle moving in $\mathbb{R}^3$ is the tensor product

$$L^2(\mathbb{R}^3) \otimes \mathbb{C}^2,$$

which is isomorphic to the space of functions $\psi : \mathbb{R}^3 \times \{-\frac{1}{2}, \frac{1}{2}\} \to \mathbb{C}$ with norm

$$||\psi||^2 = \int_{\mathbb{R}^3} |\psi(\boldsymbol{x}, \tfrac{1}{2})|^2 \,\mathrm{d}\boldsymbol{x} + \int_{\mathbb{R}^3} |\psi(\boldsymbol{x}, -\tfrac{1}{2})|^2 \,\mathrm{d}\boldsymbol{x}.$$

In this representation $|\psi(\boldsymbol{x}, \pm\frac{1}{2})|^2$ represents the probability (density) that the particle is at $\boldsymbol{x}$ and has spin in z-direction equal to $\pm\hbar/2$.

From $SU(2)$ to $SO(3)$

The similarity of commutation relations suggests there should be a relation between 2×2 matrices of the form $\exp(-\mathrm{i}(xS_x+yS_y+zS_z)/\hbar)$ and rotations in $\mathbb{R}^3$, and indeed there is. For $\boldsymbol{r}=(x,y,z)\in\mathbb{R}^3$, we denote

$$\boldsymbol{r}\cdot\boldsymbol{\sigma}=x\sigma_x+y\sigma_y+z\sigma_z$$

Every matrix of this form is Hermitian and has trace zero; in fact every Hermitian, trace-zero matrix can be written in this way. Hence we may identify $\mathbb{R}^3$ with the space of Hermitian, trace-zero 2×2-matrices. The Euclidean norm is recovered with the determinant as

$$\det(\boldsymbol{r}\cdot\sigma)=\det\begin{pmatrix} z & x-\mathrm{i}y\\ x+\mathrm{i}y & -z\end{pmatrix}=-(x^2+y^2+z^2).$$

The exponential $U=\exp(-\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{\sigma})$ of any such matrix is unitary and has determinant 1 because the determinant of the exponential is the exponential of the trace, a well-known fact from linear algebra. The set of unitary 2×2-matrices with determinant 1 forms a group. It is called the *special unitary group* and denoted $SU(2)$.

If U is in $SU(2)$ and $M=M^*$ with $\operatorname{Tr}M=0$, then

$$(UMU^*)^*=UMU^*,\quad \operatorname{Tr}UMU^*=\operatorname{Tr}M=0,\quad \det(UMU^*)=\det(M).$$

As a consequence, there exists a uniquely defined orthogonal 3×3-matrix R_U such that

$$U(\boldsymbol{r}\cdot\sigma)U^*=(R_U\boldsymbol{r})\cdot\sigma$$

for all $\boldsymbol{r}\in\mathbb{R}^3$. It turns out that the matrix is also orientation-preserving and has determinant 1, thus $R_U\in SO(3)$. Notice $R_{-U}=R_U$ and $R_{UV}=R_UR_V$.

Proposition 5.2. *The group homomorphism*

$$SU(2)\to SO(3),\quad U\mapsto R_U$$

is surjective but not injective. The preimage of each $R\in SO(3)$ is of the form $\{+U,-U\}$ for some $U\in SU(2)$.

Let's work out the image $R(t)$ of

$$U(t)=\exp(-\mathrm{i}t\sigma_z/2).$$

Differentiating

$$\mathrm{e}^{-\mathrm{i}t\sigma_z/2}(\boldsymbol{r}\cdot\boldsymbol{\sigma})\mathrm{e}^{\mathrm{i}t\sigma_z/2}=(R(t)\boldsymbol{r})\cdot\boldsymbol{\sigma}$$

at $t=0$, we get

$$-\frac{\mathrm{i}}{2}\big[\sigma_z,\boldsymbol{r}\cdot\boldsymbol{\sigma}\big]=(R'(0)\boldsymbol{r})\cdot\boldsymbol{\sigma}.$$

An explicit computation shows that the commutator on the left is equal to $2\mathrm{i}(\boldsymbol{e}_z \times \boldsymbol{r}) \cdot \boldsymbol{\sigma}$ and we deduce $R'(0)\boldsymbol{r} = \boldsymbol{e}_z \times \boldsymbol{r}$. Together with $R(0) = \mathrm{id}$, $R(t+s) = R(t)R(s)$, and the geometric considerations from Section 5.1, this shows

$$R_{U(t)} = R(t) = R(t; \boldsymbol{e}_z)$$

is the rotation of angle t around the z-axis. Notice that $U(t + 2\pi) = -U(t)$ but $R(t + 2\pi) = R(t)$.

