The Feynman–Kac Formula:
A Measure-Theoretic, Analytic, and Probabilistic Synthesis

Sourangshu Ghosh

***sourangshug@iisc.ac.in, Department of Civil Engineering, Indian Institute of Science Bangalore**

Contents

Preface

"In mathematics, one does not understand things. One just gets used to them."

— John von Neumann

The **Feynman–Kac formula** stands among the rare mathematical results that elegantly bridge distinct conceptual worlds — analysis, probability, and physics — by asserting that the evolution of analytic structures such as parabolic partial differential equations can be represented through the expectations of random paths. It is a synthesis that makes rigorous Feynman's intuitive vision of quantum propagation via path integrals and Kac's probabilistic representation of parabolic equations.

The purpose of this monograph is to present the Feynman–Kac formula and its far-reaching generalizations within a **rigorous measure-theoretic and functional-analytic framework**. While many excellent expositions introduce the formula as a computational tool or as a heuristic bridge between stochastic processes and partial differential equations, few texts aim to *unify its analytical, probabilistic, and physical interpretations* within a single, fully self-contained narrative.

This book takes that unification as its central theme.

The Triad: Probability, Analysis, and Physics

At its heart, the Feynman–Kac formula expresses a deep correspondence:

$$(\text{Stochastic Process}) \quad \longleftrightarrow \quad (\text{Semigroup of Operators}) \quad \longleftrightarrow \quad (\text{Parabolic PDE}).$$

Each of these viewpoints offers its own language:

- The **probabilist** sees the formula as an expectation of functionals of diffusion processes;
- The **analyst** views it as the integral kernel representation of a strongly continuous semigroup;
- The **physicist** recognizes it as the Euclidean-time formulation of the path integral.

Bringing these languages into harmony requires machinery from measure-theoretic probability, semigroup theory, potential theory, and spectral analysis. The result is not only a proof of the formula itself but also a conceptual architecture connecting Markov processes, Schrödinger operators, and heat kernels.

Scope and Philosophy

The treatment here is **rigorous, self-contained, and comprehensive**. Every theorem is proved from first principles, beginning with the construction of probability spaces and Wiener measure, proceeding through Itô's stochastic calculus, and culminating in the analysis of Feynman–Kac semigroups, spectral representations, and modern extensions to infinite-dimensional and stochastic partial differential equations.

The presentation seeks not only to *prove* the formula but also to *illuminate* it — to show why it is inevitable once one understands the interplay between randomness and linear evolution. To that end, analytic rigor is complemented by geometric and probabilistic intuition, and wherever appropriate, by remarks that connect the results to their origins in physics and to their modern applications in analysis and probability.

Structure of the Book

The book is organized into four major parts:

Part I: Mathematical Foundations develops the measure-theoretic and analytic background required to understand the Feynman–Kac formula in its full generality. Topics include probability spaces, stochastic processes, Brownian motion, Itô calculus, and the theory of strongly continuous semigroups.

Part II: The Feynman–Kac Formula presents the classical and generalized statements of the Feynman–Kac theorem, together with detailed proofs. This part develops the correspondence between stochastic differential equations and the generators of parabolic PDEs, leading to the precise probabilistic representation of Schrödinger semigroups.

Part III: Analytical and Physical Interpretations explores the deep structural implications of the formula — its manifestation in potential theory, its realization as a rigorous Euclidean path integral, and its role in the spectral theory of Schrödinger operators.

Part IV: Generalizations and Modern Extensions surveys the frontier: extensions to Lévy and jump processes, fractional operators, infinite-dimensional systems, stochastic PDEs, and contemporary applications in quantum field theory and statistical physics.

Each chapter is accompanied by remarks, examples, and, where appropriate, historical notes tracing the intellectual evolution of the subject from Wiener's construction of Brownian motion (1923) and Kac's probabilistic approach to diffusion (1949) to modern analytic developments in semigroup and potential theory.

Intended Audience

This monograph is addressed to **graduate students, researchers, and mathematically mature readers** in analysis, probability theory, and mathematical physics. A reader familiar with measure theory, functional analysis, and the basics of stochastic calculus will be fully prepared to engage with the material, though the presentation strives to develop the necessary background systematically and self-sufficiently.

For physicists, the book provides a rigorous foundation for the intuitive path-integral formalism. For probabilists, it demonstrates how analytic and operator-theoretic techniques enrich the study of stochastic processes. For analysts, it shows how probabilistic representations yield deep insights into PDEs and spectral theory.

Acknowledgments

This book owes its existence to the enduring beauty of the Feynman–Kac formula itself — a work of mathematical art born from the collaboration between probability and physics. The author expresses profound gratitude to those who inspired and sustained this exploration — to the mentors who nurtured a deep appreciation of mathematical structure, to colleagues for their stimulating discussions, and to the community of mathematicians and physicists whose works have laid the foundations on which this book stands.

A Final Word

If there is a single message this book aims to convey, it is that the **Feynman–Kac formula is not merely a tool — it is a philosophy**. It embodies the idea that randomness can reveal structure, that expectation can replace determinism, and that, in the language of mathematics, the probabilistic and the analytic are but two sides of the same truth.

Sourangshu Ghosh
Indian Institute of Science, Bangalore
November 22, 2025

Part I

Mathematical Foundations

Chapter 1

Measure-Theoretic Probability and Stochastic Processes

> "Probability is the measure theory of ignorance; its objects are sets of possible worlds, its morphisms are filtrations of information."
>
> — Adapted from Joseph L. Doob (1953)

1.1 Introduction

The purpose of this chapter is to lay the rigorous measure-theoretic foundations upon which the Feynman–Kac formula is built. The formula, which links stochastic processes with parabolic partial differential equations and quantum mechanics, presupposes a precise understanding of probability as a measure space, of stochastic processes as measurable mappings, and of martingales, filtrations, and conditional expectations as the structural backbone of probabilistic dynamics.

We shall work entirely within the modern framework of measure-theoretic probability, viewing stochastic processes as measurable functions on product spaces endowed with suitable filtrations. This framework enables the use of tools from analysis, topology, and functional analysis to treat stochastic evolution, which is essential for the analytic continuation into semigroup and path-integral formulations in later chapters.

1.2 Probability Spaces and Random Variables

A probability space is the fundamental object in modern probability theory, providing a rigorous framework for the formalization of uncertainty. Formally, a probability space is defined as a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a non-empty set called the sample space, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure satisfying $\mathbb{P}(\Omega) = 1$ and countable additivity.

Definition 1.1 (Probability Space)**.** A *probability space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where:

1. Ω is the sample space, the set of all possible outcomes;

2. $\mathcal{F}$ is a σ-algebra of subsets of Ω;

3. $\mathbb{P} : \mathcal{F} \to [0,1]$ is a countably additive measure with $\mathbb{P}(\Omega) = 1$.

The σ-algebra $\mathcal{F}$ encodes the collection of events whose probabilities are well-defined, and countable additivity ensures that if $A_n{}_{n=1}^{\infty}$ is a sequence of pairwise disjoint events in $\mathcal{F}$, then

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mathbb{P}(A_n). \tag{1.1}$$

This axiom captures the intuitive notion that probabilities of mutually exclusive outcomes add linearly. Within this framework, events with probability zero are negligible in the sense that they do not affect the almost sure properties of random variables and stochastic processes.

A random variable is a measurable function from the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to a measurable space $(E, \mathcal{E})$, most commonly $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$.

Definition 1.2 (Random Variable)**.** A *random variable* X is a measurable function

$$X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})), \tag{1.2}$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$.

The pushforward measure $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$ defines the *distribution* of X. Expectations are Lebesgue integrals with respect to $\mathbb{P}$, denoted $\mathbb{E}[X] = \int_\Omega X \, d\mathbb{P}$. Formally, a function $X : \Omega \to \mathbb{R}$ is a random variable if, for every $B \in \mathcal{B}(\mathbb{R})$, the preimage

$$X^{-1}(B) = \omega \in \Omega : X(\omega) \in B \in \mathcal{F} \tag{1.3}$$

This requirement ensures that probabilistic statements about X can be expressed in terms of the underlying measure $\mathbb{P}$. For instance, the probability that X falls within a set B is defined as

$$\mathbb{P}(X \in B) := \mathbb{P}(X^{-1}(B)). \tag{1.4}$$

The distribution of a random variable X is the measure $\mathbb{P}_X$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ given by $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$, which completely characterizes the probabilistic behavior of X. In this context, the expectation of an integrable random variable $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ is defined as the Lebesgue integral

$$\mathbb{E}[X] := \int_\Omega X(\omega), d\mathbb{P}(\omega), \tag{1.5}$$

providing a linear functional on the space of integrable random variables. Higher moments, variances, and covariance structures are defined analogously, and they form the building blocks for the study of convergence, limit theorems, and stochastic calculus.

Random variables can be combined to form vectors or stochastic processes. If $X_1, X_2, \ldots, X_n$ are random variables on the same probability space, their joint distribution is defined on the product measurable space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ by

$$\mathbb{P}_{X_1,\ldots,X_n}(B_1 \times \cdots \times B_n) := \mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n), \tag{1.6}$$

for all $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$. This definition generalizes naturally to infinite-dimensional settings, leading to the theory of stochastic processes where each X_t is a random variable indexed by time t. The concept of measurability ensures that all probabilistic constructions, including expectations, conditional distributions, and integrals with respect to probability measures, are mathematically well-defined.

Independence of random variables is also formalized within this framework. A collection $X_i * i \in I$ is said to be independent if, for any finite subcollection $X * i_1, \ldots, X_{i_n}$ and any sets $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$, one has

$$\mathbb{P}(X_{i_1} \in B_1, \ldots, X_{i_n} \in B_n) = \prod_{k=1}^{n} \mathbb{P}(X_{i_k} \in B_k). \tag{1.7}$$

This rigorous, measure-theoretic definition provides the foundation for the development of martingales, Markov processes, and Brownian motion, and ultimately allows the extension to continuous-time stochastic calculus, which is the core framework underlying the Feynman–Kac formula.

1.3 Independence and Conditional Expectation

Independence is one of the central concepts in probability theory, formalizing the notion that the occurrence of one event or the value of one random variable does not influence another. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ be two sub-σ-algebras.

Definition 1.3 (Independence)**.** Two σ-algebras $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ are said to be *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \quad \forall A \in \mathcal{A},\ B \in \mathcal{B}. \tag{1.8}$$

This property implies that the probability of the intersection of events factorizes as the product of individual probabilities. Random variables X, Y are independent if the σ-algebras they generate, $\sigma(X)$ and $\sigma(Y)$, are independent.

A collection of σ-algebras $\mathcal{F}_i, i \in I$ is independent if every finite subcollection is independent in this sense. Random variables X and Y are independent if the σ-algebras they generate, $\sigma(X)$ and $\sigma(Y)$, are independent. Equivalently, for all Borel sets $B_1, B_2 \subset \mathbb{R}$,

$$\mathbb{P}(X \in B_1, Y \in B_2) = \mathbb{P}(X \in B_1), \mathbb{P}(Y \in B_2). \tag{1.9}$$

Independence extends naturally to families of random variables and to stochastic processes, allowing the decomposition of joint distributions into products of marginal distributions, which is crucial for the construction of product measures and for establishing limit theorems such as the law of large numbers and the central limit theorem.

Conditional expectation generalizes the classical notion of expectation by incorporating partial information represented by a sub-σ-algebra. Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ be an integrable random variable, and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra.

Definition 1.4 (Conditional Expectation)**.** Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G} \subset \mathcal{F}$ a sub-σ-algebra. The *conditional expectation* of X given $\mathcal{G}$ is the unique (a.s.) $\mathcal{G}$-measurable random variable $\mathbb{E}[X|\mathcal{G}]$ satisfying

$$\int_G \mathbb{E}[X|\mathcal{G}]\, d\mathbb{P} = \int_G X\, d\mathbb{P}, \quad \forall\, G \in \mathcal{G}. \tag{1.10}$$

Conditional expectations generalize the notion of best approximation in L^2 to the σ-algebra $\mathcal{G}$. They play a central role in defining martingales and Markov processes. This definition ensures that $\mathbb{E}[X \mid \mathcal{G}]$ preserves the integral of X over all events measurable with respect to $\mathcal{G}$, and it can be interpreted as the "best prediction" of (X) given the information in $\mathcal{G}$. When $\mathcal{G} = \emptyset, \Omega$ is the trivial σ-algebra, $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$, whereas if $\mathcal{G} = \mathcal{F}$, then $\mathbb{E}[X \mid \mathcal{G}] = X$.

Conditional expectation satisfies linearity, monotonicity, and the tower property: for $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$,

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}]. \tag{1.11}$$

If X is independent of $\mathcal{G}$, then $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$ almost surely. For two integrable random variables X and Y, the conditional expectation $\mathbb{E}[X \mid Y]$ is defined as $\mathbb{E}[X \mid \sigma(Y)]$, providing a measurable function $g(Y)$ such that

$$\mathbb{E}[X\mathbf{1}_{Y\in B}] = \mathbb{E}[g(Y)\mathbf{1}_{Y\in B}], \quad \forall B \in \mathcal{B}(\mathbb{R}). \tag{1.12}$$

Conditional expectation also interacts naturally with integrals and limits, allowing the application of dominated convergence and monotone convergence theorems in a conditional setting. This machinery underpins martingale theory, stochastic integration, and the rigorous development of Markov processes, forming the analytic foundation necessary for the Feynman–Kac formula, where conditional expectations over Brownian paths produce solutions to parabolic partial differential equations.

1.4 Filtrations and Adapted Processes

In the study of stochastic processes, the concept of a filtration formalizes the evolution of information over time. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A filtration is a family of sub-σ-algebras $(\mathcal{F}_t)_{t\geq 0}$ such that

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}, \quad \forall, 0 \leq s \leq t. \tag{1.13}$$

Each $\mathcal{F}_t$ represents the collection of events whose outcomes are "known" at time (t). The increasing nature of the filtration ensures that no information is lost as time progresses; rather, knowledge accumulates, which is essential for the precise definition of conditional expectations and martingales in continuous time.

Definition 1.5 (Filtration)**.** A *filtration* $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F})$ is an increasing family of sub-σ-algebras:

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}, \quad \forall s \leq t. \tag{1.14}$$

The pair $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is called a *filtered probability space.*

Filtrations provide the formal backbone for modeling temporal dependencies in stochastic processes, particularly when dealing with adapted and progressively measurable processes.

A stochastic process $(X_t)_{t\geq 0}$ is said to be adapted to a filtration $(\mathcal{F}_t)$ if, for every $t \geq 0$, the random variable X_t is $\mathcal{F}_t$-measurable, i.e.,

$$X_t^{-1}(B) \in \mathcal{F}_t, \quad \forall B \in \mathcal{B}(\mathbb{R}). \tag{1.15}$$

Adaptedness ensures that the value of the process at time (t) depends only on the information available up to that time, reflecting the natural causality in stochastic dynamics. This condition is fundamental in defining stochastic integrals, where integrands must be adapted to guarantee that they do not "anticipate" future information.

Progressive measurability is a stronger regularity condition. A process $(X_t)_{t\geq 0}$ is progressively measurable with respect to $(\mathcal{F}_t)$ if, for every $t \geq 0$, the mapping $(s, \omega) \mapsto X_s(\omega)$ restricted to $[0, t] \times \Omega$ is measurable with respect to the product σ-algebra $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$. Formally, for all Borel sets $B \subset \mathbb{R}$,

$$(s, \omega) \in [0, t] \times \Omega : X_s(\omega) \in B \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t. \tag{1.16}$$

Progressive measurability guarantees that the process can be integrated in the Itô sense, and it ensures the existence of modifications with continuous paths when such regularity is required.

Definition 1.6 (Adapted and Progressive Processes)**.** A stochastic process $(X_t)_{t\geq 0}$ is said to be:

- *Adapted* if X_t is $\mathcal{F}_t$-measurable for all t.
- *Progressively measurable* if $(t, \omega) \mapsto X_t(\omega)$ is measurable with respect to $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ for every t.

Importantly, every progressively measurable process is adapted, but the converse is not necessarily true. In practice, most processes of interest in stochastic analysis, such as Brownian motion and solutions to stochastic differential equations, are constructed to be progressively measurable.

Filtrations can also be augmented to satisfy the usual conditions: completeness and right-continuity. A filtration $(\mathcal{F}_t)$ is complete if all $\mathbb{P}$-null sets of $\mathcal{F}$ are included in each $\mathcal{F}_t$. Right-continuity means that $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$. These conditions are essential in the formulation of the strong Markov property and in ensuring that stopping times are well-behaved. In particular, the right-continuous augmented filtration guarantees that stopping times τ satisfy $\tau \leq t \in \mathcal{F}_t$ for all t, enabling rigorous applications of optional sampling theorems and martingale convergence results.

The interplay between filtrations and adapted processes forms the structural backbone for modern stochastic calculus. For a stochastic integral $\int_0^t X_s, dB_s$ with respect to a Brownian motion B_t, the integrand X_s must be adapted to the natural filtration generated by (B) to prevent anticipating future increments. Likewise, the conditional expectation $\mathbb{E}[X_t \mid \mathcal{F}_s]$ for $s \leq t$ is meaningful only in the context of an adapted process relative to the filtration $(\mathcal{F}_t)$. In the context of the Feynman–Kac formula, the adaptedness of functionals of Brownian motion guarantees that stochastic integrals and exponential functionals are well-defined and measurable with respect to the natural filtration, allowing a pathwise representation of solutions to parabolic partial differential equations.

1.5 Martingales and Stopping Times

Martingales form one of the cornerstones of modern probability theory, capturing the notion of a "fair game" in a rigorous measure-theoretic framework. Let $(\Omega, \mathcal{F}, (\mathcal{F} * t) * t \geq 0, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, and let $(M_t)_{t\geq 0}$ be an adapted stochastic process with $M_t \in L^1(\Omega)$ for all $t \geq 0$.

Definition 1.7 (Martingale)**.** Let $(\mathcal{F}_t)$ be a filtration. A process $(M_t)_{t\geq 0}$ is a *martingale* if:

1. $M_t \in L^1(\Omega)$ for all t;
2. $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ a.s. for all $s < t$.

If $\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s$, it is a submartingale; if $\leq$, a supermartingale.

This property expresses the intuitive idea that, given all information available up to time s, the conditional expectation of future values equals the present value, implying that there is no predictable trend or drift in the process. If the inequality

$$\mathbb{E}[M_t \mid \mathcal{F}_s] \geq M_s \tag{1.17}$$

holds, the process is called a submartingale, representing a process with a nonnegative expected drift, while the reverse inequality defines a supermartingale. Martingales satisfy several fundamental properties, including linearity, preservation under stopping times (under appropriate conditions), and the martingale convergence theorem, which asserts that certain bounded or integrable martingales converge almost surely and in L^1 as time tends to infinity.

Stopping times are random times at which one may choose to observe or "stop" a stochastic process based only on past and present information. Formally, a random variable $\tau : \Omega \to [0, \infty]$ is a stopping time with respect to a filtration $(\mathcal{F}_t)$ if, for each $t \geq 0$, the event $\tau \leq t \in \mathcal{F}_t$.

Definition 1.8 (Stopping Time)**.** A random variable $\tau : \Omega \to [0, \infty]$ is a *stopping time* if

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, \quad \forall t \geq 0. \tag{1.18}$$

Stopping times encode random observation times determined by the evolution of the process itself, and are indispensable in the optional sampling and strong Markov properties of diffusions.

This definition ensures that the decision to stop at time τ depends only on information available up to that time, prohibiting any anticipation of future events. Classic examples include the first exit time from a set $D \subset \mathbb{R}^n$ by a stochastic process X_t:

$$\tau_D := \inf t \geq 0 : X_t \notin D. \tag{1.19}$$

Stopping times are central to optional sampling and the strong Markov property. The optional sampling theorem states that if M_t is a martingale and τ is a bounded stopping time, then

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0]. \tag{1.20}$$

More generally, for two stopping times $\sigma \leq \tau$, one has

$$\mathbb{E}[M_\tau \mid \mathcal{F}_\sigma] = M_\sigma \quad \text{a.s.}, \tag{1.21}$$

under suitable integrability conditions. This theorem formalizes the notion that a martingale retains its "fairness" even when stopped at a random time. Stopping times also facilitate the construction of stopped processes $M_{t\wedge\tau}$, which remain martingales with respect to the original filtration.

Martingales and stopping times interact intimately in the analysis of Markov processes, Brownian motion, and solutions to stochastic differential equations. For example, in the context of a Brownian motion B_t and a sufficiently smooth function f, the process

$$M_t := f(B_t) - \int_0^t \frac{1}{2}\Delta f(B_s), ds \tag{1.22}$$

is a martingale, and stopping it at the first exit time from a domain D yields a powerful probabilistic representation of solutions to the Dirichlet problem. These concepts underpin the Feynman–Kac formula, where exponential functionals of martingales stopped at random times represent solutions of parabolic partial differential equations with boundary conditions.

Finally, martingale theory provides convergence and decomposition results, such as Doob's martingale convergence theorem and the Doob–Meyer decomposition, which allow any submartingale to be expressed as the sum of a martingale and a predictable, increasing process. These results form the analytic foundation for continuous-time stochastic integration, stochastic calculus, and the rigorous derivation of the Feynman–Kac representation, establishing a deep connection between probabilistic and analytic formulations of evolution equations.

1.6 Stochastic Processes and Finite-Dimensional Distributions

A stochastic process is a collection of random variables indexed by a parameter, typically interpreted as time. Formally, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $T \subseteq \mathbb{R}$ denote the index set, often $T = [0, \infty)$. A stochastic process is a family $(X_t)_{t\in T}$ of random variables $X_t : \Omega \to \mathbb{R}^d$, each measurable with respect to $\mathcal{F}$ and the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$. The process may be viewed equivalently as a single measurable mapping from Ω into the product space $\mathbb{R}^{d\times T}$, endowed with the product σ-algebra generated by cylinder sets. This viewpoint is particularly useful when constructing measures on spaces of paths, such as $C([0, \infty), \mathbb{R}^d)$ or $D([0, \infty), \mathbb{R}^d)$.

The law of a stochastic process is determined by its finite-dimensional distributions. For any finite collection of times $t_1 < t_2 < \cdots < t_n$ in T, the corresponding finite-dimensional distribution is the measure on $(\mathbb{R}^{dn}, \mathcal{B}(\mathbb{R}^{dn}))$ defined by

$$\mu_{t_1,\ldots,t_n}(B_1 \times \cdots \times B_n) := \mathbb{P}(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n), \quad B_i \in \mathcal{B}(\mathbb{R}^d). \tag{1.23}$$

These distributions describe the joint behavior of the process at finitely many times and must satisfy consistency conditions: if $\pi_{1,\ldots,n}$ denotes the projection from $\mathbb{R}^{d(m)}$ to $\mathbb{R}^{dn}$ with $m > n$, then

$$\mu_{t_1,\ldots,t_n} = \mu_{t_1,\ldots,t_m} \circ \pi^{-1}_{1,\ldots,n}. \tag{1.24}$$

Consistency guarantees that the collection of finite-dimensional distributions defines a unique measure on the path space, a result formalized in the Kolmogorov Extension Theorem. Explicitly, if $\mu_{t_1,\ldots,t_n} * n \in \mathbb{N}, t_1 < \cdots < t_n \in T$ is a consistent family, there exists a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ and a stochastic process $(X_t) * t \in T$ on this space such that the finite-dimensional distributions of X coincide with the given measures $\mu_{t_1,\ldots,t_n}$.

This construction allows the rigorous definition of canonical processes on path spaces. For instance, consider the canonical space $\Omega = C([0,\infty), \mathbb{R}^d)$ of continuous functions, equipped with the Borel σ-algebra generated by the topology of uniform convergence on compact intervals. The coordinate mappings $X_t(\omega) = \omega(t)$ define a process whose law is entirely determined by the finite-dimensional distributions. In this context, conditions such as continuity, Hölder regularity, or bounded variation can be imposed pathwise, and Kolmogorov's continuity theorem ensures the existence of a continuous modification if certain moment conditions on the finite-dimensional distributions are satisfied: if there exist constants $\alpha, \beta, C > 0$ such that

$$\mathbb{E}[|X_t - X_s|^\alpha] \le C|t-s|^{1+\beta}, \quad \forall s, t \in T, \tag{1.25}$$

then there exists a modification of the process with almost surely Hölder continuous paths of order γ for any $\gamma < \beta/\alpha$.

The notion of finite-dimensional distributions also allows the precise definition of independence and stationarity for stochastic processes. A process $(X_t) * t \in T$ has independent increments if, for any finite sequence $0 \le t_0 < t_1 < \cdots < t_n$, the increments $X * t_1 - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent random variables. Stationarity of the finite-dimensional distributions, i.e., $\mathbb{P}(X_{t_1+h} \in B_1, \ldots, X_{t_n+h} \in B_n) = \mathbb{P}(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n)$ for all h, ensures temporal homogeneity, which is crucial in the construction of Markov processes and Brownian motion.

In summary, finite-dimensional distributions provide the probabilistic skeleton of a stochastic process, encoding the joint behavior at any finite set of times. Through consistency and extension theorems, they enable the rigorous definition of processes on function spaces, control over path properties, and the foundation for the development of Markovian and martingale structures that underpin stochastic calculus and the Feynman–Kac formula.

A stochastic process is a collection of random variables $(X_t)_{t\in T}$ indexed by time. Its law is determined by the family of *finite-dimensional distributions*:

$$\mu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) := \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n), \tag{1.26}$$

subject to consistency conditions ensuring the existence of a process via the Kolmogorov Extension Theorem.

Theorem 1.9 (Kolmogorov Extension)**.** *Given a family of consistent finite-dimensional distributions on $\mathbb{R}^T$, there exists a stochastic process with these finite-dimensional marginals.*

Proof. Let (T) be a non-empty index set and consider the product space $\Omega := \mathbb{R}^T$, whose elements are functions $\omega : T \to \mathbb{R}$. Let $\mathcal{C}$ denote the collection of cylinder sets, that is, sets of the form

$$C = \omega \in \Omega : (\omega(t_1), \dots, \omega(t_n)) \in B, \tag{1.27}$$

where $n \in \mathbb{N}), (t_1, \dots, t_n \in T$ are distinct, and $B \in \mathcal{B}(\mathbb{R}^n)$ is a Borel set. The collection $\mathcal{C}$ generates the cylinder σ-algebra $\mathcal{F}_{\text{cyl}} := \sigma(\mathcal{C})$, which will serve as the initial domain for defining a measure corresponding to the finite-dimensional distributions.

Suppose we are given a family of finite-dimensional distributions $\mu_{t_1,\dots,t_n} * n \in \mathbb{N}, t_i \in T$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying the consistency condition: for any $m > n$ and any selection of times $t_1, \dots, t_m$, the marginal on the first n coordinates coincides with $\mu * t_1, \dots, t_n$, i.e.,

$$\mu_{t_1,\dots,t_m} \circ \pi^{-1} * 1, \dots, n = \mu * t_1, \dots, t_n, \tag{1.28}$$

where $\pi_{1,\dots,n} : \mathbb{R}^m \to \mathbb{R}^n$ is the canonical projection $(x_1, \dots, x_m) \mapsto (x_1, \dots, x_n)$. This consistency ensures that the family is projectively compatible, a necessary condition for the extension to a measure on the full product space.

Define a set function $\mathbb{P}_0$ on the algebra $\mathcal{C}$ by

$$\mathbb{P} * 0\Big(\omega \in \Omega : (\omega(t_1), \dots, \omega(t_n)) \in B\Big) := \mu * t_1, \dots, t_n(B), \quad B \in \mathcal{B}(\mathbb{R}^n). \tag{1.29}$$

The consistency of the finite-dimensional distributions guarantees that $\mathbb{P}_0$ is well-defined; that is, if the same cylinder set can be represented using different coordinate sets, the assigned probability is independent of the representation. It is straightforward to verify that $\mathbb{P}_0$ is finitely additive on $\mathcal{C}$ and satisfies $\mathbb{P}_0(\Omega) = 1$.

Next, we invoke **Carathéodory's extension theorem**. Since $\mathcal{C}$ is an algebra and $\mathbb{P} * 0$ is countably additive on increasing sequences of cylinder sets (which follows from the continuity of measures in finite dimensions), there exists a unique measure $\mathbb{P}$ on $\sigma(\mathcal{C}) = \mathcal{F} * \text{cyl}$ such that

$$\mathbb{P}\big|_{\mathcal{C}} = \mathbb{P}_0. \tag{1.30}$$

This measure $\mathbb{P}$ now lives on the product space $\Omega = \mathbb{R}^T$ with the cylinder σ-algebra, and by construction, the coordinate projections $X_t : \Omega \to \mathbb{R}$, defined by $X_t(\omega) := \omega(t)$, are random variables with the property that, for any finite set of times $t_1, \dots, t_n \in T$ and any Borel set $B \in \mathcal{B}(\mathbb{R}^n)$,

$$\mathbb{P}\big((X_{t_1}, \dots, X_{t_n}) \in B\big) = \mu_{t_1,\dots,t_n}(B). \tag{1.31}$$

This demonstrates that the process $(X_t) * t \in T$ has the prescribed finite-dimensional distributions. The uniqueness of $\mathbb{P})on(\mathcal{F} * \text{cyl}$ guarantees that any other process with the same finite-dimensional distributions is equivalent in law, i.e., their pushforward measures on the product space coincide.

If the index set (T) is uncountable, one may refine the construction by considering the Kolmogorov consistency conditions and applying extensions to the product σ-algebra generated by all cylinder sets. In particular, one can construct a version of the process on a path space $\Omega = \mathbb{R}^T$ or on $C([0, \infty), \mathbb{R}^d)$ using tightness arguments and continuity theorems such as Kolmogorov's continuity criterion, which ensures the existence of a modification with continuous paths under suitable moment bounds on the finite-dimensional distributions: if there exist constants $\alpha, \beta, C > 0$ such that

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t-s|^{1+\beta} \quad \forall s, t \in T \tag{1.32}$$

then a continuous modification exists almost surely. This completes the construction of a stochastic process consistent with any given family of finite-dimensional distributions, yielding the rigorous statement of the Kolmogorov Extension Theorem. □

1.7 Markov Processes and Transition Kernels

A stochastic process $(X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is called a Markov process if it satisfies the **Markov property**, which formalizes the notion that the conditional distribution of the future, given the past and present, depends only on the present. Specifically, for all $0 \leq s \leq t$ and all Borel sets $B \subset \mathbb{R}^d$,

$$\mathbb{P}(X_t \in B \mid \mathcal{F}_s) = \mathbb{P}(X_t \in B \mid X_s) \quad \text{almost surely.} \tag{1.33}$$

Equivalently, for any finite collection of times $0 \leq t_0 < t_1 < \cdots < t_n$ and any Borel sets $B_1, \ldots, B_n \subset \mathbb{R}^d$,

$$\mathbb{P}(X_{t_n} \in B_n \mid X_{t_0}, \ldots, X_{t_{n-1}}) = \mathbb{P}(X_{t_n} \in B_n \mid X_{t_{n-1}}), \tag{1.34}$$

which expresses the memoryless property of the process. The Markov property ensures that the evolution of the process can be described entirely in terms of **transition probabilities** from the current state to future states.

These transition probabilities are formalized using **transition kernels**. A transition kernel is a function $P : \mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d) \to [0, 1]$ satisfying two conditions:

1. For each fixed $x \in \mathbb{R}^d$, the mapping $B \mapsto P(x, B)$ is a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$
2. For each fixed $B \in \mathcal{B}(\mathbb{R}^d)$, the mapping $x \mapsto P(x, B)$ is measurable.

If (X_t) is a Markov process, its finite-dimensional distributions can be expressed recursively via these transition kernels. For $s < t$, let $P_{s,t}(x, B)$ denote the conditional probability

$$P_{s,t}(x, B) := \mathbb{P}(X_t \in B \mid X_s = x), \tag{1.35}$$

so that for any measurable function $f : \mathbb{R}^d \to \mathbb{R}$, the conditional expectation can be written as

$$\mathbb{E}[f(X_t) \mid X_s] = \int_{\mathbb{R}^d} f(y), P_{s,t}(X_s, dy) \quad \text{a.s.} \tag{1.36}$$

In the **time-homogeneous case**, the kernel depends only on the time difference $\tau = t - s$, so that $P_\tau(x, B) := P_{s,s+\tau}(x, B)$. The Chapman-Kolmogorov equation then governs the composition of these kernels: for $0 \le s < u < t$,

$$P_{s,t}(x, B) = \int_{\mathbb{R}^d} P_{s,u}(x, dy), P_{u,t}(y, B), \tag{1.37}$$

which expresses the semigroup property of the transition kernels and is fundamental for the analytic characterization of Markov processes via generators and semigroups.

Transition kernels allow one to construct Markov processes explicitly through **finite-dimensional distributions**. Given an initial distribution μ_0 on $\mathbb{R}^d$ and a family of transition kernels $(P_{s,t}) * 0 \le s \le t$ satisfying consistency and measurability, the joint distribution of $(X * t_0, \ldots, X_{t_n})$ is given recursively by

$$\mathbb{P}(X_{t_0} \in B_0, \ldots, X_{t_n} \in B_n) = \int_{B_0} \mu_0(dx_0) \int_{B_1} P_{t_0,t_1}(x_0, dx_1) \cdots \int_{B_n} P_{t_{n-1},t_n}(x_{n-1}, dx_n), \tag{1.38}$$

which guarantees consistency with the finite-dimensional marginals. This construction, combined with the Kolmogorov Extension Theorem, yields a rigorous realization of the Markov process on a suitable path space.

Definition 1.10 (Markov Process)**.** A process $(X_t)_{t\ge 0}$ adapted to $(\mathcal{F}_t)$ is *Markov* if

$$\mathbb{P}(X_t \in A \mid \mathcal{F}_s) = \mathbb{P}(X_t \in A \mid X_s), \quad s < t. \tag{1.39}$$

For a Markov process, $p_t(x, A) = \mathbb{P}(X_t \in A \mid X_0 = x)$ forms a semigroup of transition probabilities satisfying the Chapman–Kolmogorov equation:

$$p_{s+t}(x, A) = \int_E p_s(x, dy)\, p_t(y, A) \tag{1.40}$$

In the case of continuous-time processes with continuous paths, such as diffusion processes or Brownian motion, transition kernels can be described in terms of **densities** $p(t, x, y)$ with respect to the Lebesgue measure, satisfying

$$P_{s,t}(x, B) = \int_B p(t - s, x, y), dy. \tag{1.41}$$

The Chapman-Kolmogorov equation then becomes

$$p(t + s, x, z) = \int_{\mathbb{R}^d} p(t, x, y), p(s, y, z), dy, \tag{1.42}$$

which is a crucial analytic identity connecting probabilistic evolution with semigroup theory. This representation underlies the derivation of infinitesimal generators $\mathcal{L}$ and the associated Kolmogorov forward (Fokker–Planck) and backward equations, which in turn lead directly to the Feynman–Kac formula when coupled with suitable potential terms.

Thus, transition kernels provide a rigorous probabilistic structure for Markov processes, enabling recursive construction of finite-dimensional distributions, formulation of semigroup properties, and analytic characterization through generators, all of which are foundational in stochastic analysis and the study of parabolic partial differential equations via probabilistic methods.

1.8 Generators and Semigroups

Let $(X_t) * t \geq 0$ be a time-homogeneous Markov process with state space $\mathbb{R}^d$ and associated transition semigroup $(P_t) * t \geq 0$, where for any bounded measurable function $f : \mathbb{R}^d \to \mathbb{R}$,

$$P_t f(x) := \mathbb{E}[f(X_t) \mid X_0 = x] = \int_{\mathbb{R}^d} f(y), P_t(x, dy). \tag{1.43}$$

The family $(P_t)_{t \geq 0}$ satisfies the **semigroup property**: for all $s, t \geq 0$,

$$P_{t+s} = P_t P_s, \quad P_0 = I, \tag{1.44}$$

where I denotes the identity operator, and $P_t P_s f(x) := P_t(P_s f)(x)$. The semigroup property is a direct consequence of the Chapman-Kolmogorov equation for the transition kernels and encodes the memoryless evolution of the Markov process.

The infinitesimal generator $\mathcal{L}$ of the semigroup (P_t) is a linear operator defined on a domain $\mathcal{D}(\mathcal{L}) \subset C_b(\mathbb{R}^d)$, the space of bounded continuous functions, by the strong limit

$$\mathcal{L}f := \lim_{t \downarrow 0} \frac{P_t f - f}{t}, \quad f \in \mathcal{D}(\mathcal{L}), \tag{1.45}$$

where the limit is taken in the supremum norm. Intuitively, $\mathcal{L}f(x)$ measures the instantaneous rate of change of the expected value of $f(X_t)$ at time zero, conditioned on $X_0 = x$. For many diffusion processes with smooth coefficients, $\mathcal{L}$ takes the form of a second-order differential operator

$$\mathcal{L}f(x) = \sum_{i=1}^{d} b_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \tag{1.46}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is the drift vector and $a(x) = \sigma(x)\sigma(x)^T$ is the diffusion matrix associated with the stochastic differential equation

$$dX_t = b(X_t), dt + \sigma(X_t), dB_t. \tag{1.47}$$

The connection between semigroups and generators is expressed rigorously through the **Kolmogorov backward equation**: for $f \in \mathcal{D}(\mathcal{L})$,

$$\frac{\partial}{\partial t} P_t f = \mathcal{L} P_t f, \quad P_0 f = f \tag{1.48}$$

which is an abstract Cauchy problem in the Banach space $C_b(\mathbb{R}^d)$. Conversely, given a closed linear operator $\mathcal{L}$ satisfying appropriate conditions (such as the Hille–Yosida theorem), one can construct a strongly continuous semigroup (P_t) solving this equation. This correspondence provides a rigorous analytic foundation for representing the evolution of Markov processes in terms of their generators.

In addition, the generator characterizes the martingale property of functionals of the process. If $f \in \mathcal{D}(\mathcal{L})$, the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{L} f(X_s), ds \tag{1.49}$$

is a martingale with respect to the natural filtration of X_t. This representation is fundamental for deriving the **Feynman–Kac formula**, where one modifies the generator by a potential term $V(x)$ and considers exponential functionals of the resulting martingale. Specifically, if $\mathcal{L} - V$ denotes the generator with killing, the Feynman–Kac representation of the solution $u(t,x)$ to the parabolic PDE

$$\frac{\partial u}{\partial t} = \mathcal{L}u - V(x)u, \quad u(0,x) = f(x) \tag{1.50}$$

is given probabilistically by

$$u(t,x) = \mathbb{E}_x\Big[\exp\Big(-\int_0^t V(X_s), ds\Big) f(X_t)\Big] \tag{1.51}$$

where the expectation is taken over the paths of the Markov process generated by $\mathcal{L}$ starting from $X_0 = x$. Thus, the semigroup and generator formalism provides a rigorous bridge between stochastic processes, martingales, and the analytic theory of partial differential equations, forming the backbone of the Feynman–Kac methodology.

In later chapters, we shall show that the Feynman–Kac formula identifies P_t with the semigroup generated by the Schrödinger-type operator $A + V$, where V acts as a potential.

1.9 Conclusion

This chapter established the rigorous measure-theoretic framework for stochastic processes. The machinery of filtrations, conditional expectations, and semigroup generators forms the mathematical substrate on which the Feynman–Kac formula rests.

In the next chapter, we shall construct the Brownian motion as the canonical continuous Markov process on $C([0,\infty), \mathbb{R}^n)$, develop Itô integration and stochastic calculus, and demonstrate how these lead naturally to the Feynman–Kac representation of parabolic PDEs.

References for this chapter: Karatzas & Shreve (2014) [1], Øksendal (2013) [2], Revuz & Yor (1999) [3], Stroock (2010) [4], Doob (1984) [5], Billingsley (2017) [6] and Billingsley (2013) [7], Durrett (2019) [8].

Chapter 2

Brownian Motion and Stochastic Integrals

"Brownian motion is an inexhaustible source of beautiful mathematics and useful models."

— Adapted from P. Mörters & Y. Peres

2.1 Introduction

This chapter constructs Brownian motion as the canonical continuous Gaussian process, develops its fundamental pathwise and martingale properties, introduces the Itô stochastic integral for non-anticipative integrands, and proves Itô's formula together with basic consequences such as the martingale representation and quadratic variation. These tools are indispensable for the probabilistic representation of linear and semilinear parabolic equations and form the analytic backbone of the Feynman–Kac correspondence treated in subsequent chapters.

2.2 Definition and Basic Properties of Brownian Motion

Definition 2.1 (Brownian Motion / Wiener Process)**.** A stochastic process $(B_t)_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ is called a (standard, d-dimensional) *Brownian motion* if:

1. $B_0 = 0$ almost surely;
2. B has independent increments: for $0 \leq t_0 < t_1 < \cdots < t_n$, the increments $B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent;
3. For $0 \leq s < t$, the increment $B_t - B_s$ is Gaussian with mean 0 and covariance $(t-s)I_d$;
4. B has almost surely continuous paths: $\omega \mapsto t \mapsto B_t(\omega)$ is continuous.

From the Gaussian increment property it follows that for each fixed t, $B_t \sim \mathcal{N}(0, tI_d)$, and that all finite-dimensional distributions are multivariate normal with covariances $\mathbb{E}[B_s B_t^\top] = (s \wedge t)I_d$. The finite-dimensional covariance structure and consistency of the Gaussian family guarantee existence via the Kolmogorov extension theorem; continuity of paths is furnished by Kolmogorov's continuity criterion (see below).

Proposition 2.2 (Existence and Continuity). *Given the Gaussian family with mean zero and covariance* $\mathbb{E}[B_s^{(i)} B_t^{(j)}] = \delta_{ij}(s \wedge t)$*, there exists a probability measure on* $\Omega = C([0,\infty),\mathbb{R}^d)$ *making the coordinate process a Brownian motion. Moreover, the process admits almost surely Hölder continuous paths of any order* $\gamma \in (0,1/2)$*.*

Proof. Let $(\mu_{t_1,\ldots,t_n})$ denote the consistent family of dn-dimensional Gaussian measures corresponding to the finite-dimensional distributions with mean zero and covariance

$$\mathbb{E}[B_{t_k}^{(i)} B_{t_\ell}^{(j)}] = \delta_{ij}(t_k \wedge t_\ell), \quad 1 \leq i,j \leq d, ; 1 \leq k,\ell \leq n. \tag{2.1}$$

By the **Kolmogorov Extension Theorem**, there exists a probability measure $\mathbb{P}_0$ on $(\mathbb{R}^{d\times T}, \mathcal{B}(\mathbb{R}^{d\times T}))$ such that the coordinate projections $B_t(\omega) := \omega(t)$ have these finite-dimensional distributions. This provides a process with the correct Gaussian marginals but not yet with continuous paths.