5.5 Spin $n/2$, Ehrenfest model, Kravchuk polynomials

Motivated by Theorem 5.1 and Section 5.4 we can look for matrices that satisfy the given commutation relations and are associate with half-integer ℓ. The matrices should be $(2\ell + 1) \times (2\ell + 1)$ matrices. In algebraic language, we look for irreducible unitary representations of the special unitary group $SU(2)$ with dimension $2\ell + 1$.

The existence and uniqueness (up to isometry) of such representations is well-known. We provide here a representation which at first sight looks different from the standard physicist's and analyst's representation but lends itself more easily to a probabilistic interpretation.

Spin $n/2$. Kac matrix

Let $\ell \in \mathbb{N}$ and $\ell = n/2$. We label the canonical basis vectors of $\mathbb{C}^{2\ell+1} = \mathbb{C}^{n+1}$ as $\boldsymbol{e}_0, \ldots, \boldsymbol{e}_n$ and define linear maps by

$$J^+\boldsymbol{e}_k = (n-k)\boldsymbol{e}_{k+1}, \quad J^-\boldsymbol{e}_k = k\boldsymbol{e}_{k-1}, \quad J_z\boldsymbol{e}_k = \big(k - \frac{n}{2}\big)\boldsymbol{e}_k.$$

A straightforward computation shows

$$[J_z, J^+] = J_+, \quad [J_z, J^-] = -J^-, \quad [J^+, J^-] = 2J_z$$

and we recognize the commutation relations from (5.4). Set $J_x = \frac{1}{2}(J^+ + J^-)$ and $J_y = \frac{1}{2\mathrm{i}}(J^+ - J^-)$, then $J^\pm = J_x \pm iJ_y$ and

$$J_x^2 + J_y^2 + J_z^2 = \frac{n}{2}\Big(\frac{n}{2} + 1\Big) = \frac{\ell(\ell+1)}{2}$$

(times identity matrix). The eigenvalues of J_z are $-\ell, -\ell+1, \ldots, \ell-1, \ell$. Thus we have indeed found a set of matrices associated with half-integer $\ell \in \frac{1}{2}\mathbb{N}$.

We note in passing that

$$2J_x = J^+ + J^- = \begin{pmatrix} 0 & 1 & & & & \\ n & 0 & 2 & & & \\ & n-1 & 0 & 3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 2 & 0 & n \\ & & & & 1 & 0 \end{pmatrix} \tag{5.5}$$

is sometimes called *Kac matrix* after [Kac47, Eq. (47)] but it was also investigated by others, see the references in [TT91]. Notice that for $\ell = 1/2$ ($n = 1$), it is equal to the Pauli matrix σ_x.

The matrices J_x and J_y are not Hermitian but this can be remedied as follows. Let

$$D = \operatorname{diag}\Big(\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}\Big).$$

be the diagonal matrix with binomial coefficients on the diagonal. Then $J^- D = (J^+ D)^{\mathsf{T}}$. Set $\tilde{J}_u = D^{-1/2} J_u D^{1/2}$, then $\tilde{J}_x$, $\tilde{J}_y$, $\tilde{J}_z$ are Hermitian matrices.

Ehrenfest model. Elastic random walker

The matrix $2J_x$ has non-negative entries and constant column sums equal to n, therefore

$$Q = (2J_x)^{\mathsf{T}} - n\,\mathrm{id}$$

is a Q-matrix. As usual we identify $\mathbb{C}^{n+1}$ with functions $f : \{0, \dots, n\} \to \mathbb{C}$, then

$$(Qf)(k) = (n-k)\big(f(k+1) - f(k)\big) + k\big(f(k-1) - f(k)\big). \tag{5.6}$$

The binomial distribution $\mathsf{Bin}(n, \frac{1}{2})$ is reversible.