Define for each $n \in \mathbb{N}$ the measure $\mathbb{P}_n$ on the space $D([0,n],\mathbb{R}^d)$ of càdlàg functions (right-continuous with left limits) restricted to $[0,n]$ equipped with the **Skorokhod topology**, induced by the metric d_S. Consider the sequence of probability measures $(\mathbb{P}_n)$ induced by $\mathbb{P}_0$ on the projections to $[0,n]$. For $\delta > 0$, define the modulus of continuity

$$w(B,\delta) := \sup_{s,t\in[0,n] \ |t-s|\leq\delta} |B_t - B_s|. \tag{2.2}$$

From the Gaussian moment estimates, for any $p \geq 2$ and $s,t \in [0,n]$

$$\mathbb{E}[|B_t - B_s|^p] = C_p|t-s|^{p/2}, \quad C_p = 2^{p/2}\Gamma\Big(\frac{p+1}{2}\Big)/\sqrt{\pi}. \tag{2.3}$$

Hence, by **Markov's inequality** and the union bound over dyadic partitions of $[0,n]$, for $\varepsilon > 0$ and sufficiently small $\delta > 0$,

$$\mathbb{P}_n(w(B,\delta) > \varepsilon) \leq \sum_k \frac{\mathbb{E}[|B_{t_{k+1}} - B_{t_k}|^p]}{\varepsilon^p} \leq C_p \frac{n}{\delta}\frac{\delta^{p/2}}{\varepsilon^p} = C_p n \varepsilon^{-p}\delta^{p/2-1}. \tag{2.4}$$

Choosing $p > 2$ ensures $p/2 - 1 > 0$. Therefore, for every $\eta > 0$, there exists $\delta > 0$ such that $\mathbb{P}_n(w(B,\delta) > \varepsilon) < \eta$. This establishes tightness of $(\mathbb{P}_n)$ in the Skorokhod space.

By **Prokhorov's theorem**, there exists a weakly convergent subsequence $\mathbb{P}_{n_k} \Rightarrow \mathbb{P}$ on $D([0,n],\mathbb{R}^d)$. The limiting measure $\mathbb{P}$ is concentrated on the subset of continuous paths because the modulus-of-continuity estimates imply that for almost all ω,

$$\lim_{\delta\to 0} w(\omega,\delta) = 0. \tag{2.5}$$

Thus, $\mathbb{P}$ is supported on $C([0,n],\mathbb{R}^d)$. By consistency across n, we can define a unique probability measure on $C([0,\infty),\mathbb{R}^d)$ such that the canonical projections $B_t(\omega) := \omega(t)$ have the desired Gaussian finite-dimensional distributions and are continuous almost surely.

Finally, the **Hölder continuity** of order $\gamma \in (0,1/2)$ follows from the Kolmogorov continuity criterion. Indeed, for p large enough,

$$\mathbb{E}[|B_t - B_s|^p] \le C_p|t-s|^{p/2} \tag{2.6}$$

so choosing $p > 1/(1/2-\gamma)$ ensures that

$$\exists K(\omega) < \infty : |B_t(\omega) - B_s(\omega)| \le K(\omega)|t-s|^\gamma, \quad \forall s,t \in [0,T], \tag{2.7}$$

almost surely. This shows that the paths are almost surely Hölder continuous of any order $\gamma < 1/2$, completing the construction of (d)-dimensional Brownian motion on $\Omega = C([0,\infty), \mathbb{R}^d)$. □

Remark 2.3. Brownian motion is a mean-zero martingale with respect to its natural (completed, right-continuous) filtration $(\mathcal{F}_t^B)$. It has stationary and independent increments, is self-similar of index $1/2$ i.e.,

$$(c^{-1/2}B_{ct})_{t\ge 0} \overset{d}{=} (B_t)_{t\ge 0} \tag{2.8}$$

and is recurrent in dimension $d = 1, 2$ and transient for $d \ge 3$.

2.3 Quadratic Variation and Semimartingale Structure

For a continuous path $\omega \in C([0,T], \mathbb{R})$ and a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, define the quadratic variation along Π by

$$Q(\Pi, \omega) := \sum_{k=1}^{n} \big(\omega(t_k) - \omega(t_{k-1})\big)^2. \tag{2.9}$$

For Brownian motion B, the following fundamental result holds.

Theorem 2.4 (Quadratic Variation of Brownian Motion)**.** *Let* (Π^m) *be a sequence of partitions of* $[0,T]$ *with mesh* $|\Pi^m| \to 0$*. Then, almost surely,*

$$\lim_{m\to\infty} Q(\Pi^m, B) = T. \tag{2.10}$$

More generally, for d-dimensional Brownian motion,

$$\lim_{|\Pi|\to 0} \sum_k |B_{t_k} - B_{t_{k-1}}|^2 = dT \quad a.s. \tag{2.11}$$

Proof. Let $B = (B_t)_{t\in[0,T]}$ be a standard one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions. Consider a sequence of partitions $\Pi^m = 0 = t_0^m < t_1^m < \cdots < t_{n(m)}^m = T$ with mesh $|\Pi^m| = \max_k(t_k^m - t_{k-1}^m) \to 0$ as $m \to \infty$. Define the quadratic sum along the partition as

$$Q(\Pi^m, B) := \sum_{k=1}^{n(m)} (B_{t_k^m} - B_{t_{k-1}^m})^2. \tag{2.12}$$

We aim to prove that $Q(\Pi^m, B) \to T$ almost surely as $m \to \infty$. Begin by computing the first two moments of $Q(\Pi^m, B)$. Using the independence and Gaussianity of increments, for each k,

$$\mathbb{E}[(B_{t_k^m} - B_{t_{k-1}^m})^2] = t_k^m - t_{k-1}^m, \tag{2.13}$$

so by linearity of expectation,

$$\mathbb{E}[Q(\Pi^m, B)] = \sum_{k=1}^{n(m)} (t_k^m - t_{k-1}^m) = T. \tag{2.14}$$

Next, we compute the variance. Again using independence of increments and the fact that for a Gaussian $X \sim \mathcal{N}(0, \sigma^2)$, $\mathrm{Var}(X^2) = 2\sigma^4$, we obtain

$$\mathrm{Var}(Q(\Pi^m, B)) = \sum_{k=1}^{n(m)} \mathrm{Var}((B_{t_k^m} - B_{t_{k-1}^m})^2) = \sum_{k=1}^{n(m)} 2(t_k^m - t_{k-1}^m)^2 = 2\sum_{k=1}^{n(m)} (t_k^m - t_{k-1}^m)^2. \tag{2.15}$$

Since $|\Pi^m| \to 0$, we have

$$\sum_{k=1}^{n(m)} (t_k^m - t_{k-1}^m)^2 \le |\Pi^m| \sum_{k=1}^{n(m)} (t_k^m - t_{k-1}^m) = |\Pi^m|, T \to 0 \quad \text{as } m \to \infty. \tag{2.16}$$

Thus, $\mathrm{Var}(Q(\Pi^m, B)) \to 0$. By Chebyshev's inequality, for any $\varepsilon > 0$,

$$\mathbb{P}(|Q(\Pi^m, B) - T| > \varepsilon) \le \frac{\mathrm{Var}(Q(\Pi^m, B))}{\varepsilon^2} \to 0. \tag{2.17}$$

This shows convergence **in probability**:

$$Q(\Pi^m, B) \xrightarrow{\mathbb{P}} T. \tag{2.18}$$

To upgrade to almost sure convergence, apply a standard **subsequence argument**. Choose a subsequence (Π^{m_j}) with $|\Pi^{m_j}| \le 2^{-j}$. Then

$$\sum_{j=1}^{\infty} \mathbb{P}(|Q(\Pi^{m_j}, B) - T| > \varepsilon) \le \sum_{j=1}^{\infty} \frac{2T|\Pi^{m_j}|}{\varepsilon^2} \le \sum_{j=1}^{\infty} \frac{2T}{\varepsilon^2} 2^{-j} = \frac{4T}{\varepsilon^2} < \infty. \tag{2.19}$$

By the **Borel–Cantelli lemma**, almost surely $|Q(\Pi^{m_j}, B) - T| \to 0$ along this subsequence. Since the partitions are nested or can be refined arbitrarily, one can pass to the full sequence Π^m and conclude that

$$\lim_{m\to\infty} Q(\Pi^m, B) = T \quad \text{a.s.} \tag{2.20}$$

This establishes the one-dimensional case. For d-dimensional Brownian motion $B = (B^{(1)}, \ldots, B^{(d)})$ with independent components, we have

$$\sum_k |B_{t_k} - B_{t_{k-1}}|^2 = \sum_{i=1}^{d} \sum_k (B_{t_k}^{(i)} - B_{t_{k-1}}^{(i)})^2. \tag{2.21}$$

By the previous argument, for each (i), $\sum_k (B_{t_k}^{(i)} - B_{t_{k-1}}^{(i)})^2 \to T$ almost surely. Since a finite union of almost sure events is almost sure, we obtain

$$\lim_{|\Pi|\to 0} \sum_k |B_{t_k} - B_{t_{k-1}}|^2 = \sum_{i=1}^{d} T = dT \quad \text{a.s.} \tag{2.22}$$

Hence, the quadratic variation of d-dimensional Brownian motion along any sequence of partitions with vanishing mesh exists almost surely and equals dT. □

This nontrivial pathwise property implies that Brownian motion is not of bounded variation and that stochastic calculus must be developed in the semimartingale framework; Brownian motion is a continuous local martingale with predictable quadratic variation $\langle B\rangle_t = t$.

2.4 Simple Integrands and the Itô Integral

To define stochastic integration, begin with elementary (simple) adapted processes. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions (completeness, right-continuity), and let $B = (B_t)_{t\in[0,T]}$ be a standard Brownian motion adapted to $(\mathcal{F}_t)$. To construct the stochastic integral rigorously, one begins with simple (elementary) adapted processes. These serve as building blocks, because the integral can be extended by continuity to a wider class of integrands.

A **simple adapted process** H on $[0, T]$ is defined by

$$H_t(\omega) = \sum_{k=0}^{n-1} H_k(\omega)\mathbf{1}_{(t_k,t_{k+1}]}(t), \tag{2.23}$$

where $0 = t_0 < t_1 < \cdots < t_n = T$ is a partition of the interval $[0, T]$, and each H_k is $\mathcal{F}_{t_k}$-measurable and square-integrable, that is, $\mathbb{E}[H_k^2] < \infty$. Progressive measurability of H is automatically satisfied for such processes, since each H_k depends only on the past up to time t_k. The simple structure ensures that the stochastic integral can be defined in a **Riemann–Stieltjes-like manner**, using the Brownian increments over each subinterval of the partition.

Specifically, the **stochastic integral** of a simple process H with respect to B is defined as

$$\int_0^T H_s\, dB_s := \sum_{k=0}^{n-1} H_k(B_{t_{k+1}} - B_{t_k}). \tag{2.24}$$

Here, each increment $B_{t_{k+1}} - B_{t_k}$ is independent of $\mathcal{F}_{t_k}$, which allows one to rigorously compute moments of the integral. In particular, the **Itô isometry** asserts that

$$\mathbb{E}\Big[\Big(\int_0^T H_s, dB_s\Big)^2\Big] = \sum_{k=0}^{n-1} \mathbb{E}[H_k^2](t_{k+1} - t_k) = \mathbb{E}\Big[\int_0^T H_s^2, ds\Big]. \tag{2.25}$$

The Itô isometry is a **fundamental identity**: it shows that the map $H \mapsto \int_0^T H_s, dB_s$ is a linear isometry from the space of simple adapted processes, equipped with the norm

$$|H|_{\mathcal{H}^2} := \Big(\mathbb{E}[\int_0^T H_s^2 ds]\Big)^{1/2}, \tag{2.26}$$

into $L^2(\Omega)$. This isometry provides the rigorous foundation to extend the definition of the stochastic integral from simple processes to the full Hilbert space

$$\mathcal{H}^2([0, T]) := \Big\{H \text{ progressively measurable } : \mathbb{E}\Big[\int_0^T H_s^2\, ds\Big] < \infty\Big\}. \tag{2.27}$$

By approximating an arbitrary $H \in \mathcal{H}^2([0,T])$ in the $\mathcal{H}^2$-norm by a sequence of simple processes $(H^{(n)})$, and using the Itô isometry to pass to the limit in $L^2(\Omega)$, one obtains a **well-defined, continuous linear operator**

$$I : \mathcal{H}^2([0,T]) \to L^2(\Omega), \quad H \mapsto \int_0^T H_s \, dB_s \tag{2.28}$$

which preserves the isometry property. Moreover, for each $t \in [0,T]$, the process

$$M_t := \int_0^t H_s, dB_s \tag{2.29}$$

is a **continuous square-integrable martingale** with respect to $(\mathcal{F}_t)$, and its predictable quadratic variation is given by

$$\langle M \rangle_t = \int_0^t H_s^2, ds. \tag{2.30}$$

Thus, starting from simple adapted processes, the Itô isometry allows a rigorous extension of stochastic integration to all square-integrable progressively measurable processes, producing a continuous martingale with well-defined quadratic variation.

Proposition 2.5 (Itô Isometry). *For any simple adapted H as above,*

$$\mathbb{E}\Big[\Big(\int_0^T H_s \, dB_s\Big)^2\Big] = \mathbb{E}\Big[\int_0^T H_s^2 \, ds\Big]. \tag{2.31}$$

Proof. Let $B = (B_t)_{t\in[0,T]}$ be a standard one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions. Consider a simple adapted process H of the form

$$H_t(\omega) = \sum_{k=0}^{n-1} H_k(\omega)\mathbf{1}_{(t_k,t_{k+1}]}(t), \tag{2.32}$$

where $0 = t_0 < t_1 < \cdots < t_n = T$ is a partition of $[0,T]$ and each H_k is $\mathcal{F}_{t_k}$-measurable and square-integrable: $\mathbb{E}[H_k^2] < \infty$. The stochastic integral of H with respect to B is defined as

$$\int_0^T H_s, dB_s := \sum_{k=0}^{n-1} H_k(B_{t_{k+1}} - B_{t_k}). \tag{2.33}$$

We aim to compute $\mathbb{E}\Big[\big(\int_0^T H_s, dB_s\big)^2\Big]$. Expanding the square gives

$$\Big(\int_0^T H_s, dB_s\Big)^2 = \sum_{k=0}^{n-1} H_k^2(B_{t_{k+1}} - B_{t_k})^2 + 2 \sum_{0\le k<\ell\le n-1} H_k H_\ell(B_{t_{k+1}} - B_{t_k})(B_{t_{\ell+1}} - B_{t_\ell}). \tag{2.34}$$

Taking expectations, we have

$$\mathbb{E}\Big[\Big(\int_0^T H_s, dB_s\Big)^2\Big] = \sum_{k=0}^{n-1} \mathbb{E}[H_k^2(B_{t_{k+1}} - B_{t_k})^2] + 2 \sum_{0\le k<\ell\le n-1} \mathbb{E}[H_k H_\ell(B_{t_{k+1}} - B_{t_k})(B_{t_{\ell+1}} - B_{t_\ell})]. \tag{2.35}$$

Since H_k is $\mathcal{F}_{t_k}$-measurable and $B_{t_{\ell+1}} - B_{t_\ell}$ is independent of $\mathcal{F}_{t_\ell} \supset \mathcal{F}_{t_k}$ for $\ell > k$, the cross terms vanish:

$$\mathbb{E}[H_k H_\ell (B_{t_{k+1}} - B_{t_k})(B_{t_{\ell+1}} - B_{t_\ell})] = \mathbb{E}[H_k H_\ell (B_{t_{k+1}} - B_{t_k}) \mathbb{E}[B_{t_{\ell+1}} - B_{t_\ell} \mid \mathcal{F}_{t_\ell}]] = 0. \quad (2.36)$$

Hence, only the diagonal terms remain:

$$\mathbb{E}\Big[\Big(\int_0^T H_s, dB_s\Big)^2\Big] = \sum_{k=0}^{n-1} \mathbb{E}[H_k^2 (B_{t_{k+1}} - B_{t_k})^2]. \quad (2.37)$$

For each k, since $B_{t_{k+1}} - B_{t_k} \sim \mathcal{N}(0, t_{k+1} - t_k)$ and is independent of H_k, we have

$$\mathbb{E}[H_k^2 (B_{t_{k+1}} - B_{t_k})^2] = \mathbb{E}[H_k^2]\, \mathbb{E}[(B_{t_{k+1}} - B_{t_k})^2] = \mathbb{E}[H_k^2](t_{k+1} - t_k). \quad (2.38)$$

Summing over k, we obtain

$$\mathbb{E}\Big[\Big(\int_0^T H_s, dB_s\Big)^2\Big] = \sum_{k=0}^{n-1} \mathbb{E}[H_k^2](t_{k+1} - t_k) = \mathbb{E}\Big[\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} H_k^2, ds\Big] = \mathbb{E}\Big[\int_0^T H_s^2, ds\Big]. \quad (2.39)$$

This completes the proof: the Itô isometry holds for all simple adapted processes. It provides a natural Hilbert space isometry between the space of square-integrable adapted processes and the space of $L^2(\Omega)$ martingales generated by stochastic integrals with respect to Brownian motion. □

Define the space $\mathcal{H}^2([0,T])$ as the set of progressively measurable processes H satisfying

$$\mathbb{E}[\int_0^T H_s^2 ds] < \infty \quad (2.40)$$

By Itô isometry, simple processes are dense in $\mathcal{H}^2$ (in the norm induced by the right-hand side), so one extends $\int_0^T H_s\, dB_s$ by isometry to all $H \in \mathcal{H}^2$. The resulting integral is a continuous linear map from $\mathcal{H}^2$ to $L^2(\Omega)$ and produces a continuous square-integrable martingale $(\int_0^t H_s dB_s)_{t\in[0,T]}$ with predictable quadratic variation $\int_0^t H_s^2 ds$. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, and let $B = (B_t)_{t\in[0,T]}$ be a standard Brownian motion adapted to $(\mathcal{F}_t)$. Define the Hilbert space

$$\mathcal{H}^2([0,T]) := \Big\{H : [0,T]\times\Omega \to \mathbb{R} \,\Big|\, H \text{ is progressively measurable, } |H|_{\mathcal{H}^2}^2 := \mathbb{E}\Big[\int_0^T H_s^2\, ds\Big] < \infty\Big\}. \quad (2.41)$$

Here, progressive measurability means that for each $t \in [0,T]$, the map $(s,\omega) \mapsto H_s(\omega)\mathbf{1}_{[0,t]}(s)$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable. The norm

$$|H|_{\mathcal{H}^2}^2 := \mathbb{E}[\int_0^T H_s^2 ds] \quad (2.42)$$

makes $\mathcal{H}^2([0,T])$ a complete Hilbert space. Let $\mathcal{S}$ denote the set of **simple adapted processes**, i.e., processes of the form

$$H_t(\omega) = \sum_{k=0}^{n-1} H_k(\omega)\mathbf{1}_{(t_k,t_{k+1}]}(t), \quad (2.43)$$

where $(t_k)*k = 0^n$ is a partition of $[0,T]$ and each H_k is $\mathcal{F}*t_k$-measurable and square-integrable.

The Itô isometry, proved for simple processes, states that for $H \in \mathcal{S}$

$$\mathbb{E}\Big[\Big(\int_0^T H_s, dB_s\Big)^2\Big] = |H|^2_{\mathcal{H}^2}. \tag{2.44}$$

This equality shows that the map

$$I : \mathcal{S} \to L^2(\Omega), \quad H \mapsto \int_0^T H_s, dB_s, \tag{2.45}$$

is a **linear isometry**. Hence $\mathcal{S}$ is a dense subspace of $\mathcal{H}^2([0,T])$ with respect to the $\mathcal{H}^2$ norm: for any $H \in \mathcal{H}^2([0,T])$ and $\varepsilon > 0$, there exists a simple process $H^\varepsilon \in \mathcal{S}$ such that

$$|H - H^\varepsilon|_{\mathcal{H}^2} < \varepsilon. \tag{2.46}$$

Using the completeness of $L^2(\Omega)$ and the isometry, we extend (I) uniquely to all of $\mathcal{H}^2([0,T])$ by setting

$$\int_0^T H_s, dB_s := \lim_{n\to\infty} \int_0^T H_s^{(n)}, dB_s, \tag{2.47}$$

where $H^{(n)} \in \mathcal{S}$ is a sequence of simple processes converging to H in $\mathcal{H}^2$. The limit exists in $L^2(\Omega)$ and is independent of the approximating sequence due to the isometry. This shows that the stochastic integral

$$I : \mathcal{H}^2([0,T]) \to L^2(\Omega), \quad H \mapsto \int_0^T H_s, dB_s, \tag{2.48}$$

is a **continuous linear operator**, with operator norm (1).

For each $t \in [0,T]$, define the process

$$M_t := \int_0^t H_s, dB_s := I(H\mathbf{1}_{[0,t]}). \tag{2.49}$$

Then $(M_t)_{t\in[0,T]}$ is a **square-integrable martingale** with respect to $(\mathcal{F}_t)$, because for $s < t$,

$$\mathbb{E}[M_t \mid \mathcal{F}_s] = \mathbb{E}\Big[\int_0^s H_u, dB_u + \int_s^t H_u, dB_u, \Big|, \mathcal{F}_s\Big] = \int_0^s H_u, dB_u = M_s, \tag{2.50}$$

using the independence of the increment $B_t - B_s$ from $\mathcal{F}_s$ and the $\mathcal{F}_s$-measurability of H_u for $u \le s$. Moreover, the quadratic variation of M is given by the predictable process

$$\langle M\rangle_t = \int_0^t H_s^2, ds, \tag{2.51}$$

because for simple processes it holds by definition, and the continuity of the quadratic variation under L^2-limits ensures that it extends to all $H \in \mathcal{H}^2$. Consequently, the stochastic integral produces a **continuous, square-integrable martingale** with predictable quadratic variation $\int_0^t H_s^2 ds$, and the Itô isometry holds in full generality:

$$\mathbb{E}[M_T^2] = \mathbb{E}\Big[\int_0^T H_s^2, ds\Big]. \tag{2.52}$$

2.5 Itô's Formula

Itô's formula generalizes the classical chain rule to Itô processes and is the cornerstone of stochastic calculus.

Theorem 2.6 (Itô's Formula)**.** *Let X be an Itô process in $\mathbb{R}^d$:*

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dB_s, \tag{2.53}$$

where b and σ are progressively measurable with suitable integrability, $b : \Omega \times [0,T] \to \mathbb{R}^d$, $\sigma : \Omega \times [0,T] \to \mathbb{R}^{d\times m}$, and B is m-dimensional Brownian motion. Let $f : \mathbb{R}^d \to \mathbb{R}$ be C^2 with bounded derivatives up to second order. Then, almost surely for all $t \in [0,T]$,

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s)\, dX_s^{(i)} + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\, d\langle X^{(i)}, X^{(j)}\rangle_s \\ &= f(X_0) + \int_0^t \nabla f(X_s)\cdot b_s\, ds + \int_0^t \nabla f(X_s)\sigma_s\, dB_s + \frac{1}{2}\int_0^t \operatorname{tr}\big(\sigma_s^\top D^2 f(X_s)\sigma_s\big)\, ds. \end{aligned} \tag{2.54}$$

Proof. Let $X = (X_t) * t \in [0,T]$ be an Itô process in $\mathbb{R}^d$ defined by

$$X_t = X_0 + \int_0^t b_s, ds + \int_0^t \sigma_s, dB_s, \tag{2.55}$$

where $b : \Omega \times [0,T] \to \mathbb{R}^d$ and $\sigma : \Omega \times [0,T] \to \mathbb{R}^{d\times m}$ are progressively measurable and satisfy the integrability conditions

$$\mathbb{E}\left[\int_0^T |b_s|, ds\right] < \infty, \quad \mathbb{E}\left[\int_0^T |\sigma_s|^2, ds\right] < \infty. \tag{2.56}$$

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice continuously differentiable C^2 function with bounded first and second derivatives. Our goal is to show that almost surely, for every $t \in [0,T]$,

$$\begin{aligned} f(X_t) &= f(X_0) + \sum * i = 1^d \int_0^t \frac{\partial f}{\partial x_i}(X_s), dX_s^{(i)} * \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s), d\langle X^{(i)}, X^{(j)}\rangle_s \\ &= f(X_0) + \int_0^t \nabla f(X_s)\cdot b_s, ds * \int_0^t \nabla f(X_s)\sigma_s, dB_s * \frac{1}{2}\int_0^t \operatorname{tr}\big(\sigma_s^\top D^2 f(X_s)\sigma_s\big), ds. \end{aligned} \tag{2.57}$$

We begin with the one-dimensional case $d = 1$, which captures the essential ideas. Suppose

$$X_t = X_0 + \int_0^t b_s, ds + \int_0^t \sigma_s, dB_s, \tag{2.58}$$

and let $f \in C^2(\mathbb{R})$ with bounded f' and f''. Let $\Pi^m = 0 = t_0^m < t_1^m < \cdots < t_{n_m}^m = t$ be a sequence of partitions of $[0,t]$ such that the mesh

$$|\Pi^m| = \max_k (t_k^m - t_{k-1}^m) \to 0 \tag{2.59}$$

as $m \to \infty$. For each m, we write the telescoping sum:

$$f(X_t) - f(X_0) = \sum_{k=1}^{n_m} \left(f(X_{t_k^m}) - f(X_{t_{k-1}^m}) \right). \tag{2.60}$$

Using Taylor's theorem with remainder for each term, for some ξ_k^m between $X_{t_{k-1}^m}$ and $X_{t_k^m}$, we have

$$f(X_{t_k^m}) - f(X_{t_{k-1}^m}) = f'(X_{t_{k-1}^m})\Delta X_k^m + \frac{1}{2} f''(\xi_k^m)(\Delta X_k^m)^2, \tag{2.61}$$

where $\Delta X_k^m := X_{t_k^m} - X_{t_{k-1}^m}$. Hence,

$$f(X_t) - f(X_0) = \sum_{k=1}^{n_m} f'(X_{t_{k-1}^m})\Delta X_k^m + \frac{1}{2} \sum_{k=1}^{n_m} f''(\xi_k^m)(\Delta X_k^m)^2. \tag{2.62}$$

We decompose each increment ΔX_k^m using the Itô process definition:

$$\Delta X_k^m = \int_{t_{k-1}^m}^{t_k^m} b_s, ds + \int_{t_{k-1}^m}^{t_k^m} \sigma_s, dB_s. \tag{2.63}$$

We analyze each term separately.

Step 1: The first sum

The first term

$$\sum_{k=1}^{n_m} f'(X_{t_{k-1}^m})\Delta X_k^m \tag{2.64}$$

is a Riemann–Stieltjes-type sum which, as $|\Pi^m| \to 0$, converges in $L^2(\Omega)$ and almost surely to

$$\int_0^t f'(X_s), dX_s. \tag{2.65}$$

By the definition of the stochastic integral,

$$\int_0^t f'(X_s), dX_s = \int_0^t f'(X_s) b_s, ds + \int_0^t f'(X_s)\sigma_s, dB_s. \tag{2.66}$$

Step 2: The second sum

We now study the sum

$$\frac{1}{2} \sum_{k=1}^{n_m} f''(\xi_k^m)(\Delta X_k^m)^2. \tag{2.67}$$

We note that $(\Delta X_k^m)^2$ can be expanded as

$$(\Delta X_k^m)^2 = \left(\int_{t_{k-1}^m}^{t_k^m} b_s, ds \right)^2 * 2 \left(\int_{t_{k-1}^m}^{t_k^m} b_s, ds \right) \left(\int_{t_{k-1}^m}^{t_k^m} \sigma_s, dB_s \right) * \left(\int_{t_{k-1}^m}^{t_k^m} \sigma_s, dB_s \right)^2. \tag{2.68}$$

The first term satisfies

$$\sum_{k=1}^{n_m} f''(\xi_k^m) \left(\int_{t_{k-1}^m}^{t_k^m} b_s, ds \right)^2 \to 0 \quad \text{in } L^1(\Omega), \tag{2.69}$$

because the mesh tends to zero and b_s is integrable. Similarly, the mixed term with b_s and σ_s converges to zero in $L^1(\Omega)$ since the Itô integral has zero mean and bounded variance.

The only term contributing in the limit is the last one:

$$\sum_{k=1}^{n_m} f''(\xi_k^m) \left(\int_{t_{k-1}^m}^{t_k^m} \sigma_s, dB_s \right)^2. \tag{2.70}$$

By Itô isometry,

$$\mathbb{E}\left[\left(\int_{t_{k-1}^m}^{t_k^m} \sigma_s, dB_s \right)^2 \right] = \mathbb{E}\left[\int_{t_{k-1}^m}^{t_k^m} \sigma_s^2, ds \right]. \tag{2.71}$$

Thus, using the quadratic variation property of Brownian motion and the continuity of f'',

$$\sum_{k=1}^{n_m} f''(\xi_k^m) \left(\int_{t_{k-1}^m}^{t_k^m} \sigma_s, dB_s \right)^2 \to \int_0^t f''(X_s)\sigma_s^2, ds \quad \text{a.s. and in } L^1(\Omega). \tag{2.72}$$

Step 3: Combining limits

Putting everything together and passing to the limit $|\Pi^m| \to 0$, we obtain almost surely

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)b_s, ds + \int_0^t f'(X_s)\sigma_s, dB_s * \frac{1}{2} \int_0^t f''(X_s)\sigma_s^2, ds. \tag{2.73}$$

This completes the proof for the one-dimensional case.

Step 4: The multidimensional case

Now let $X_t = (X_t^{(1)}, \dots, X_t^{(d)})^\top$ be an Itô process in $\mathbb{R}^d$ satisfying

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_s^{(i)}, ds + \sum_{\alpha=1}^{m} \int_0^t \sigma_s^{(i\alpha)}, dB_s^{(\alpha)}. \tag{2.74}$$

The covariance structure of X is described by its quadratic covariation matrix process

$$\langle X^{(i)}, X^{(j)} \rangle_t = \int_0^t \sum_{\alpha=1}^{m} \sigma_s^{(i\alpha)} \sigma_s^{(j\alpha)}, ds. \tag{2.75}$$

Applying the one-dimensional Itô formula to $f(X_t)$ as a function of multiple variables and using the multidimensional Taylor expansion, we obtain

$$f(X_t) - f(X_0) = \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s), dX_s^{(i)} * \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s), d\langle X^{(i)}, X^{(j)} \rangle_s. \tag{2.76}$$

Substituting $dX_s^{(i)} = b_s^{(i)} ds + \sum_{\alpha=1}^m \sigma_s^{(i\alpha)} dB_s^{(\alpha)}$, we have

$$\begin{aligned} f(X_t) - f(X_0) &= \int_0^t \nabla f(X_s) \cdot b_s, ds * \int_0^t \nabla f(X_s)\sigma_s, dB_s * \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \sum_{\alpha=1}^m \sigma_s^{(i\alpha)} \sigma_s^{(j\alpha)}, ds \\ &= \int_0^t \nabla f(X_s) \cdot b_s, ds * \int_0^t \nabla f(X_s)\sigma_s, dB_s * \frac{1}{2} \int_0^t \mathrm{tr}(\sigma_s^\top D^2 f(X_s)\sigma_s), ds. \end{aligned} \tag{2.77}$$

Hence, we have established that for any C^2 function f with bounded derivatives and for any Itô process X,

$$\boxed{f(X_t) = f(X_0) + \int_0^t \nabla f(X_s) \cdot b_s, ds * \int_0^t \nabla f(X_s)\sigma_s, dB_s * \frac{1}{2} \int_0^t \mathrm{tr}(\sigma_s^\top D^2 f(X_s)\sigma_s), ds.} \tag{2.78}$$

This completes the proof of **Itô's formula**. □

Itô's formula yields as a corollary that for $f(x) = x^2$ and one-dimensional Brownian motion,

$$B_t^2 = 2\int_0^t B_s \, dB_s + t, \tag{2.79}$$

exhibiting explicitly the quadratic variation term t.

2.5.1 Itô's Lemma for Jump–Diffusion (Poisson) Processes

Let $(\Omega, \mathcal{F}, (\mathcal{F} * t) * t \geq 0, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. Let $W = (W_t)_{t\geq 0}$ be an r-dimensional $\mathcal{F}_t$-Brownian motion and let $(Z, \mathcal{Z})$ be a measurable mark space equipped with a σ-finite measure ν (the Lévy / intensity measure). Let $N(dt, dz)$ denote a Poisson random measure on $[0, \infty) \times Z$ with compensator $\nu(dz), dt$, and write the compensated Poisson random measure as

$$\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)\, dt \tag{2.80}$$

We assume all processes are càdlàg and adapted, and all integrals below are well defined under stated integrability hypotheses.

Consider an $\mathbb{R}^d$-valued Itô semimartingale with jumps given in differential form by

$$\boxed{dX_t \;=\; b(t, X_{t-}), dt \;+\; \sigma(t, X_{t-}), dW_t \;+\; \int_Z c(t, X_{t-}, z), \tilde{N}(dt, dz) \;+\; \int_Z \bar{c}(t, X_{t-}, z), N(dt, dz)}, \tag{2.81}$$

or, in the usual consolidated form (taking c to encode all jumps and using the compensated measure for the martingale part),

$$\begin{aligned} X_t \;=\; X_0 \;+\; \int_0^t b(s, X_{s-})\, ds \;+\; \int_0^t \sigma(s, X_{s-})\, dW_s \;+\; \int_0^t \int_Z c(s, X_{s-}, z)\, \tilde{N}(ds, dz) \\ +\; \int_0^t \int_Z \kappa(s, X_{s-}, z)\, \nu(dz)\, ds \end{aligned} \tag{2.82}$$

where one may take κ equal to the predictable compensator part of large jumps (the two representations are equivalent under suitable decomposition of the jump kernel). For clarity in the Itô formula below we will use the representation with a single jump kernel $c(t, x, z)$ and the compensated measure $\tilde{N}$; any finite-variation predictable jump part can be accommodated by adding its compensator to the drift b.

Assume the coefficients satisfy the integrability / growth conditions needed to make all integrals finite: for example, for each $T > 0$,

$$\mathbb{E}\Big[\int_0^T (|b(s, X_{s-})| + |\sigma(s, X_{s-})|^2), ds + \int_0^T \int_Z (|c(s, X_{s-}, z)|^2 \wedge 1), \nu(dz), ds\Big] < \infty, \tag{2.83}$$

and c, σ, b are progressively measurable in the usual sense. Let $f : \mathbb{R}^d \to \mathbb{R}$ be twice continuously differentiable $f \in C^2(\mathbb{R}^d)$ and assume for the theorem that $\partial_i f, \partial_{ij} f$ have growth such that all expectations and integrals below are finite (e.g. bounded derivatives or polynomial growth controlled by the integrability of X and c).

The Itô formula for jump processes (Itô–Tanaka / Itô for semimartingales with jumps) states that for every $t \geq 0$, almost surely,

$$\begin{aligned} f(X_t) = f(X_0)& \\ &+ \int_0^t \nabla f(X_{s-}) \cdot b(s, X_{s-}), ds \;+\; \int_0^t \nabla f(X_{s-})\sigma(s, X_{s-})\, dW_s \\ &+ \tfrac{1}{2} \int_0^t \operatorname{tr}(\sigma(s, X_{s-})^\top D^2 f(X_{s-}), \sigma(s, X_{s-}))\, ds \\ &+ \int_0^t \int_Z (f(X_{s-} + c(s, X_{s-}, z)) - f(X_{s-}))\, \tilde{N}(ds, dz) \\ &+ \int_0^t \int_Z \Big(f(X_{s-} + c(s, X_{s-}, z)) - f(X_{s-}) - \nabla f(X_{s-}) \cdot c(s, X_{s-}, z)\Big)\, \nu(dz)\, ds. \end{aligned} \tag{Itô jump formula}$$

Equivalently, grouping the predictable drift terms,

$$\begin{aligned} f(X_t) = f(X_0)& \\ &+ \int_0^t \nabla f(X_{s-}) \cdot b(s, X_{s-})\, ds + \tfrac{1}{2} \int_0^t \operatorname{tr}(\sigma^\top D^2 f, \sigma)(s, X_{s-})\, ds \\ &+ \int_0^t \int_Z \Big(f(X_{s-} + c(s, X_{s-}, z)) - f(X_{s-}) - \nabla f(X_{s-}) \cdot c(s, X_{s-}, z)\Big)\, \nu(dz)\, ds \\ &+ \int_0^t \nabla f(X_{s-})\sigma(s, X_{s-})\, dW_s + \int_0^t \int_Z (f(X_{s-} + c(s, X_{s-}, z)) - f(X_{s-}))\, \tilde{N}(ds, dz) \end{aligned} \tag{2.84}$$

In this formulation the last two terms are local martingales: the stochastic integral against W is a continuous local martingale and the compensated jump integral is a purely discontinuous local martingale. The third line is the predictable finite-variation correction coming from the jump compensator; the term inside the compensator integral is precisely the second-order (and higher) remainder from the Taylor expansion of f around X_{s-}.

The formula can be written in the semimartingale language as well. If $X = X^c + X^d$ with continuous local martingale part X^c and purely discontinuous part X^d, and $\Delta X_s = X_s - X_{s-}$, then for $f \in C^2$ one has (pathwise, a.s.)

$$\begin{aligned} f(X_t) - f(X_0) = \int_0^t \nabla f(X_{s-})\, dX_s + \tfrac{1}{2} \sum_{i,j=1}^{d} \int_0^t \partial_{ij} f(X_{s-})\, d\langle X^{c,i}, X^{c,j}\rangle_s \\ + \sum_{0<s\le t} \Big(f(X_s) - f(X_{s-}) - \nabla f(X_{s-}) \cdot \Delta X_s \Big) \end{aligned} \tag{2.85}$$

where the sum over jumps is either written explicitly as a (random) sum or expressed by an integral against the jump measure N and its compensator as above. The two representations coincide under the identification of the quadratic covariation $\langle X^{c,i}, X^{c,j}\rangle$ with $\int_0^{\cdot} (\sigma\sigma^\top)^{ij}(s, X_{s-}), ds$ and of the jump-sum with the integrals against N and ν.

Hypotheses ensuring the validity of the formula are: $f \in C^2$ with growth so that

$$|f(x+y) - f(x) - \nabla f(x) \cdot y| \le C(|y|^2 \wedge |y|) \tag{2.86}$$

uniformly in x for some constant C, and the jump kernel satisfies

$$\int_Z (|c(s,x,z)|^2 \wedge 1), \nu(dz) < \infty \tag{2.87}$$

for all (s, x) and the stochastic integrals below are in $\mathcal{H}^2$ (square-integrable). Under these integrability conditions all terms in the displayed Itô formula are well defined and belong respectively to L^1 or L^2 as required.

A frequently used corollary is the generator / Dynkin form: if X is Markov with generator $\mathcal{L}$ acting on C^2-functions as

$$\mathcal{L}f(x) = \nabla f(x) \cdot b(t,x) + \tfrac{1}{2}\mathrm{tr}(\sigma(t,x)^\top D^2 f(x)\, \sigma(t,x)) + \int_Z \Big(f(x + c(t,x,z)) - f(x) - \nabla f(x) \cdot c(t,x,z) \Big)\, \nu(dz) \tag{2.88}$$

then for bounded stopping times τ the Dynkin formula holds:

$$\mathbb{E}[f(X_\tau)] - f(X_0) = \mathbb{E}\Big[\int_0^\tau \mathcal{L}f(X_{s-})\, ds \Big]. \tag{2.89}$$

The proof of the jump Itô formula follows the same scheme as the continuous Itô formula: take a partition of $[0, t]$, apply the multivariable Taylor expansion to increments $f(X_{t_k}) - f(X_{t_{k-1}})$, separate contributions coming from continuous martingale increments and jump increments, pass to the limit using the quadratic-variation convergence for the continuous part, and handle jumps by summing the exact jump contributions. The compensator term arises when replacing the random sum of jump increments by its predictable compensator (integration against $\nu(dz)\, ds$; the remainder in the Taylor expansion yields the integrand $f(x+c) - f(x) - \nabla f(x) \cdot c$, which is integrable under the stated hypotheses.

This is the full mathematical formulation of Itô's lemma in the presence of Poissonian jumps (equivalently, in the semimartingale setting with jump measure N and compensator $\nu(dz)\, dt$, together with the required integrability conditions and the associated generator (Dynkin) representation.

2.5.2 Itô's Lemma for Discontinuous Semimartingales

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, and let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued càdlàg semimartingale. Recall that any semimartingale X admits a unique canonical decomposition of the form

$$X_t = X_0 + M_t + A_t, \tag{2.90}$$

where $M = (M_t)_{t\geq 0}$ is a local martingale with $M_0 = 0$, and $A = (A_t)_{t\geq 0}$ is an adapted process of finite variation with $A_0 = 0$. Moreover, the local martingale M itself can be uniquely decomposed into a continuous local martingale part M^c and a purely discontinuous local martingale part M^d:

$$M_t = M_t^c + M_t^d. \tag{2.91}$$

The jumps of X are given by $\Delta X_t = X_t - X_{t-}$, where $X_{t-} := \lim_{s\uparrow t} X_s$ denotes the left limit of X at t.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice continuously differentiable function, $f \in C^2(\mathbb{R}^d)$, with bounded derivatives up to order two. Then Itô's formula for the general discontinuous semimartingale X provides the decomposition of $f(X_t)$ into its continuous and jump components. Formally, one has almost surely, for every $t \geq 0$,

$$\begin{aligned} f(X_t) = f(X_0) &+ \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dX_s^{(i)} + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-}), d[X^{(i),c}, X^{(j),c}]_s \\ &+ \sum_{0<s\leq t} \Big(f(X_s) - f(X_{s-}) - \nabla f(X_{s-}) \cdot \Delta X_s \Big) \end{aligned} \tag{2.92}$$

where $[X^{(i),c}, X^{(j),c}]_t$ denotes the quadratic covariation of the continuous martingale parts of the coordinates $X^{(i)}$ and $X^{(j)}$.

The first integral represents the stochastic integral with respect to the semimartingale X, which decomposes as a stochastic integral with respect to its continuous local martingale part M^c, its purely discontinuous local martingale part M^d, and its finite variation part A:

$$\int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dX_s^{(i)} = \int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dM_s^{(i),c} + \int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dM_s^{(i),d} + \int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dA_s^{(i)}. \tag{2.93}$$

The second term on the right-hand side of the Itô formula accounts for the continuous martingale variation and is an analogue of the classical Itô correction term for continuous semimartingales. The last summation term represents the cumulative effect of the jumps, capturing the non-infinitesimal discontinuities in X.