The associated continuous-time Markov process with state space $\{0, \dots, n\}$ is the *Ehrenfest model*. Suppose you have a left and right container separated by a membrane and n particles distributed over the two containers (if you prefer, think of two urns containing n balls in total). The particles move independently and change containers at rate 1. If the left container contains k particles, then the net rate that one out of the k particles moves to the right is k, resulting in a decrease $k \to k-1$ in the number of particles on the left. Similarly, the net rate for one of the $n-k$ partices on the right container to move to the left is $n-k$, resulting in an increase $k \to k+1$ in the number of particles on the left.

Another interpretation is in terms of a random walk on a hypercube $\{0,1\}^n$. Suppose that the rate for one coordinate to flip is 1 and you want to count the number of 0's in the time-dependent vector. Then the rate for one out of the k coordinates to flip from zero to 1 is k and the rate for one out of the $n-k$ coordinates equal to 1 to flip back to zero is $n-k$.

Yet another interpretation is obtained by recentering. In the long-time limit, we expect there should be on average $n/2$ particles in both containers and we might decide to keep track of the deviations $k - \frac{n}{2}$, $\frac{n}{2} + k$ instead of the absolute number counts k and $n-k$. The change of variable $x = k - \frac{n}{2}$ leads to a process with state space $\{-\frac{n}{2}, \dots, \frac{n}{2} - 1, \frac{n}{2}\}$ and generator

$$\big(Lf\big)(x) = \frac{n}{2}\big(f(x+1) - 2f(x) + f(x-1)\big) - x\big(f(x+1) - f(x-1)\big),$$

which looks like a discrete version of an Ornstein-Uhlenbeck generator. The associated process is a continuous-time random walk with position-dependent transition rates that push the walker back to the origin. Loosely following Kac [Kac47] we may call the walk *elastic*.

Eigenfunctions. Kravchuk polynomials

The similarity of the elastic random walk with the Ornstein-Uhlenbeck process carries over to the spectrum and the eigenfunctions, given by the *Kravchuk polynomials*, a discrete version of the Hermite polynomials. (Another frequently used transliteration of the Ukrainian name is *Krawtchouk*.) For $r \in \{0, \dots, n\}$, define a function $\mathcal{K}_r : \{0, \dots, n\} \to \mathbb{R}$ by

$$\mathcal{K}_r(k) = \sum_{j=0}^{r} (-1)^j \binom{k}{j} \binom{n-k}{r-j}.$$

(Binomial coefficients $\binom{N}{m}$ with $N < m$ are defined as zero.) For example,

$$\mathcal{K}_0(k) = 1, \quad \mathcal{K}_1(k) = n - 2k.$$

More generally, $\mathcal{K}_r$ is a polynomial of degree r. The Kravchuk polynomials satisfy the orthogonality relation

$$\sum_{k=0}^{n} \binom{n}{k} \mathcal{K}_r(k)\mathcal{K}_s(k) = 2^n \binom{n}{r} \delta_{r,s}.$$

The following proposition is an analogue of Propositions 4.1 and 4.6 on the Ornstein-Uhlenbeck process and the birth and death process.

Proposition 5.3. *The generator of the Ehrenfest model given by* (5.6) *satisfies*

$$Q\mathcal{K}_r = -r\mathcal{K}_r, \quad r = 0, \dots, n.$$

Proof. The generating function of the Kravchuk polynomials is

$$G(t,k) = \sum_{r=0}^{n} t^r \mathcal{K}_r(k) = (1-t)^k (1+t)^{n-k}.$$

It follows that

$$\begin{aligned} -\sum_{r=0}^{n} r t^r \mathcal{K}_r(k) &= -t \frac{\mathrm{d}}{\mathrm{d}t} G(t,k) \\ &= tk(1-t)^{k-1}(1+t)^{n-k} - t(n-k)(1-t)^k(1+t)^{n-k-1}. \end{aligned}$$

We write $\pm t = 1 \pm t - 1$ and find

$$-\sum_{r=0}^{n} r t^r \mathcal{K}_r(k) = k\big(G(t,k-1) - G(t)\big) + (n-k)\big(G(t,k+1) - G(t)\big).$$

As this holds true for all t, the coefficients of t^r on the left side and on the right side must be equal, for all r, and the proposition follows. □

Remark 5.4. Kravchuk polynomials also enter the analysis of random walks on hypercubes [Dia88, DGM90].