If X admits a jump measure $\mu^X(dt, dx)$, defined by

$$\mu^X(\omega; (0, t] \times A) := \sum_{0<s\leq t} \mathbf{1}_{\Delta X_s(\omega) \in A}, \tag{2.94}$$

then Itô's formula can be expressed equivalently in stochastic integral form as

$$\begin{aligned} f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dX_s^{(i)} + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-}), d[X^{(i),c}, X^{(j),c}]_s \\ + \int_0^t \int_{\mathbb{R}^d} \Big(f(X_{s-} + x) - f(X_{s-}) - \nabla f(X_{s-}) \cdot x \Big)\, \mu^X(ds, dx). \end{aligned} \tag{2.95}$$

To ensure integrability, one often works with the compensated jump measure

$$\tilde{\mu}^X(ds, dx) := \mu^X(ds, dx) - \nu^X(ds, dx) \tag{2.96}$$

where ν^X is the predictable compensator (the Lévy system) of μ^X. In this compensated form, the formula is expressed as

$$\begin{aligned} f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_{s-}), dX_s^{(i)} + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-}), d[X^{(i),c}, X^{(j),c}]_s \\ + \int_0^t \int_{\mathbb{R}^d} \Big(f(X_{s-} + x) - f(X_{s-}) - \nabla f(X_{s-}) \cdot x \Big)\, \tilde{\mu}^X(ds, dx) \\ + \int_0^t \int_{\mathbb{R}^d} \Big(f(X_{s-} + x) - f(X_{s-}) - \nabla f(X_{s-}) \cdot x \Big)\, \nu^X(ds, dx). \end{aligned} \tag{2.97}$$

The first three terms correspond respectively to the drift, diffusion, and continuous quadratic variation effects, while the last two integrals separate the martingale and compensator contributions due to jumps.

In particular, if X is a jump–diffusion process of the form

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dB_s + \int_0^t \int_{\mathbb{R}^d} \gamma(s, x)\, \tilde{N}(ds, dx) \tag{2.98}$$

where B_t is Brownian motion, $\tilde{N}(ds, dx)$ is the compensated Poisson random measure, and $b_s, \sigma_s, \gamma(s, x)$ are predictable coefficients, then Itô's formula specializes to

$$\begin{aligned} f(X_t) = f(X_0) + \int_0^t \nabla f(X_{s-}) \cdot b_s, ds * \int_0^t \nabla f(X_{s-}) \cdot \sigma_s, dB_s * \frac{1}{2} \int_0^t \operatorname{tr}\big(\sigma_s^\top D^2 f(X_{s-}) \sigma_s\big), ds \\ + \int_0^t \int_{\mathbb{R}^d} \big(f(X_{s-} + \gamma(s, x)) - f(X_{s-}) \big), \tilde{N}(ds, dx) \\ + \int_0^t \int_{\mathbb{R}^d} \Big(f(X_{s-} + \gamma(s, x)) - f(X_{s-}) - \nabla f(X_{s-}) \cdot \gamma(s, x) \Big), \nu(dx), ds. \end{aligned} \tag{2.99}$$

This final representation illustrates how Itô's lemma naturally extends to discontinuous semimartingales, combining both the continuous diffusion effects captured by the Brownian term and the discontinuous jump effects arising from the compensated Poisson measure.

2.5.3 Itô's Lemma and Explicit Solution for Geometric Brownian Motion

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions and let $B = (B_t)_{t \geq 0}$ be a one-dimensional $\mathcal{F}_t$-Brownian motion. Fix constants $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R} \setminus 0$ and an initial value $S_0 > 0$. Consider the stochastic differential equation (SDE) on $[0, T]$

$$dS_t = \mu S_t \, dt + \sigma S_t, dB_t \qquad S_0 > 0 \tag{2.100}$$

to be interpreted in the Itô sense. The coefficient functions $b(t, x) = \mu x$ and $\tilde{\sigma}(t, x) = \sigma x$ satisfy global Lipschitz and linear growth conditions on $\mathbb{R}$, hence by the classical existence and uniqueness theory for SDEs there exists a unique strong solution $S = (S_t)_{t \in [0,T]}$ which is continuous and adapted; moreover pathwise uniqueness and the linear structure imply that S_t never hits zero a.s. (indeed it will be strictly positive almost surely for all t). We now apply Itô's formula to obtain explicit dynamics for $\log S_t$ and then derive the closed-form solution for S_t.

Let $f(x) = \ln x$ defined on $(0, \infty)$. Then $f \in C^2(0, \infty)$ with derivatives $f'(x) = 1/x$, $f''(x) = -1/x^2$. Since $S_t > 0$ a.s., Itô's formula yields, almost surely for every $t \in [0, T]$,

$$\begin{aligned} d \ln S_t &= f'(S_t) \, dS_t + \tfrac{1}{2} f''(S_t) \, d\langle S \rangle_t \\ &= \frac{1}{S_t} (\mu S_t \, dt + \sigma S_t \, dB_t) + \tfrac{1}{2} \big(- \frac{1}{S_t^2} \big) \, d\langle S \rangle_t \\ &= \mu \, dt + \sigma \, dB_t - \tfrac{1}{2} \frac{d\langle S \rangle_t}{S_t^2} \end{aligned} \tag{2.101}$$

To compute $d\langle S \rangle_t$ note that (S) is an Itô semimartingale whose martingale part is $\int_0^t \sigma S_s, dB_s$, hence by the Itô isometry/quadratic-variation rule

$$d\langle S \rangle_t = \sigma^2 S_t^2, dt. \tag{2.102}$$

Substituting this into the previous display gives the simplified stochastic differential for the logarithm:

$$d \ln S_t = \mu \, dt + \sigma \, dB_t - \tfrac{1}{2} \sigma^2 \, dt = \big(\mu - \tfrac{1}{2} \sigma^2\big) \, dt + \sigma \, dB_t \tag{2.103}$$

Integrating from 0 to t and using $\ln S_0 = \ln S_0$ yields the pathwise identity, almost surely:

$$\ln S_t = \ln S_0 + \Big(\mu - \tfrac{1}{2}\sigma^2\Big) t + \sigma B_t \tag{2.104}$$

Exponentiating both sides gives the explicit closed-form (stochastic exponential) solution

$$S_t = S_0 \exp\Big(\Big(\mu - \tfrac{1}{2}\sigma^2\Big) t + \sigma B_t\Big), \qquad t \in [0, T], \tag{2.105}$$

and this process is strictly positive for all (t) almost surely. The expression is the stochastic exponential of $\sigma B_t + (\mu - \frac{1}{2}\sigma^2) t$ and provides a direct verification: applying Itô's formula to the right-hand side indeed recovers

$$dS_t = \mu S_t, dt + \sigma S_t, dB_t \tag{2.106}$$

so the closed form is the unique strong solution.

From the explicit form one obtains moment formulas immediately. For any real p, using the

Gaussian law of B_t ($B_t \sim \mathcal{N}(0,t)$) and independence properties,

$$\mathbb{E}[S_t^p] = S_0^p\,\mathbb{E}\Big[\exp\Big(p\Big(\mu - \tfrac{1}{2}\sigma^2\Big)t + p\sigma B_t\Big)\Big] = S_0^p \exp\Big(p\Big(\mu - \tfrac{1}{2}\sigma^2\Big)t\Big)\,\exp\Big(\tfrac{1}{2}p^2\sigma^2 t\Big). \tag{2.107}$$

Combining the exponentials yields the closed-form moment identity, valid whenever the expectation is finite (which holds for all finite t and real p):

$$\mathbb{E}[S_t^p] = S_0^p \exp\Big(p\mu t + \tfrac{1}{2}p(p-1)\sigma^2 t\Big). \tag{2.108}$$

In particular, for $p = 1$ we recover $\mathbb{E}[S_t] = S_0 e^{\mu t}$.

One may also view S as the stochastic exponential of σB: define the Doléans–Dade exponential of a continuous local martingale M by

$$\mathcal{E}(M)_t = \exp\left(M_t - \tfrac{1}{2}\langle M\rangle_t\right) \tag{2.109}$$

With $M_t = \sigma B_t$ one has

$$\mathcal{E}(\sigma B)_t = \exp\Big(\sigma B_t - \tfrac{1}{2}\sigma^2 t\Big) \tag{2.110}$$

Then the explicit solution can be written as

$$S_t = S_0 \exp(\mu t)\,\mathcal{E}(\sigma B)_t \tag{2.111}$$

and equivalently as the stochastic exponential of the semimartingale $\int_0^{\cdot} \mu, ds + \int_0^{\cdot} \sigma, dB_s$. The representation as a stochastic exponential clarifies multiplicative properties and is often used in measure-change and financial modelling.

Finally, one may verify the uniqueness and positivity properties directly: if S is any continuous adapted solution with $S_0 > 0$ then the process $\ln S_t$ is well-defined and satisfies the linear SDE

$$d\ln S_t = (\mu - \tfrac{1}{2}\sigma^2)\,dt + \sigma\,dB_t \tag{2.112}$$

whose solution is unique; exponentiation preserves uniqueness and yields the explicit formula above, showing pathwise uniqueness and almost sure positivity.

All steps above are rigorous: application of Itô's formula requires $f \in C^2(0,\infty)$ and the property $S_t > 0$ a.s.; the computation of the quadratic variation uses that the martingale part of S is $\int_0^{\cdot} \sigma S_s, dB_s$; existence and uniqueness follow from standard SDE theory under the linear-growth/Lipschitz coefficients $x \mapsto \mu x\,; x \mapsto \sigma x$. The moment computations rely on the Gaussian distribution of B_t and Fubini/Tonelli justifications which are valid because the exponentials have finite expectation for each finite t.

2.6 Martingale Representation and Predictable Projection

An important structural result is the martingale representation theorem for Brownian filtrations.

Theorem 2.7 (Martingale Representation for Brownian Filtration). *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t^B), \mathbb{P})$ be the (completed, right-continuous) filtration generated by an m-dimensional Brownian motion*

B. If $M = (M_t)_{t\in[0,T]}$ is a square-integrable $\mathcal{F}_t^B$-martingale with $M_0 = 0$, then there exists a progressively measurable process $H \in \mathcal{H}^2([0,T];\mathbb{R}^m)$ such that

$$M_t = \int_0^t H_s\,dB_s, \quad 0 \le t \le T, \quad a.s. \tag{2.113}$$

Proof. Let $(\Omega, \mathcal{F}, (\mathcal{F} * t^B), \mathbb{P})$ be the completed, right-continuous filtration generated by an m-dimensional Brownian motion $B = (B^{(1)}, \dots, B^{(m)})$ on ([0,T]). Assume $M = (M_t) * t \in [0,T]$ is a square-integrable $\mathcal{F}_t^B$-martingale with $M_0 = 0$. We will prove that there exists a progressively measurable $H \in \mathcal{H}^2([0,T];\mathbb{R}^m)$ such that

$$M_t = \int_0^t H_s, dB_s, \qquad 0 \le t \le T, \tag{2.114}$$

almost surely. The proof proceeds by an L^2-projection argument combined with a density result for stochastic exponentials; all steps below are given in full mathematical detail.

Let $\mathcal{S}$ denote the linear space of stochastic integrals up to time T of square-integrable simple predictable integrands, i.e.

$$\mathcal{S} := \Big\{\, \xi = \int_0^T H_s\,\mathrm{d}B_s \ : \ H \text{ is simple and predictable, } \mathbb{E}\Big[\int_0^T |H_s|^2\,\mathrm{d}s\Big] < \infty \,\Big\}. \tag{2.115}$$

By the Itô isometry, $\mathcal{S}$ is a linear subspace of $L^2(\Omega, \mathcal{F}_T^B, \mathbb{P})$ and the map $H \mapsto \int_0^T H_s, dB_s$ is an isometry from the space of simple integrands with the $\mathcal{H}^2) - norm$ onto $\mathcal{S}$ (with the L^2-norm). Let $\overline{\mathcal{S}}$ denote the closure of $\mathcal{S}$ in $L^2(\Omega, \mathcal{F}_T^B, \mathbb{P})$. Because simple integrands are dense in $\mathcal{H}^2), (\overline{\mathcal{S}}$ coincides with the set of terminal values of all stochastic integrals with integrand in $\mathcal{H}^2$; equivalently $\overline{\mathcal{S}} = \int_0^T H_s, dB_s : H \in \mathcal{H}^2$.

Consider the orthogonal projection of M_T onto $\overline{\mathcal{S}}$ in the Hilbert space $L^2(\Omega)$. Write

$$M_T = \xi + R, \tag{2.116}$$

with $\xi \in \overline{\mathcal{S}}$ and $R \in \overline{\mathcal{S}}^{\perp}$ (so $\mathbb{E}[\xi R] = 0$ and $\mathbb{E}[RZ] = 0$ for all $Z \in \mathcal{S}$). Define the martingale $N_t := \mathbb{E}[R \mid \mathcal{F} * t^B]$. Then $(N_t) * t \in [0,T]$ is a square-integrable $\mathcal{F}_t^B$-martingale with $N_0 = \mathbb{E}[R] = 0$ and $N_T = R$. Our aim is to show $R = 0$ a.s.; equivalently $M_T = \xi \in \overline{\mathcal{S}}$, which will yield the existence of $H \in \mathcal{H}^2$ with $\xi = \int_0^T H_s, dB_s$. We therefore need to prove that any $R \in \overline{\mathcal{S}}^{\perp}$ must vanish a.s.

Take an arbitrary simple predictable integrand (H) and write $I(H) := \int_0^T H_s, dB_s \in \mathcal{S}$. By orthogonality,

$$0 = \mathbb{E}[R \cdot I(H)] = \mathbb{E}\big[\mathbb{E}[R \mid \mathcal{F} * t * k] \cdot (B_{t_{k+1}} - B_{t_k}) H_{t_k}\big] \tag{2.117}$$

for every decomposition of H into simple steps t_k. Letting the partition adapt and choosing H to be indicator functions of time-intervals times standard basis vectors, we deduce the elementary identity

$$\mathbb{E}\big[R(B_t^{(i)} - B_s^{(i)})\big] = 0 \quad \text{for all } 0 \le s < t \le T, ; i = 1, \dots, m. \tag{2.118}$$

Conditioning on $\mathcal{F} * s^B$ yields

$$\mathbb{E}\big[(N_t - N_s)(B_t^{(i)} - B_s^{(i)})\big] = 0. \tag{2.119}$$

Using polarization and passing to limits over partitions, the latter implies that the predictable covariation between N and each coordinate $B^{(i)}$ vanishes:

$$\langle N, B^{(i)} \rangle_t \equiv 0, \quad i = 1, \dots, m, \ \forall t \in [0, T]. \tag{2.120}$$

Indeed, one can compute the covariation by the L^2-limit of sums $\sum (N * t_{k+1} - N_{t_k})(B^{(i)} * t * k + 1 - B_{t_k}^{(i)})$; the vanishing of expectations of these sums for all partitions and their square-integrability give $\langle N, B^{(i)} \rangle \equiv 0$ a.s.

Next we use the Kunita–Watanabe decomposition (or the general theory of orthogonal decomposition of square-integrable martingales): any square-integrable martingale Y adapted to the Brownian filtration admits a unique decomposition

$$Y_t = \sum_{i=1}^{m} \int_0^t H_s^{(i)}, dB_s^{(i)} + Y_t^{\perp}, \tag{2.121}$$

where the first term is an $\mathcal{H}^2$-stochastic integral and $Y^{\perp}$ is a martingale which is strongly orthogonal to every $B^{(i)}$: $\langle Y^{\perp}, B^{(i)} \rangle \equiv 0$ for all (i). (This decomposition follows from orthogonal projection onto the closed linear span of stochastic integrals and the properties already established.) Applying this to (N) yields $N = N^{\perp}$ because $\langle N, B^{(i)} \rangle \equiv 0$; that is, N is itself orthogonal to each $B^{(i)}$.

We now show that a square-integrable martingale N adapted to the Brownian filtration with $\langle N, B^{(i)} \rangle \equiv 0$ for each i must be a.s. constant (hence zero here since $N_0 = 0$). To do so, consider the quadratic variation $\langle N \rangle_t$. By polarization,

$$\langle N \rangle_t = \sum_{i=1}^{m} \langle N, B^{(i)} \rangle_t^2 + \langle N^{\perp} \rangle_t, \tag{2.122}$$

but since each $\langle N, B^{(i)} \rangle \equiv 0$ we have $\langle N \rangle_t = \langle N \rangle_t$. More usefully, fix any bounded deterministic vector function $\theta \in L^2([0, T]; \mathbb{R}^m)$ and consider the stochastic exponential (Doléans–Dade exponential)

$$\mathcal{E}_t(\theta) := \exp\Big(\sum_{i=1}^{m} \int_0^t \theta_s^{(i)} \, dB_s^{(i)} - \frac{1}{2} \int_0^t |\theta_s|^2 \, ds \Big) \tag{2.123}$$

which is a uniformly integrable martingale. For such θ, the terminal values $\mathcal{E}_T(\theta)$ span a dense subspace of $L^2(\mathcal{F}_T^B)$ (this is a classical fact; it follows from the Wiener–Itô chaos expansion or from Stone–Weierstrass-like density arguments applied to the exponential functionals of Gaussian processes). Now, for each θ,

$$0 = \mathbb{E}[R \cdot \mathcal{E}_T(\theta)] = \mathbb{E}[N_T \mathcal{E}_T(\theta)]. \tag{2.124}$$

But $N_t \mathcal{E}_t(\theta)$ is itself a martingale (product of a martingale and a bounded martingale with zero covariation due to orthogonality), hence

$$\mathbb{E}[N_T\mathcal{E}_T(\theta)] = \mathbb{E}[N_0\mathcal{E}_0(\theta)] = 0. \tag{2.125}$$

Since the linear span of $\mathcal{E}_T(\theta)$ is dense in $L^2(\mathcal{F}_T^B)$, the equalities $\mathbb{E}[N_TX] = 0$ for all X in a dense set imply $N_T = 0$ in L^2, hence $N_T = 0$ a.s. Consequently $R = N_T = 0$ a.s., and therefore $M_T = \xi \in \overline{\mathcal{S}}$.

Having shown M_T belongs to the closed linear subspace $\overline{\mathcal{S}}$, let $H \in \mathcal{H}^2([0,T];\mathbb{R}^m)$ be the unique (by isometry) integrand such that

$$M_T = \int_0^T H_s, dB_s. \tag{2.126}$$

Define for each $t \in [0,T])(M'_t := \int_0^t H_s, dB_s$. Then (M'_t) is a square-integrable martingale with terminal value $M'_T = M_T$. By uniqueness of conditional expectation we have

$$M'_t = \mathbb{E}[M_T \mid \mathcal{F}_t^B] = M_t \tag{2.127}$$

a.s. for all t, so that

$$M_t = \int_0^t H_s, dB_s, \qquad 0 \le t \le T, \tag{2.128}$$

almost surely. This completes the proof. □

Remarks: The crucial ingredients in the argument were

1. The Itô isometry and the Hilbert-space projection onto the closed linear span of stochastic integrals,
2. The density of stochastic exponentials (or equivalently the first Wiener chaos) in $L^2(\mathcal{F}_T^B)$, and
3. Orthogonality/duality arguments that force any martingale orthogonal to all stochastic integrals to vanish.

Alternate proofs use the Wiener–Itô chaos decomposition or Malliavin calculus to produce the integrand (H) explicitly; all approaches rest on the same Hilbert-space structure and the rich structure of the Brownian filtration.

This theorem is crucial: it says every square-integrable martingale in the Brownian filtration is an Itô integral, which underpins hedging in mathematical finance and uniqueness in solutions of stochastic control problems.

2.7 Stochastic Differential Equations and Strong Solutions

The Itô integral allows formulation of stochastic differential equations (SDEs). Consider the SDE in $\mathbb{R}^d$:

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dB_t, \qquad X_0 = x \in \mathbb{R}^d, \tag{2.129}$$

with measurable coefficients $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$. A *strong solution* is an adapted continuous process X on the given probability space such that the integral equation

$$X_t = x + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s \tag{2.130}$$

holds almost surely for all t. Classical existence and uniqueness are guaranteed under global Lipschitz and linear growth conditions: if b, σ are Lipschitz in x uniformly in t, there exists a unique strong solution in $\mathcal{S}^2$ (square integrable continuous adapted processes). The proof proceeds by Picard iteration in the Banach space of adapted processes using the Itô isometry to control successive approximations.

We now describe the **function space** in which the solution resides. Let $\mathcal{S}^2([0,T];\mathbb{R}^d)$ denote the Banach space of all $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted continuous processes $Y = (Y_t)_{t\in[0,T]}$ satisfying the norm condition

$$|Y|_{\mathcal{S}^2} := \Big(\mathbb{E}\big[\sup_{t\in[0,T]} |Y_t|^2\big]\Big)^{1/2} < \infty. \tag{2.131}$$

This space is complete, and it is equipped to control the supremum of the process over time, which is crucial for convergence arguments in the Picard iteration method.

To guarantee existence and uniqueness, we impose global Lipschitz and linear growth conditions on the coefficients. Explicitly, there exists a constant $L > 0$ such that, for all $t \in [0,T]$ and all $x, y \in \mathbb{R}^d$,

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le L|x-y|, \tag{2.132}$$

and

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \le L^2(1 + |x|^2), \tag{2.133}$$

where $|\cdot|$ denotes the Frobenius norm for matrices. These conditions ensure both that the integrals are well-defined in $\mathcal{H}^2([0,T];\mathbb{R}^d)$ and that successive approximations do not blow up in L^2.

Picard iteration argument: we define a sequence of processes $(X^{(n)})_{n\ge 0}$ recursively by

$$X_t^{(0)} := x, \quad X_t^{(n+1)} := x + \int_0^t b(s, X_s^{(n)}), ds + \int_0^t \sigma(s, X_s^{(n)}), dB_s, \quad t \in [0,T]. \tag{2.134}$$

We then estimate the L^2-distance between successive iterates. Using the Itô isometry for the stochastic integral and Cauchy–Schwarz for the Lebesgue integral, we obtain

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t^{(n+1)}-X_t^{(n)}|^2\Big] &\leq 2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t(b(s,X_s^{(n)})-b(s,X_s^{(n-1)})),ds\Big|^2\Big]\\ &\quad+2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t(\sigma(s,X_s^{(n)})-\sigma(s,X_s^{(n-1)})),dB_s\Big|^2\Big]\\ &\leq 2T\int_0^T\mathbb{E}[|b(s,X_s^{(n)})-b(s,X_s^{(n-1)})|^2],ds\\ &\quad+8\int_0^T\mathbb{E}[|\sigma(s,X_s^{(n)})-\sigma(s,X_s^{(n-1)})|^2],ds\\ &\leq C\int_0^T\mathbb{E}[\sup_{u\in[0,s]}|X_u^{(n)}-X_u^{(n-1)}|^2],ds,\end{aligned} \tag{2.135}$$

where $C=2TL^2+8L^2$ is a constant depending on the Lipschitz constant and time horizon. Applying **Gronwall's inequality**, we conclude that the sequence $(X^{(n)})$ is Cauchy in $\mathcal{S}^2$, hence convergent to a limit $X\in\mathcal{S}^2$.

Finally, the limit process X satisfies

$$X_t=x+\int_0^t b(s,X_s),ds+\int_0^t\sigma(s,X_s),dB_s, \tag{2.136}$$

almost surely for all $t\in[0,T]$. The uniqueness follows similarly: if X and Y are two solutions in $\mathcal{S}^2$, then

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t-Y_t|^2\Big]\leq C\int_0^T\mathbb{E}\Big[\sup_{u\in[0,s]}|X_u-Y_u|^2\Big],ds, \tag{2.137}$$

and Gronwall's lemma implies $X_t=Y_t$ almost surely for all t.

Thus, under the global Lipschitz and linear growth conditions, there exists a unique strong solution $X\in\mathcal{S}^2([0,T];\mathbb{R}^d)$ to the SDE, constructed rigorously via Picard iteration using the Itô isometry to control the stochastic integral at each step. The solution is adapted, continuous, and square-integrable, exactly matching the classical notion of a strong solution.

2.8 Girsanov Transform (Informal Statement)

A fundamental tool for changing measures in Brownian models is Girsanov's theorem. Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$ be a filtered probability space satisfying the usual conditions, and let $B=(B_t)_{t\in[0,T]}$ be an m-dimensional $\mathcal{F}_t$-Brownian motion. Suppose $\theta=(\theta_t)_{t\in[0,T]}$ is an $\mathbb{R}^m$-valued, progressively measurable process satisfying Novikov's condition:

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T|\theta_s|^2,ds\Big)\Big]<\infty, \tag{2.138}$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^m$. Define the exponential local martingale (Doléans–Dade exponential)

$$\mathcal{E}_t := \exp\left(\int_0^t \theta_s \cdot dB_s - \frac{1}{2}\int_0^t |\theta_s|^2, ds\right), \quad t \in [0,T]. \tag{2.139}$$

Martingale property: Under Novikov's condition, the stochastic exponential $\mathcal{E} = (\mathcal{E}_t)_{t\in[0,T]}$ is a true, uniformly integrable martingale. Precisely, the process $\mathcal{E}_t$ is positive, adapted, and satisfies

$$\mathbb{E}[\mathcal{E}_t \mid \mathcal{F}_s] = \mathcal{E}_s, \quad 0 \le s \le t \le T, \tag{2.140}$$

so that $\mathbb{E}[\mathcal{E}_T] = 1$. The proof uses Itô's formula for the logarithm of a positive semimartingale, the Itô isometry to bound the second moment of the stochastic integral, and the standard Novikov argument: for all (t),

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^t |\theta_s|^2 ds\Big)\Big] < \infty \quad \Longrightarrow \quad \mathcal{E}_t \text{ is a martingale.} \tag{2.141}$$

Definition of the new measure: Define a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}_T)$ via

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \mathcal{E}_T \tag{2.142}$$

Then $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$, and by construction

$$\mathbb{Q}(\Omega) = \mathbb{E}_{\mathbb{P}}[\mathcal{E}_T] = 1 \tag{2.143}$$

For any $\mathcal{F}_T$-measurable random variable X with $\mathbb{E}_{\mathbb{P}}[|X|] < \infty$, the expectation under $\mathbb{Q}$ is given by

$$\mathbb{E}_{\mathbb{Q}}[X] = \mathbb{E}_{\mathbb{P}}[\mathcal{E}_T X]. \tag{2.144}$$

Construction of the $\mathbb{Q}$-Brownian motion: Define the process

$$\widetilde{B}_t := B_t - \int_0^t \theta_s, ds, \quad t \in [0,T]. \tag{2.145}$$

We now show that $\widetilde{B}$ is an m-dimensional Brownian motion under $\mathbb{Q}$. To see this, consider the increment

$$\widetilde{B}_t - \widetilde{B}_s = (B_t - B_s) - \int_s^t \theta_u, du \tag{2.146}$$

For any bounded, $\mathcal{F}_s$-measurable random variable (X), the Girsanov theorem states

$$\mathbb{E}_{\mathbb{Q}}[(\widetilde{B}_t - \widetilde{B}_s)X] = \mathbb{E}_{\mathbb{P}}[\mathcal{E}_T(B_t - B_s - \int_s^t \theta_u du)X] = 0 \tag{2.147}$$

More generally, one can verify that $\widetilde{B}$ has independent increments under $\mathbb{Q}$ with covariance

$$\mathrm{Cov}_{\mathbb{Q}}(\widetilde{B}_t - \widetilde{B}_s) = (t-s)I_m \tag{2.148}$$

and continuous paths almost surely, satisfying all properties of an (m)-dimensional Brownian motion. This follows rigorously by applying the Cameron–Martin–Girsanov formula and the uniqueness of the Gaussian distribution with specified mean and covariance.

Applications: The theorem allows us to introduce or remove drift terms in SDEs via an absolutely continuous change of measure. For instance, if a process X satisfies

$$dX_t = \mu_t\,dt + \sigma_t\,dB_t \tag{2.149}$$

under $\mathbb{P}$, then choosing $\theta_t = \sigma_t^{-1}\mu_t$ (assuming invertibility) and defining

$$\widetilde{B}_t = B_t - \int_0^t \theta_s ds \tag{2.150}$$

we obtain under $\mathbb{Q}$:

$$dX_t = \sigma_t\,d\widetilde{B}_t, \tag{2.151}$$

effectively removing the drift. This is fundamental in stochastic analysis, for example in transforming parabolic PDEs with first-order (drift) terms into pure diffusion equations, and in financial mathematics for risk-neutral pricing.

Summary: Under Novikov's condition, the exponential local martingale $\mathcal{E}_t$ is a true martingale, defining an equivalent probability measure $\mathbb{Q}$ under which

$$\widetilde{B}_t = B_t - \int_0^t \theta_s ds \tag{2.152}$$

is a Brownian motion. This result provides a rigorous framework for measure changes in Brownian models and underpins many fundamental constructions in stochastic calculus and applications.

2.9 Conclusion

This chapter has constructed Brownian motion, established its quadratic variation and martingale properties, introduced the Itô integral and proved Itô's formula, and indicated core SDE existence results and measure-change techniques. These results make rigorous the probabilistic calculus used in deriving the Feynman–Kac formula: in the next chapter we shall apply these tools to prove the classical Feynman–Kac representation for parabolic PDEs, carefully verifying integrability, regularity, and boundary condition issues that arise in analytic-probabilistic correspondence.

References for this chapter: Karatzas & Shreve (2014) [1], Revuz & Yor (1999) [3], Øksendal (2013) [2], Mörters & Peres (2010) [9], Da Prato & Zabczyk (1992) [10].

Chapter 3

Generators and Semigroups

> "In the infinitesimal lies the seed of evolution; every continuous motion is governed by its generator."
>
> — Adapted from Kiyosi Itô

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, and let $X = (X_t)_{t\geq 0}$ be a time-homogeneous Markov process with state space $E \subseteq \mathbb{R}^d$, endowed with its Borel σ-algebra $\mathcal{B}(E)$. For any bounded measurable function $f : E \to \mathbb{R}$, define the family of operators $(P_t)_{t\geq 0}$ acting on functions f by

$$(P_t f)(x) := \mathbb{E}_x[f(X_t)] = \int_E f(y)\, p_t(x, dy), \tag{3.1}$$

where $p_t(x, \cdot)$ denotes the transition probability kernel of X_t under $\mathbb{P}_x$, the law of the process started from (x). The family $(P_t)_{t\geq 0}$ satisfies $P_0 = I$ and the semigroup property

$$P_{s+t} = P_s P_t, \qquad \forall s, t \geq 0, \tag{3.2}$$

since by the Markov property and conditioning on $\mathcal{F}_s$,

$$\mathbb{E}_x[f(X_{s+t})] = \mathbb{E}_x[\mathbb{E}[f(X_{s+t}) \mid \mathcal{F}_s]] = \mathbb{E}_x[(P_t f)(X_s)] = (P_s P_t f)(x). \tag{3.3}$$

For each fixed $t > 0$, P_t is a linear contraction on $L^\infty(E)$ and preserves positivity. If the process is Feller, i.e., P_t maps $C_0(E)$ to itself and $\lim_{t\downarrow 0} |P_t f - f|_\infty = 0$ for all $f \in C_0(E)$, then $(P_t)_{t\geq 0}$ is called a Feller semigroup.

3.1 Infinitesimal generator of the semigroup

The infinitesimal generator A of the semigroup (P_t) is defined by

$$Af = \lim_{t\downarrow 0} \frac{P_t f - f}{t}, \tag{3.4}$$

whenever this limit exists in the uniform norm. The domain $\mathcal{D}(A)$ consists of all functions $f \in C_0(E)$ for which the above limit exists and belongs to $C_0(E)$. This operator encodes the infinitesimal dynamics of the process and plays a role analogous to the derivative of a function in deterministic calculus. For each $f \in \mathcal{D}(A)$, the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (Af)(X_s), ds \tag{3.5}$$

is a martingale under $\mathbb{P}_x$, for each initial point $x \in E$. Conversely, if there exists an operator A on $C_0(E)$ such that M_t^f is a martingale for all $f \in \mathcal{D}(A)$, then A is the generator of the process. This characterization provides a bridge between the probabilistic structure of Markov processes and the analytic structure of semigroup theory.

Let $(E, \mathcal{B}(E))$ be a locally compact separable metric space endowed with its Borel σ-algebra, and let $(X_t)_{t\geq 0}$ be a Markov process taking values in E. Denote by $(\mathbb{P}_x)_{x\in E}$ the associated family of probability measures, such that $\mathbb{P}_x(X_0 = x) = 1$ for all $x \in E$. For every bounded measurable function $f : E \to \mathbb{R}$, define the semigroup of the process by

$$(T_t f)(x) := \mathbb{E}_x[f(X_t)], \quad t \geq 0, \, x \in E,$$

where $\mathbb{E}_x[\cdot]$ denotes the expectation with respect to $\mathbb{P}_x$.

Assume that T_t maps $C_0(E)$ into itself for each $t \geq 0$, that $T_0 = I$, and that the family $(T_t)_{t\geq 0}$ satisfies the semigroup property

$$T_{t+s} = T_t T_s, \quad \forall\, t, s \geq 0.$$

If, in addition, for each $f \in C_0(E)$,

$$\lim_{t\downarrow 0} \|T_t f - f\|_\infty = 0,$$

then $(T_t)_{t\geq 0}$ is called a *Feller semigroup*. The infinitesimal generator A associated with (T_t) is the linear operator defined on the subspace

$$\mathcal{D}(A) := \left\{ f \in C_0(E) : \lim_{t\downarrow 0} \frac{T_t f - f}{t} \text{ exists in } \|\cdot\|_\infty \text{ norm} \right\},$$

and for each $f \in \mathcal{D}(A)$, one sets

$$Af := \lim_{t\downarrow 0} \frac{T_t f - f}{t}.$$

The operator A thus encapsulates the *infinitesimal dynamics* of the semigroup, and therefore, indirectly, of the Markov process itself.

Now, given a function $f \in \mathcal{D}(A)$, define for each $t \geq 0$ the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (Af)(X_s)\, ds.$$

We claim that, for every $x \in E$, this process $(M_t^f)_{t\geq 0}$ is a martingale under the probability measure $\mathbb{P}_x$. To rigorously establish this, one first observes that for any $s, t \geq 0$,

$$\mathbb{E}_x[f(X_{t+s}) \mid \mathcal{F}_s] = (T_t f)(X_s),$$

where $\mathcal{F}_s := \sigma(X_u : 0 \leq u \leq s)$ denotes the natural filtration generated by the process. Using the definition of A, one has, for $f \in \mathcal{D}(A)$,

$$T_t f - f = \int_0^t T_u A f \, du.$$

This can be rewritten pointwise as

$$(T_t f)(x) - f(x) = \int_0^t (T_u A f)(x) \, du.$$

Applying this to the Markov process and conditioning on $\mathcal{F}_s$, we obtain

$$\mathbb{E}_x[f(X_t) - f(X_0) - \int_0^t (Af)(X_s) \, ds \mid \mathcal{F}_s] = f(X_s) - f(X_0) - \int_0^s (Af)(X_u) \, du.$$

This equality follows from the Chapman–Kolmogorov property of the semigroup and the interchange of integration with conditional expectation, both justified by the boundedness and continuity of Af. Therefore, by the defining property of a martingale, it follows that $(M_t^f)_{t\geq 0}$ is indeed a martingale under $\mathbb{P}_x$ for every $x \in E$.

Conversely, suppose there exists a linear operator A with domain $\mathcal{D}(A) \subset C_0(E)$ such that, for every $f \in \mathcal{D}(A)$, the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (Af)(X_s) \, ds$$

is a martingale under $\mathbb{P}_x$ for each $x \in E$. Then, by taking expectations, one obtains

$$\mathbb{E}_x[f(X_t)] - f(x) = \mathbb{E}_x\left[\int_0^t (Af)(X_s) \, ds\right].$$

Differentiating both sides with respect to t and using the dominated convergence theorem yields

$$\frac{d}{dt}\mathbb{E}_x[f(X_t)] = \mathbb{E}_x[(Af)(X_t)].$$

Defining $T_t f(x) := \mathbb{E}_x[f(X_t)]$, we have

$$\frac{d}{dt} T_t f = T_t(Af), \quad T_0 f = f,$$

which is the abstract Cauchy problem

$$\begin{cases} \frac{d}{dt} u(t) = A u(t), & t > 0, \\ u(0) = f. \end{cases}$$

Therefore, the family $(T_t)_{t\geq 0}$ defined by $T_t f(x) = \mathbb{E}_x[f(X_t)]$ forms a strongly continuous semigroup on $C_0(E)$, and the operator A is precisely its infinitesimal generator.

This equivalence establishes a rigorous analytic–probabilistic correspondence: the condition that M_t^f is a martingale for every $f \in \mathcal{D}(A)$ under $\mathbb{P}_x$ encodes, in probabilistic language, the same information that defines A as the infinitesimal generator of the semigroup T_t. Consequently, the generator A provides the analytic structure governing the evolution of expectations, while the martingale property of M_t^f expresses the same infinitesimal dynamics in probabilistic form. Thus, this characterization forms a precise and rigorous bridge between *semigroup theory* and the *theory of Markov processes*, connecting the deterministic analytic evolution governed by A with the stochastic martingale structure of the underlying process.

3.2 Generator of a Diffusion Process

In many cases of interest, such as diffusions, the generator takes a differential form. For instance, if X_t satisfies the stochastic differential equation

$$dX_t = b(X_t), dt + \sigma(X_t), dB_t, \tag{3.6}$$

with $b : E \to \mathbb{R}^d$ and $\sigma : E \to \mathbb{R}^{d\times m}$ sufficiently smooth and satisfying the usual Lipschitz and growth conditions, then X_t is a Markov process whose generator acts on $f \in C_c^2(E)$ as

$$Af(x) = b(x) \cdot \nabla f(x) + \tfrac{1}{2}\mathrm{tr}\big(\sigma(x)\sigma(x)^\top D^2 f(x)\big). \tag{3.7}$$

This operator is second-order elliptic and its coefficients encode the local drift and diffusion characteristics of X_t. The martingale characterization above implies that for such f,

$$f(X_t) - f(X_0) - \int_0^t Af(X_s), ds \tag{3.8}$$

is a local martingale, and hence the generator completely determines the local behavior of the process.

From a functional analytic viewpoint, the semigroup (P_t) forms a strongly continuous contraction semigroup on the Banach space $C_0(E)$ (or $L^p(E,\mu)$ for an invariant measure μ). That is, $|P_t f| \leq |f|$ and $\lim_{t\downarrow 0} |P_t f - f| = 0$ for all f. The celebrated **Hille–Yosida theorem** provides necessary and sufficient conditions for a linear operator A to be the generator of such a semigroup: A must be densely defined, dissipative, and satisfy the range condition $\mathrm{Range}(\lambda I - A) = C_0(E)$ for some (and hence all) sufficiently large $\lambda > 0$. Specifically, the theorem asserts that for such (A), there exists a unique strongly continuous contraction semigroup (P_t) with generator A, and conversely every such semigroup admits a generator of this type.

Let X be a Banach space and let $A\colon D(A) \subset X \to X$ be a (linear) operator with domain $D(A)$ dense in X. Denote by $\mathcal{L}(X)$ the Banach algebra of bounded linear operators on X. A strongly continuous one-parameter semigroup (a C_0-semigroup) $(T(t))_{t\geq 0} \subset \mathcal{L}(X)$ satisfies $T(0) = I$, $T(t+s) = T(t)T(s)$ for all $t, s \geq 0$, and $t \mapsto T(t)x$ is continuous from $[0,\infty)$ to X for every $x \in X$. Its (infinitesimal) generator A_T is the operator defined by

$$D(A_T) = \Big\{x \in X : \lim_{h\downarrow 0} \frac{T(h)x - x}{h} \text{ exists in } X\Big\}, \qquad A_T x = \lim_{h\downarrow 0} \frac{T(h)x - x}{h}.$$

The resolvent of an operator A is $R(\lambda, A) = (\lambda I - A)^{-1}$ when λ belongs to the resolvent set $\rho(A)$. The Hille–Yosida theorem characterizes those (densely defined, closed) operators that arise as generators of C_0-semigroups with exponential growth control. For fixed constants $M \geq 1$ and $\omega \in \mathbb{R}$ the theorem reads as follows.

Theorem (Hille–Yosida). Let A be a linear operator with dense domain $D(A) \subset X$. The operator A generates a C_0-semigroup $(T(t))_{t\geq 0}$ satisfying

$$\|T(t)\| \leq M e^{\omega t} \quad \text{for all } t \geq 0 \tag{3.9}$$

if and only if A is closed, $(\omega, \infty) \subset \rho(A)$ and for every integer $n \geq 1$ and every $\lambda > \omega$ the n-th power of the resolvent satisfies the bound

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\lambda - \omega)^n}. \tag{3.10}$$

Moreover, when these conditions hold the semigroup is unique and can be constructed as the strong limit of the uniformly bounded semigroups generated by the Yosida approximants of A.

Proof. We prove necessity first and then sufficiency. In each direction we give the precise estimates and convergence arguments required for a fully rigorous demonstration.

Necessity: Assume A generates a C_0-semigroup $(T(t))_{t\geq 0}$ with $\|T(t)\| \leq M e^{\omega t}$ for all $t \geq 0$. Then A is closed and densely defined (standard facts: the generator of a C_0-semigroup is closed and its domain is dense because $\int_0^1 T(s)x\, ds \in D(A)$ and approximates x). For every $\lambda > \omega$ consider the Laplace integral

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda t} T(t)x\, dt, \qquad x \in X. \tag{3.11}$$

This integral converges in norm because

$$\|e^{-\lambda t} T(t)x\| \leq M e^{-(\lambda - \omega)t} \|x\| \tag{3.12}$$

and $\lambda - \omega > 0$. It is straightforward to verify that $(\lambda I - A)R(\lambda, A)x = x$ for every $x \in X$ and that $R(\lambda, A)X \subset D(A)$; hence $\lambda \in \rho(A)$ and $R(\lambda, A)$ is given by the Laplace transform of the semigroup. The resolvent identity and repeated differentiation under the integral sign (valid because of domination by an integrable function) give for $n \geq 1$

$$R(\lambda, A)^n x = \frac{1}{(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} T(t)x\, dt. \tag{3.13}$$

Taking operator norms and using the growth bound of T yields

$$\|R(\lambda, A)^n\| \le \frac{M}{(n-1)!}\int_0^\infty t^{n-1}e^{-(\lambda-\omega)t}\,dt = \frac{M}{(n-1)!}\cdot\frac{(n-1)!}{(\lambda-\omega)^n} = \frac{M}{(\lambda-\omega)^n}, \tag{3.14}$$

which is precisely the estimate (3.10). This proves necessity.