Bibliography

[]

[Abd13] Abdelmalek Abdesselam, *QFT, RG, and all that, for mathematicians, in 11 pages*, Electronic preprint arXiv:1311.4897 [math.PR], 2013.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014.

[Cha18] Sourav Chatterjee, *Introduction to quantum field theory for mathematicians*, Online lecture notes `https://statweb.stanford.edu/~souravc/qft-lectures-combined.pdf`, 2018, Based on a forthcoming book by Michel Talagrand.

[Dav69] Edward B. Davies, *Quantum stochastic processes*, Comm. Math. Phys. **15** (1969), 277–304.

[Dav76] ———, *Quantum theory of open systems*, Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York, 1976.

[dFdM10] Edson de Faria and Welington de Melo, *Mathematical aspects of quantum field theory*, Cambridge studies in advanced mathematics, vol. 127, Cambridge University Press, 2010.

[DGM90] Persi Diaconis, R. L. Graham, and J. A. Morrison, *Asymptotic analysis of a random walk on a hypercube with many dimensions*, Random Structures Algorithms **1** (1990), no. 1, 51–72.

[Dia88] Persi Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 11, Institute of Mathematical Statistics, Hayward, CA, 1988.

[Fan10] Mark Fannes, *An introduction to quantum probability*, Theoretical foundations of quantum information processing and communication, Lecture Notes in Phys., vol. 787, Springer, Berlin, 2010, pp. 1–38.

[Far10] William G. Faris, *An outline of quantum mechanics*, Entropy and the Quantum (Robert Sims and Daniel Ueltschi, eds.), Contemp. Math., vol. 529, American Mathematical Society, 2010, `http://www.mathphys.org/AZschool/material/AZ09-faris.pdf`, pp. 1–52.

[FK06] Jin Feng and Thomas G. Kurtz, *Large deviations for stochastic processes*, Mathematical Surveys and Monographs, vol. 131, American Mathematical Society, Providence, RI, 2006.

[Fol08] Gerald B. Folland, *Quantum field theory: A tourist guide for mathematicians*, Mathematical Surveys and Monographs, vol. 149, American Mathematical Society, 2008.

[FW84] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer-Verlag, New York, 1984, Translated from the Russian by Joseph Szücs.

[GJ87] James Glimm and Arthur Jaffe, *Quantum physics. A functional integral point of view*, Springer, New-York, 1987.

[GS11] Stephen J. Gustafson and Israel Michael Sigal, *Mathematical concepts of quantum mechanics*, Universitext, Springer, 2011.

[Hal13] Brian C. Hall, *Quantum theory for mathematicians*, Graduate Texts in Mathematics, vol. 267, 2013, 2013.

[Hel02] Bernard Helffer, *Semiclassical analysis, Witten Laplacians, and statistical mechanics*, Series in Partial Differential Equations and Applications, vol. 1, World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

[Hol11] Alexander Holevo, *Probabilistic and statistical aspects of quantum theory*, second ed., Quaderni/Monographs, vol. 1, Edizioni della Normale, Pisa, 2011, With a foreword from the second Russian edition by K. A. Valiev.

[Hsu02] Elton P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, 2002.

[Kac47] Mark Kac, *Random walk and the theory of Brownian motion*, Amer. Math. Monthly **54** (1947).

[KM98] B. Kümmerer and H. Maassen, *Elements of quantum probability*, Quantum probability communications, QP-PQ, X, World Sci. Publ., River Edge, NJ, 1998, pp. 73–100.

[LL01] Elliott H. Lieb and Michael Loss, *Analysis*, Graduate Studies in Mathematics, vol. 14, American Mathematical Societty, 2001.

[LS09] Elliott H. Lieb and Robert Seiringer, *The stability of matter in quantum mechanics*, Cambridge University Press, 2009.

[Mes14] Albert Messiah, *Quantum mechanics. Two volumes bound as one*, Dover Publications, 2014, Reprint from the 1961-1962 original.

[Mey95] Paul A. Meyer, *Quantum probability for probabilists*, 2 ed., Lecture Notes in Mathematics, vol. 1538, Springer, 1995.

[MP11] Domenico Marinucci and Giovanni Peccati, *Random fields on the sphere*, London Mathematical Society Lecture Note Series, vol. 389, Cambridge University Press, Cambridge, 2011, Representation, limit theorems and cosmological applications.

[Nur19] Hendra I. Nurdin, *Quantum stochastic processes and the modelling of quantum noise*, pp. 1–8, Springer London, London, 2019.

[Par92] K. R. Parthasarathy, *An introduction to quantum stochastic calculus*, Monographs in Mathematics, vol. 85, Birkhäuser Verlag, Basel, 1992.

[PT11] Giovanni Peccati and Murad S. Taqqu, *Wiener chaos: Moments, cumulants, and diagrams. a survey with computer implementation*, Bocconi & Springer book series, vol. 1, Springer, Milano, 2011.

[RS75] Michael Reed and Barry Simon, *Fourier analysis, self-adjointness*, Methods of modern mathematical physics, vol. 2, Academic Press-Elsevier, 1975.

[RS78] ———, *Analysis of operators*, Methods of modern mathematical physics, vol. 4, Academic Press-Elsevier, 1978.

[RS81] ———, *Functional analysis*, Methods of modern mathematical physics, vol. 1, Academic Press-Elsevier, 1981.

[Sch00] Wim Schoutens, *Stochastic processes and orthogonal polynomials*, Lecture Notes in Statistics, vol. 146, Springer, New York, NY, 2000.

[Sch08] Martin Schottenloher, *A mathematical introduction to conformal field theory*, Lecture Notes in Physics, vol. 759, Springer-Verlag Berlin Heidelberg, 2008.

[She07] Scott Sheffield, *Gaussian free field for mathematicians*, Probab. Theory Relat. Fields **139** (2007), 521–541.

[Shu01] Mikhail A. Shubin, *Pseudodifferential operators and spectral theory*, 2 ed., Springer-Verlag, Berlin-Heidelbert, 2001.

[Sim05] Barry Simon, *Functional integration and quantum physics*, 2 ed., AMS Chelsea Publishing-American Mathematical Society, 2005.

[Spe] Roland Speicher, *L]ecture notes on "Free probability", howpublished = arXiv:1908.08125 [math.OA], year = 2019.*

[Str00] R. F. Streater, *Classical and quantum probability*, J. Math. Phys. **41** (2000), no. 6, 3556–3603.

[Swa20] Jan Swart, *Introduction to quantum probability*, Lecture notes. Available at `http://staff.utia.cas.cz/swart/lecture_notes/qua20_04_15.pdf`. Downloaded 15 January 2021, 2020.

[TT91] Olga Taussky and John Todd, *Another look at a matrix of Mark Kac*, Proceedings of the First Conference of the International Linear Algebra Society (Provo, UT, 1989), vol. 150, 1991, pp. 341–360.

[Uel20] Daniel Ueltschi, *Quantum mechanics: basic principles and probabilistic methods*, Online lecture notes `http://www.ueltschi.org/teaching/2019-MA4A7.html`. Last consulted 3 April 2020, 2020.

[Yos49] Kosaku Yosida, *Brownian motion on the surface of the 3-sphere*, Ann. Math. Statist. **22** (1949), no. 2, `https://projecteuclid.org/euclid.aoms/1177730038`.

[Zei06] Eberhard Zeidler, *Quantum field theory I: Basics in mathematics and physics. A bridge between mathematicians and physicists*, Springer-Verlag Berlin Heidelberg, 2006.

[Zei09] ———, *Quantum field theory II: Quantum electrodynamics. A bridge between mathematicians and physicists*, Springer-Verlag Berlin Heidelberg, 2009.