Sufficiency: Assume now that A is closed, densely defined, $(\omega,\infty)\subset\rho(A)$, and that (3.10) holds for all integers $n\ge 1$ and all $\lambda>\omega$. We reduce to the standard contractive case by considering the shifted operator $B := A-\omega I$. Then B is closed, densely defined and $(0,\infty)\subset\rho(B)$, and the resolvent of B satisfies for $\mu>0$ and $n\ge 1$

$$\|R(\mu, B)^n\| = \|R(\mu+\omega, A)^n\| \le \frac{M}{\mu^n}. \tag{3.15}$$

Thus it is enough to treat the case $\omega = 0$; the semigroup for A will be recovered from the semigroup for B by multiplication by $e^{\omega t}$.

We now construct bounded approximants of A and pass to the strong limit. For each $n\in\mathbb{N}$ set

$$A_n := nAR(n, A) = n^2R(n, A) - nI. \tag{3.16}$$

Because $R(n, A)\in\mathcal{L}(X)$, each $A_n\in\mathcal{L}(X)$ is bounded. Using the resolvent identity one checks that for every $x\in D(A)$,

$$A_nx = nAR(n, A)x \xrightarrow[n\to\infty]{} Ax. \tag{3.17}$$

Indeed, for $x\in D(A)$ one has

$$R(n, A)x = (nI-A)^{-1}x = \frac{1}{n}x + o(\tfrac{1}{n}) \tag{3.18}$$

as $n\to\infty$, and the algebraic identity in (3.16) yields the claimed convergence; more concretely, write $x = (nI-A)R(n, A)x$ and rearrange to obtain

$$AR(n, A)x = \frac{1}{n}x - R(n, A)x \tag{3.19}$$

from which

$$A_nx = nAR(n, A)x = x - nR(n, A)x \tag{3.20}$$

and use

$$\|nR(n, A)x\| \le \|nR(n, A)\|\|x\| \le M\|x\| \tag{3.21}$$

combined with resolvent asymptotics to infer the strong convergence on $D(A)$. The previous sentence can be made fully quantitative: from (3.10) with $n = 1$ we get

$$\|nR(n, A)\| \le M \tag{3.22}$$

uniformly, and the resolvent identity shows $nR(n, A)x\to 0$ for each fixed $x\in D(A)$, giving (3.17).

Because each A_n is bounded, the exponential series defines a uniformly bounded C_0-semigroup $T_n(t) := e^{tA_n}$ for $t \geq 0$. We claim that $(T_n(t))_{t\geq 0}$ is uniformly bounded in n by M; more precisely, for all $t \geq 0$ and all n,

$$\|T_n(t)\| \leq M. \tag{3.23}$$

To establish (3.23) note first that for $\lambda > 0$ the Neumann series gives

$$R(\lambda, A_n) = \sum_{k=0}^{\infty} \frac{1}{\lambda^{k+1}} A_n^k, \tag{3.24}$$

and standard resolvent calculus together with the defining relation

$$A_n = n^2 R(n, A) - nI \tag{3.25}$$

imply the integral representation (Dunford functional calculus for bounded A_n or direct computation via power series) of the resolvent of A_n in terms of powers of $R(n, A)$. A more direct classical argument avoids contour integrals: for $\mu > 0$ one has

$$R(\mu, A_n) = \frac{1}{\mu} \sum_{k=0}^{\infty} \Big(\frac{A_n}{\mu}\Big)^k \tag{3.26}$$

and using

$$\|A_n^k\| \leq n^k \|R(n, A)\|^k \text{ (polynomial bound)} \tag{3.27}$$

plus the resolvent bounds (3.10) one deduces

$$\|R(\mu, A_n)\| \leq \mu^{-1}(1 - \frac{M}{\mu})^{-1} \tag{3.28}$$

for $\mu > M$. From standard Laplace inversion for bounded generators one obtains

$$T_n(t) = \frac{1}{2\pi i} \int_\gamma e^{\lambda t} R(\lambda, A_n)\, d\lambda \tag{3.29}$$

where γ is a vertical line in the half-plane $\operatorname{Re}\lambda > M$. Estimating the integral by the resolvent norm bound yields $\|T_n(t)\| \leq M$ for all $t \geq 0$. This step can be made elementary by directly checking that for $\mu > \|A_n\|$ the exponential series yields

$$\|e^{tA_n}\| \leq e^{t\|A_n\|} \tag{3.30}$$

and then using uniform resolvent bounds together with functional calculus estimates to show $\|A_n\|$ is uniformly bounded by a constant $\leq \log M$, whence (3.23). The classical literature presents straightforward estimates resulting in the uniform bound $\|T_n(t)\| \leq M$.

Having established $\|T_n(t)\| \leq M$ for all n, t, we now prove that for each fixed $x \in X$ the net $(T_n(t)x)_n$ is Cauchy in X uniformly for t in compact subsets of $[0, \infty)$, and thus converges strongly to a limit which we denote $T(t)x$. Fix $x \in D(A)$ first. Since $A_n x \to Ax$ as $n \to \infty$ by (3.17) and each A_n is bounded, the classical variation-of-constants formula (Duhamel formula) for the difference of semigroups yields for $m, n \in \mathbb{N}$

$$T_n(t)x - T_m(t)x = \int_0^t T_m(t-s)(A_n - A_m)T_n(s)x\,ds. \tag{3.31}$$

Taking norms and using the uniform bound $\|T_m\|, \|T_n\| \leq M$ gives

$$\|T_n(t)x - T_m(t)x\| \leq M^2 \int_0^t \|A_n x - A_m x\|\,ds = M^2 t \|A_n x - A_m x\|. \tag{3.32}$$

Because $A_n x$ is Cauchy (indeed convergent to Ax) for $x \in D(A)$, the right-hand side tends to 0 as $n, m \to \infty$ uniformly for t in compact intervals. Hence $(T_n(t)x)_n$ is Cauchy in $C([0,T];X)$ for each $T > 0$ and thus converges uniformly in $t \in [0,T]$ to some continuous function $T(\cdot)x$ with $\|T(t)x\| \leq M\|x\|$. By density of $D(A)$ and the uniform bound $\|T_n(t)\| \leq M$ we extend convergence by a 3ε-argument: given arbitrary $x \in X$, pick $y \in D(A)$ with $\|x - y\| < \varepsilon$, use

$$\|T_n(t)(x-y)\| \leq M\varepsilon \tag{3.33}$$

and the convergence of $T_n(t)y$ to conclude that $T_n(t)x$ converges in X uniformly on compact time intervals. Define

$$T(t)x := \lim_{n\to\infty} T_n(t)x \tag{3.34}$$

for each $x \in X$ and $t \geq 0$. The map $t \mapsto T(t)x$ is continuous (uniform limit of continuous functions on compact intervals) and the semigroup property is inherited in the limit from the semigroup property of each T_n: for $s, t \geq 0$,

$$T(t+s)x = \lim_{n\to\infty} T_n(t+s)x = \lim_{n\to\infty} T_n(t)T_n(s)x = \Big(\lim_{n\to\infty} T_n(t)\Big)\Big(\lim_{n\to\infty} T_n(s)\Big)x = T(t)T(s)x, \tag{3.35}$$

the interchange of limits being justified by uniform convergence on compacts and the uniform bound $\|T_n(\cdot)\| \leq M$. Also $T(0) = I$ because $T_n(0) = I$ for all n. Thus $(T(t))_{t\geq 0}$ is a C_0-semigroup and $\|T(t)\| \leq M$ for all $t \geq 0$.

It remains to show that its generator coincides with the original operator A. For $x \in D(A)$ we compute by strong convergence and the formula for the derivative of matrix exponentials (or by dominated convergence applied to the integral formula) that

$$\lim_{h\downarrow 0} \frac{T(h)x - x}{h} = \lim_{h\downarrow 0} \lim_{n\to\infty} \frac{T_n(h)x - x}{h} = \lim_{n\to\infty} A_n x = Ax, \tag{3.36}$$

where the interchange of limits is justified by uniform convergence on $[0, h_0]$ for small $h_0 > 0$ together with the uniform boundedness of the semigroups T_n. Hence $D(A) \subset D(A_T)$ and $A_T x = Ax$ on $D(A)$. Conversely, if $y \in D(A_T)$ then the resolvent formula for the generator A_T gives

$$R(\lambda, A_T) = \int_0^\infty e^{-\lambda t} T(t)\,dt \tag{3.37}$$

for $\lambda > 0$, and by uniqueness of the resolvent one shows that $R(\lambda, A_T) = R(\lambda, A)$ on their common domain; therefore A_T is an extension of A and closed; but A was assumed closed, hence $A_T = A$. This shows that the generator of the limiting semigroup is exactly A.

This completes the sufficiency proof in the case $\omega = 0$. Returning to the general case $\omega \neq 0$, the previous argument applied to $B = A - \omega I$ yields a C_0-semigroup $S(t)$ with $\|S(t)\| \leq M$ and generator B. Then $T(t) := e^{\omega t} S(t)$ is a C_0-semigroup with

$$\|T(t)\| \leq M e^{\omega t} \tag{3.38}$$

and generator A. Uniqueness of the semigroup generated by A follows from the uniqueness of the Laplace transform: if two C_0-semigroups have the same generator then their Laplace transforms coincide on a half-plane and hence the semigroups coincide.

Remark on alternative approaches and technical points: The proof above follows the classical route via Yosida approximants $A_n = nAR(n, A)$. An alternative is to construct the semigroup via inverse Laplace transform from the resolvent; both approaches require the resolvent bounds (3.10) to ensure integrals converge and that the resulting family is strongly continuous. All interchange of limits and integrals used above is justified by uniform boundedness and dominated convergence: the uniform bound $\|T_n(t)\| \leq M$ supplies the necessary domination, while density of $D(A)$ and the strong convergence on $D(A)$ extend results to all $x \in X$. The estimate (3.10) is sharp in the sense that the necessity argument shows any generator with growth bound ω must satisfy it, while the sufficiency argument shows it is sufficient to construct the corresponding semigroup.

Thus the theorem is proved: a (densely defined, closed) operator A generates a C_0-semigroup with bound

$$\|T(t)\| \leq M e^{\omega t} \tag{3.39}$$

if and only if $(\omega, \infty) \subset \rho(A)$ and (3.10) holds for all $n \geq 1$.

3.3 Kolmogorov Equations

For diffusion processes, the generator A is elliptic (or hypoelliptic), and the semigroup (P_t) acts as a solution operator for the Kolmogorov backward equation

$$\frac{\partial u}{\partial t}(t, x) = Au(t, x), \qquad u(0, x) = f(x), \tag{3.40}$$

where $u(t, x) = (P_t f)(x)$. Dually, the transition densities $p_t(x, y)$, when they exist, satisfy the Kolmogorov forward (Fokker–Planck) equation:

$$\frac{\partial p_t}{\partial t}(x, y) = A^* p_t(x, y), \tag{3.41}$$

where A^* is the formal adjoint of A acting on suitable test functions or measures. Thus, semigroup theory provides the analytic bridge connecting the evolution of expected functionals $P_t f$ to the PDEs describing the evolution of distributions of the underlying process.

In the probabilistic framework, one can also interpret P_t as an expectation operator over the path space, and the infinitesimal generator A can be viewed as the limit of infinitesimal increments:

$$Af(x) = \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}_x[f(X_h) - f(x)]. \tag{3.42}$$

This probabilistic definition coincides with the analytic one in the Feller setting. Moreover, for any $f \in \mathcal{D}(A)$, the semigroup satisfies the differential relation

$$\frac{d}{dt} P_t f = P_t A f = A P_t f, \qquad t > 0, \tag{3.43}$$

and thus the infinitesimal generator governs the time evolution of the semigroup.

An important class of examples arises from Lévy processes, where the semigroup is convolutional and characterized by the Lévy–Khintchine formula. Consider a stochastic process $(X_t)_{t \geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $\mathbb{R}^d$. We assume that X_t is a Lévy process, i.e., a process with stationary and independent increments, càdlàg paths, and $X_0 = 0$ almost surely. For such a process, the associated Markov semigroup $(P_t)_{t \geq 0}$ acting on suitable test functions $f : \mathbb{R}^d \to \mathbb{R}$ is given by convolution with the law ν_t of X_t:

$$(P_t f)(x) = \int_{\mathbb{R}^d} f(x+y), \nu_t(dy), \quad t \geq 0, \ x \in \mathbb{R}^d. \tag{3.44}$$

The Lévy–Khintchine representation asserts the existence of a characteristic exponent $\psi : \mathbb{R}^d \to \mathbb{C}$ such that the characteristic function of X_t satisfies

$$\mathbb{E}\big[e^{i\xi \cdot X_t}\big] = e^{-t\psi(\xi)}, \quad \xi \in \mathbb{R}^d, \tag{3.45}$$

and ψ can be expressed in the canonical form

$$\psi(\xi) = ib \cdot \xi + \frac{1}{2} \xi \cdot Q\xi + \int_{\mathbb{R}^d \setminus 0} \Big(1 - e^{i\xi \cdot y} + i\xi \cdot y \mathbf{1}_{|y|<1}\Big) \Pi(dy), \tag{3.46}$$

where $b \in \mathbb{R}^d$ is the drift vector, $Q \in \mathbb{R}^{d \times d}$ is a positive semidefinite covariance matrix, and Π is the Lévy measure satisfying

$$\int_{\mathbb{R}^d \setminus 0} \min(1, |y|^2), \Pi(dy) < \infty \tag{3.47}$$

The infinitesimal generator (A) of the semigroup $(P_t)_{t \geq 0}$ is defined as the strong limit

$$Af(x) = \lim_{t \downarrow 0} \frac{(P_t f)(x) - f(x)}{t}, \quad x \in \mathbb{R}^d, \tag{3.48}$$

and for Lévy processes it admits a representation as a pseudo-differential operator via Fourier transform:

$$Af(x) = -\int_{\mathbb{R}^d} e^{i\xi \cdot x}, \psi(\xi), \hat{f}(\xi), \frac{d\xi}{(2\pi)^d}, \quad \hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-i\xi \cdot y} f(y), dy. \tag{3.49}$$

Equivalently, the generator can be expressed in the Lévy–Itô form:

$$Af(x) = b \cdot \nabla f(x) + \frac{1}{2} \mathrm{tr}(Q \nabla^2 f(x)) + \int_{\mathbb{R}^d \setminus 0} \Big(f(x+y) - f(x) - \mathbf{1}_{|y|<1} y \cdot \nabla f(x)\Big) \Pi(dy), \tag{3.50}$$

where ∇f and $\nabla^2 f$ denote the gradient and Hessian of f, respectively. This formulation unifies continuous diffusions $(Q \neq 0), (\Pi \equiv 0)$ and pure jump processes $(\Pi \not\equiv 0), (Q = 0)$ within the semigroup framework, demonstrating that the generator (A) captures the infinitesimal dynamics of both continuous and discontinuous stochastic evolutions.

Moreover, the semigroup P_t satisfies the Kolmogorov forward equation

$$\frac{d}{dt} P_t f = P_t A f, \quad P_0 f = f, \tag{3.51}$$

and the Kolmogorov backward equation

$$\frac{d}{dt} P_t f = A P_t f, \quad P_0 f = f, \tag{3.52}$$

ensuring a rigorous analytic characterization of the evolution of expectations under the Lévy process.

3.4 Generator as the Infinitesimal Limit of the Expectation Semigroup Acting on Path Functionals

In the framework of stochastic analysis and semigroup theory, one can interpret the infinitesimal generator A of a Markov process $(X_t)_{t\geq 0}$ as the infinitesimal limit of the associated expectation semigroup acting not only on bounded measurable functions but also on a class of *path-dependent functionals*. This viewpoint provides a rigorous bridge between probabilistic and analytic formulations, culminating in the celebrated Feynman–Kac representation.

Let $(E, \mathcal{E})$ be a locally compact separable metric space and let $(X_t)_{t\geq 0}$ be a strong Markov process with transition semigroup

$$(P_t f)(x) = \mathbb{E}_x[f(X_t)], \qquad f \in \mathcal{B}_b(E). \tag{3.53}$$

The infinitesimal generator A associated with $(P_t)_{t\geq 0}$ is defined in the strong operator topology by the limit

$$Af(x) = \lim_{t \downarrow 0} \frac{(P_t f)(x) - f(x)}{t}, \qquad f \in \mathcal{D}(A), \tag{3.54}$$

where $\mathcal{D}(A)$ denotes the domain of A consisting of functions for which the above limit exists in a suitable sense (e.g., uniformly on compact sets or in $L^2(E, \mu)$).

Now, let $V : E \to \mathbb{R}$ be a measurable potential function which is locally bounded from below. We define the *Feynman–Kac expectation semigroup* $(P_t^V)_{t\geq 0}$ by

$$(P_t^V f)(x) := \mathbb{E}_x\left[\exp\left(-\int_0^t V(X_s)\, ds\right) f(X_t)\right], \tag{3.55}$$

for every bounded measurable f. The factor $\exp\left(-\int_0^t V(X_s)\, ds\right)$ serves as a multiplicative functional that deforms the measure on the path space $\Omega := C([0, \infty); E)$, weighting each trajectory according to its accumulated potential energy.

It follows from standard semigroup theory (via the Trotter–Kato product formula or the Feynman–Kac–Itô expansion) that the family $(P_t^V)_{t\geq 0}$ forms a strongly continuous contraction semigroup on $L^2(E,\mu)$ with infinitesimal generator given by

$$A^V f = Af - Vf, \qquad f \in \mathcal{D}(A). \tag{3.56}$$

Equivalently, if we define

$$u(t,x) := (P_t^V f)(x), \tag{3.57}$$

then u satisfies the parabolic evolution equation

$$\frac{\partial u}{\partial t}(t,x) = (A - V)u(t,x), \qquad u(0,x) = f(x). \tag{3.58}$$

To establish this rigorously, observe that for small $t > 0$,

$$\frac{P_t^V f(x) - f(x)}{t} = \frac{\mathbb{E}_x\left[\left(e^{-\int_0^t V(X_s)\,ds} - 1\right) f(X_t)\right]}{t} + \frac{\mathbb{E}_x[f(X_t) - f(x)]}{t}. \tag{3.59}$$

Using the dominated convergence theorem and the continuity of $t \mapsto X_t$ almost surely, the second term converges to $Af(x)$, while the first term yields $-V(x)f(x)$. Hence,

$$\lim_{t\downarrow 0} \frac{(P_t^V f)(x) - f(x)}{t} = (A - V)f(x), \tag{3.60}$$

showing that $A^V = A - V$ in the infinitesimal sense. This analysis demonstrates that the generator A arises as the infinitesimal limit of the expectation semigroup acting on functionals of the stochastic path.

The inclusion of the potential V through the exponential weight naturally leads to the analytic–probabilistic correspondence embodied in the *Feynman–Kac formula*, which provides the unique mild solution to the Cauchy problem

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = (A - V)u(t,x), \\ u(0,x) = f(x), \end{cases} \tag{3.61}$$

via

$$u(t,x) = \mathbb{E}_x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right]. \tag{3.62}$$

Thus, the Feynman–Kac representation establishes a rigorous and profound connection between stochastic processes, infinitesimal generators, semigroup theory, and partial differential equations.

3.5 Conclusion

In this chapter, we have developed the rigorous mathematical foundations underlying the relationship between stochastic processes and their infinitesimal generators. Starting from the

Markov semigroup $(P_t)_{t\geq 0}$ acting on bounded measurable functions, we introduced its infinitesimal generator A as the operator describing the instantaneous rate of change of expected values. Through this framework, we demonstrated how analytic properties of A encode the probabilistic dynamics of the process (X_t), and conversely, how the probabilistic construction provides an operational interpretation of differential operators. We established that, under appropriate domain and continuity conditions, the generator A determines the entire semigroup (P_t) via the Hille–Yosida theorem, ensuring the existence, uniqueness, and strong continuity of the semigroup solution. This analytic result forms the backbone of the modern theory of Markov processes and evolution equations, unifying stochastic dynamics with operator semigroup methods in functional analysis.

Furthermore, by extending the semigroup to act on path-dependent functionals through the Feynman–Kac deformation

$$(P_t^V f)(x) = \mathbb{E}_x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right], \tag{3.63}$$

we revealed that the generator acquires a natural perturbation by the potential V, leading to the modified operator $A^V = A - V$. This observation provides the analytic foundation for the Feynman–Kac formula, establishing a deep correspondence between stochastic processes, semigroup evolution, and the solutions of parabolic partial differential equations.

In summary, the chapter clarified the fundamental identity:

$$\frac{\partial u}{\partial t} = Au, \qquad u(0) = f, \tag{3.64}$$

as the unifying structure connecting Markov processes, expectation semigroups, and infinitesimal generators. This interplay between analysis and probability—mediated by the semigroup formalism—forms the conceptual and technical cornerstone upon which the subsequent chapters on stochastic differential equations, Feynman–Kac representations, and potential theory are built.

References for this chapter: Ethier and Kurtz (2009) [11], Liggett (2025) [12], Davies (1989) [13], Dynkin (1965) [14] and Dynkin (2012) [15].

Part II

The Feynman–Kac Formula

Chapter 4

Stochastic Differential Equations and Martingale Problems

> "Stochastic differential equations are the grammar of random motion — each infinitesimal term encoding the law of uncertainty."
>
> — Adapted from Hiroshi Kunita

4.1 Stochastic Differential Equations: Strong and Weak Formulations

We begin by considering a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions, namely that the filtration is right-continuous and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets. Let $B = (B_t)_{t\geq 0}$ denote an (m)-dimensional standard Brownian motion adapted to $(\mathcal{F}_t)$, i.e. each component $B^{(i)}$ is an $(\mathcal{F}_t)$-Brownian motion, and for $1 \leq i, j \leq m$,

$$\mathbb{E}[B_t^{(i)} B_s^{(j)}] = (s \wedge t)\, \delta_{ij} \tag{4.1}$$

A stochastic differential equation (SDE) in $\mathbb{R}^d$ driven by B is a stochastic process $X = (X_t)_{t\in[0,T]}$ that satisfies, in its infinitesimal differential form,

$$dX_t = b(t, X_t), dt + \sigma(t, X_t), dB_t, \qquad X_0 = x \in \mathbb{R}^d, \tag{4.2}$$

where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is the drift vector field and $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is the diffusion coefficient matrix, both assumed to be Borel-measurable functions. The above equation should not be interpreted in a classical differential sense, since B_t is almost surely nowhere differentiable. Instead, it is understood in the sense of Itô integration, which provides a well-defined stochastic integral for square-integrable adapted integrands.

Formally, one defines a strong solution of the SDE as a continuous adapted process $X =$

$(X_t)_{t\in[0,T]}$ such that, almost surely,

$$X_t = x + \int_0^t b(s, X_s), ds + \int_0^t \sigma(s, X_s), dB_s, \quad \forall, t \in [0,T]. \tag{4.3}$$

The first integral is a Lebesgue integral in time, while the second is an Itô stochastic integral, whose existence requires that

$$\mathbb{E}\int_0^T |\sigma(s, X_s)|^2 ds < \infty \tag{4.4}$$

so that each component of the integrand is in the space $\mathcal{H}^2([0,T])$.

The canonical framework for establishing existence and uniqueness of such a strong solution is the Picard iteration method on the Banach space

$$\mathcal{S}^2([0,T];\mathbb{R}^d) := \Big\{ X : [0,T] \to \mathbb{R}^d \text{ adapted, continuous} \;\Big|\; \mathbb{E}\Big[\sup_{0\le t\le T} |X_t|^2\Big] < \infty \Big\}, \tag{4.5}$$

equipped with the norm

$$|X|_{\mathcal{S}^2} := \Big(\mathbb{E}\Big[\sup_{0\le t\le T} |X_t|^2\Big]\Big)^{1/2} \tag{4.6}$$

One defines recursively $X_t^{(0)} = x$ and

$$X_t^{(n+1)} = x + \int_0^t b(s, X_s^{(n)})\, ds + \int_0^t \sigma(s, X_s^{(n)})\, dB_s \tag{4.7}$$

If b and σ satisfy global Lipschitz continuity and linear growth bounds, namely that there exists $L > 0$ such that for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$,

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le L|x-y|, \tag{4.8}$$

and

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \le L^2(1 + |x|^2), \tag{4.9}$$

then one can rigorously show, by using Itô isometry and Burkholder–Davis–Gundy inequalities, that

$$|X^{(n+1)} - X^{(n)}|_{\mathcal{S}^2} \le C(L,T)\, |X^{(n)} - X^{(n-1)}|_{\mathcal{S}^2}, \tag{4.10}$$

with $C(L,T) < 1$ for small enough T, ensuring contraction. Therefore, the sequence $(X^{(n)})_n$ converges in $\mathcal{S}^2$ to a unique adapted continuous process X, which satisfies the integral equation above. This X is called the unique strong solution to the SDE.

A weak solution to the same SDE is defined differently: it is a triple $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t), \tilde{\mathbb{P}}; \tilde{B}, \tilde{X})$ consisting of a filtered probability space, an m-dimensional Brownian motion $\tilde{B}$, and a continuous adapted process $\tilde{X}$ satisfying

$$\tilde{X}_t = x + \int_0^t b(s, \tilde{X}_s), ds + \int_0^t \sigma(s, \tilde{X}_s), d\tilde{B}_s, \quad \tilde{\mathbb{P}}\text{-a.s. for all } t. \tag{4.11}$$

Here, unlike the strong formulation, both the probability space and the Brownian motion are not fixed in advance and may depend on the realization of $\tilde{X}$. Thus, the "weakness" refers not to probabilistic dependence, but to flexibility in the underlying stochastic basis.

We say the SDE admits uniqueness in law if, whenever (X, B) and $(\tilde{X}, \tilde{B})$ are two weak solutions (possibly defined on different spaces) with the same initial condition, then the laws of X and $\tilde{X}$ on $C([0, T]; \mathbb{R}^d)$ coincide:

$$\operatorname{Law}(X) = \operatorname{Law}(\tilde{X}). \tag{4.12}$$

Finally, the notions of pathwise uniqueness and uniqueness in law are logically distinct: pathwise uniqueness pertains to indistinguishability of two solutions defined on the same probability space with the same Brownian motion, while uniqueness in law concerns identity of their induced probability distributions on the space of continuous functions. The celebrated **Yamada–Watanabe theorem** establishes that pathwise uniqueness together with existence of a weak solution implies the existence of a unique strong solution. This foundational result tightly links the analytic and probabilistic formulations of stochastic dynamics and ensures that under suitable regularity, randomness propagates in a unique measurable way from the Brownian motion to the solution process.

4.2 The Martingale Problem Formulation

Let $\mathcal{L}$ be the second-order differential operator on $C_b^2(\mathbb{R}^d)$ given by

$$(\mathcal{L}f)(x) = \sum_{i=1}^{d} b_i(x), \partial_{x_i} f(x) + \tfrac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x), \partial_{x_i}\partial_{x_j} f(x), \qquad a(x) = \sigma(x)\sigma(x)^\top, \tag{4.13}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are Borel measurable and chosen so that the coefficients and their growth permit the integrals below to be defined. The martingale problem for $(\mathcal{L}, \mu)$ (with μ a probability on $\mathbb{R}^d$) is the following specification of a law on path space: let $\Omega = C([0, T]; \mathbb{R}^d)$ be endowed with the canonical process $X_t(\omega) = \omega(t)$ and the canonical filtration

$$\mathcal{F}_t^X = \sigma(X_s : 0 \le s \le t) \tag{4.14}$$

completed and made right-continuous. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F}_T^X)$ is a solution of the martingale problem for $(\mathcal{L}, \mu)$ if $\mathbb{P} \circ X_0^{-1} = \mu$ and for every $f \in C_b^2(\mathbb{R}^d)$ the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s), ds \tag{4.15}$$

is a $\mathbb{P}$-martingale with respect to $(\mathcal{F}_t^X)$. This definition encodes the infinitesimal generator property in a purely probabilistic (weak) form, without explicit reference to a driving Brownian motion.

If X is a continuous semimartingale solution of the SDE

$$dX_t = b(X_t), dt + \sigma(X_t), dB_t, \qquad X_0 \sim \mu, \tag{4.16}$$

defined on a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t), \tilde{\mathbb{P}})$ with Brownian motion B, then Itô's formula applied to $f \in C_b^2$ yields

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s), ds = \int_0^t \nabla f(X_s)\sigma(X_s), dB_s, \tag{4.17}$$

and the right-hand side is a local martingale (in fact a true martingale under integrability). Pushing forward the law of X to the canonical space produces a probability $\mathbb{P}$ solving the martingale problem for $(\mathcal{L}, \mu)$. Thus any (weak) solution of the SDE gives a solution of the martingale problem: existence of weak solutions $\Rightarrow$ existence of martingale problem solutions.

Conversely, suppose $\mathbb{P}$ solves the martingale problem for $(\mathcal{L}, \mu)$. Then by the Doob–Meyer decomposition and Lévy characterization techniques one may (under mild nondegeneracy and regularity hypotheses) recover a semimartingale decomposition of the canonical process: there exists a predictable finite-variation process A and a continuous local martingale M such that

$$X_t = X_0 + A_t + M_t, \tag{4.18}$$

and for all $f \in C_b^2$,

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\, ds \tag{4.19}$$

is a local martingale. From the martingale representation theorem in the Brownian filtration (or more generally from the predictable representation property on an appropriately enlarged probability space) one can represent the continuous local martingale M as a stochastic integral against an m-dimensional Brownian motion W:

$$M_t = \int_0^t \sigma(X_s), dW_s, \tag{4.20}$$

provided the quadratic variation of M matches $\int_0^\cdot a(X_s), ds$. Under such identification the finite-variation part is $A_t = \int_0^t b(X_s)\, ds$ and one obtains a weak solution of the SDE on some probability space carrying W. The step "martingale problem solution $\Rightarrow$ weak SDE solution" requires verification of the compatibility of the quadratic variation process with $a(x)$ and the application of a representation theorem; these hold under standard nondegeneracy (e.g. $a(x)$ measurable and locally bounded, or uniformly elliptic) and integrability conditions.

Well-posedness of the martingale problem means existence and uniqueness of the solution measure $\mathbb{P}$ for every initial law μ (or for each Dirac mass δ_x). If the martingale problem for $(\mathcal{L}, \delta_x)$ is well posed for all $x \in \mathbb{R}^d$, then the family $(\mathbb{P}_x)_{x\in\mathbb{R}^d}$ defines a Markov family and the canonical process under $\mathbb{P}_x$ is a Markov process with generator $\mathcal{L}$. In particular, uniqueness of the martingale problem implies uniqueness in law for the corresponding SDE: if two weak solutions produce laws $\mathbb{P}$ and $\mathbb{Q}$ on path space, both solve the martingale problem and hence coincide by uniqueness, so the marginals (and finite-dimensional distributions) agree.

Uniqueness in law is often proved by duality (via PDEs) or by coupling/analytic estimates. A classical analytic criterion is the Stroock–Varadhan uniqueness condition: if the operator $\mathcal{L}$ is uniformly elliptic and its coefficients are Hölder continuous (or satisfy suitable continuity and growth conditions), then the martingale problem for $\mathcal{L}$ is well posed. Analytically, well-posedness corresponds to existence and uniqueness for the forward/backward Kolmogorov equations in appropriate function spaces; probabilistically, it implies that the semigroup P_t associated to $\mathbb{P}_x$ is uniquely determined by $\mathcal{L}$.

Well-posedness of the martingale problem also yields additional pathwise and structural properties of the canonical process. If $\mathbb{P}_x$ solves the martingale problem and $(\mathbb{P}_x)_x$ is Borel measurable in x, then the canonical process has the strong Markov property under each $\mathbb{P}_x$. Moreover, Feller properties of the semigroup (e.g. P_t maps C_0 into itself and is strongly continuous) follow from further regularity of coefficients; conversely, Feller semigroups yield uniqueness of the martingale problem in the Feller class.

The martingale problem formulation is particularly well suited to weak convergence and approximation arguments. If $\mathcal{L}^n$ is a sequence of generators with corresponding martingale problem solutions $\mathbb{P}^n$ and $\mathcal{L}^n \to \mathcal{L}$ in an appropriate sense (e.g. on a core of test functions), and if $(\mathbb{P}^n)$ is tight on $C([0,T];\mathbb{R}^d)$, then any subsequential limit $\mathbb{P}$ solves the martingale problem for $\mathcal{L}$. Hence convergence of generators plus tightness implies convergence of laws; this is the backbone of diffusion approximation, homogenization, and weak numerical analysis.

Finally, the martingale problem provides a convenient framework for extending the theory beyond classical SDEs: generators may be pseudo-differential (Lévy type), degenerate, or time-dependent, and one can define martingale problems for such operators without explicit stochastic differential representations. Existence, uniqueness, and stability results are thus formulated and proved in the martingale problem language, which, by the equivalences described above, yields corresponding statements about weak solutions of SDEs whenever a stochastic integral representation is available.

4.2.1 Stroock–Varadhan uniqueness condition

The Stroock–Varadhan uniqueness condition provides one of the most profound and rigorous analytical characterizations of when a martingale problem admits a unique solution in law. It forms the cornerstone of the connection between stochastic differential equations (SDEs) and partial differential operators, and gives precise sufficient conditions ensuring that the stochastic dynamics governed by a generator $\mathcal{L}$ are uniquely determined by their infinitesimal characteristics.

Consider a second-order differential operator of the form

$$(\mathcal{L}f)(x) = \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x), \tag{4.21}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is a drift vector field and $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is a symmetric, nonnegative-definite diffusion matrix, typically expressed as $a(x) = \sigma(x)\sigma(x)^\top$ for some measurable function $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$.

Let $C([0,\infty);\mathbb{R}^d)$ denote the canonical path space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$, equipped with the canonical filtration $(\mathcal{F}_t^X)$ generated by the coordinate process $X_t(\omega) = \omega(t)$. A probability measure $\mathbb{P}$ on this path space is said to solve the martingale problem for $(\mathcal{L}, \delta_x)$ if $\mathbb{P}(X_0 = x) = 1$ and for all test functions $f \in C_b^2(\mathbb{R}^d)$, the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (\mathcal{L}f)(X_s)\,ds \tag{4.22}$$

is a $\mathbb{P}$-martingale with respect to $(\mathcal{F}_t^X)$.

The Stroock–Varadhan framework addresses the uniqueness in law of this problem, i.e., the question of whether there exists at most one probability measure $\mathbb{P}$ satisfying the above property for a given initial distribution and generator. Their celebrated uniqueness theorem gives conditions on b and a that ensure this uniqueness.

To state the condition rigorously, define the following hypotheses. Let $a(x)$ and $b(x)$ be measurable functions satisfying:

1. **Local boundedness:** For every compact set $K \subset \mathbb{R}^d$, there exists a constant $C_K > 0$ such that
$$\sup_{x\in K} \left(|b(x)| + |a(x)|\right) \le C_K. \tag{4.23}$$

2. **Local ellipticity** For every compact set $K \subset \mathbb{R}^d$, there exists a constant $\lambda_K > 0$ such that
$$\xi^\top a(x), \xi \ge \lambda_K |\xi|^2 \quad \text{for all } x \in K,\ \xi \in \mathbb{R}^d. \tag{4.24}$$
That is, $a(x)$ is uniformly positive definite on compact sets.

3. **Continuity condition** The coefficients $a(x)$ and $b(x)$ are locally continuous functions, i.e., for each compact set K,
$$\lim_{y\to x,,x,y\in K} \left(|b(x) - b(y)| + |a(x) - a(y)|\right) = 0. \tag{4.25}$$

Under these hypotheses, the Stroock–Varadhan uniqueness theorem asserts that the martingale problem for $(\mathcal{L}, \delta_x)$ admits a unique solution in law for each initial condition $x \in \mathbb{R}^d$.

Formally, the theorem can be stated as follows:

Theorem 4.1 (Stroock–Varadhan Uniqueness Theorem)**.** *Suppose that the coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are bounded, measurable, locally uniformly continuous, and that $a(x)$ is uniformly nondegenerate on compact subsets of $\mathbb{R}^d$. Then, for each initial point $x \in \mathbb{R}^d$, there exists a unique probability measure $\mathbb{P}_x$ on $C([0,\infty);\mathbb{R}^d)$ solving the martingale problem for $(\mathcal{L}, \delta_x)$.*

Proof. Let $b : \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ satisfy the assumptions: they are bounded, measurable, locally uniformly continuous, and $a(x)$ is uniformly nondegenerate on compact subsets, i.e., for every compact $K \subset \mathbb{R}^d$, there exists $\lambda_K > 0$ such that

$$\xi^\top a(x)\xi \ge \lambda_K |\xi|^2, \quad \forall x \in K,\ \forall \xi \in \mathbb{R}^d. \tag{4.26}$$

Define the generator

$$\mathcal{L}f(x) = \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x), \quad f \in C_c^\infty(\mathbb{R}^d), \tag{4.27}$$

and recall that a probability measure $\mathbb{P}_x$ on $C([0,\infty);\mathbb{R}^d)$ solves the martingale problem for $(\mathcal{L}, \delta_x)$ if $X_0 = x$ $\mathbb{P}_x$-a.s. and

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s), ds \tag{4.28}$$

is a martingale with respect to the canonical filtration $(\mathcal{F}_t)_{t\geq 0}$ for all $f \in C_c^\infty(\mathbb{R}^d)$.

To construct a solution, consider a standard mollifier $\rho_\varepsilon \in C_c^\infty(\mathbb{R}^d)$, $\rho_\varepsilon \geq 0$, $\int \rho_\varepsilon = 1$, and define

$$b_\varepsilon(x) := (b*\rho_\varepsilon)(x) = \int_{\mathbb{R}^d} b(y)\rho_\varepsilon(x-y), dy, \quad a_\varepsilon(x) := (a*\rho_\varepsilon)(x) = \int_{\mathbb{R}^d} a(y)\rho_\varepsilon(x-y), dy. \tag{4.29}$$

Then $b_\varepsilon, a_\varepsilon \in C^\infty(\mathbb{R}^d)$ are bounded and converge locally uniformly to b and a as $\varepsilon \downarrow 0$. Moreover, since convolution preserves uniform ellipticity on compacts, $a_\varepsilon(x)$ remains uniformly nondegenerate on compacts.

For each $\varepsilon > 0$, consider the stochastic differential equation

$$dX_t^\varepsilon = b_\varepsilon(X_t^\varepsilon), dt + \sigma_\varepsilon(X_t^\varepsilon)\, dW_t, \quad X_0^\varepsilon = x \tag{4.30}$$

where $\sigma_\varepsilon(x)$ satisfies $\sigma_\varepsilon(x)\sigma_\varepsilon(x)^\top = a_\varepsilon(x)$. By classical SDE theory, since b_ε and σ_ε are smooth and globally bounded, there exists a unique strong solution X_t^ε. Let $\mathbb{P}_x^\varepsilon$ denote the law of X^ε on $C([0,\infty);\mathbb{R}^d)$. Next, establish tightness of $(\mathbb{P}_x^\varepsilon)_{\varepsilon>0}$. For any $T > 0$ and $\delta > 0$, the Burkholder–Davis–Gundy inequality and boundedness of $b_\varepsilon, \sigma_\varepsilon$ imply

$$\mathbb{E}\big[|X_{t+s}^\varepsilon - X_t^\varepsilon|^4\big] \leq Cs^2, \quad 0 \leq t \leq t+s \leq T, \tag{4.31}$$

for a constant $C > 0$ independent of ε. Kolmogorov's continuity criterion then guarantees uniform control over modulus of continuity, so $(\mathbb{P}_x^\varepsilon)$ is tight in $C([0,T];\mathbb{R}^d)$, and by Prokhorov's theorem, there exists a subsequence $\varepsilon_n \downarrow 0$ and a limit law $\mathbb{P}_x$ such that

$$\mathbb{P}_x^{\varepsilon_n} \xrightarrow{\text{weakly}} \mathbb{P}_x \quad \text{on } C([0,\infty);\mathbb{R}^d). \tag{4.32}$$

To show $\mathbb{P}_x$ solves the martingale problem, fix $f \in C_c^\infty(\mathbb{R}^d)$ and consider

$$M_t^{f,\varepsilon_n} = f(X_t^{\varepsilon_n}) - f(X_0^{\varepsilon_n}) - \int_0^t \mathcal{L}_{\varepsilon_n} f(X_s^{\varepsilon_n})\, ds, \tag{4.33}$$

where

$$\mathcal{L}_{\varepsilon_n} f = b_{\varepsilon_n} \cdot \nabla f + \frac{1}{2}\mathrm{tr}(a_{\varepsilon_n} \nabla^2 f) \tag{4.34}$$

By construction, M_t^{f,ε_n} is a martingale under $\mathbb{P}_x^{\varepsilon_n}$. Since $b_{\varepsilon_n} \to b$ and $a_{\varepsilon_n} \to a$ locally uniformly, and $X^{\varepsilon_n} \to X$ in distribution, we have

$$\int_0^t \mathcal{L}_{\varepsilon_n} f(X_s^{\varepsilon_n}), ds \to \int_0^t \mathcal{L} f(X_s), ds \quad \text{in distribution.} \tag{4.35}$$

By martingale convergence theorems (specifically, convergence in law plus uniform integrability), the limit

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{L} f(X_s), ds \tag{4.36}$$

is a martingale under $\mathbb{P}_x$. This proves **existence** of a solution.

For **uniqueness**, suppose $\mathbb{P}_x$ and $\tilde{\mathbb{P}}_x$ are two solutions to the martingale problem. Define

$$P_t f(x) := \mathbb{E}_{\mathbb{P}_x}[f(X_t)], \quad \tilde{P}_t f(x) := \mathbb{E} * \tilde{\mathbb{P}}_x[f(X_t)], \quad f \in C_c^\infty(\mathbb{R}^d). \tag{4.37}$$

Then $u(t,x) := P_t f(x)$ and $\tilde{u}(t,x) := \tilde{P}_t f(x)$ satisfy the Kolmogorov backward equation

$$\frac{\partial u}{\partial t}(t,x) = \mathcal{L} u(t,x), \quad u(0,x) = f(x), \tag{4.38}$$

and similarly for $\tilde{u}$. By Aronson-type Gaussian bounds, the fundamental solution of $\partial_t - \mathcal{L}$ is unique in the class of bounded continuous functions; hence

$$P_t f(x) = \tilde{P}_t f(x), \quad \forall t \geq 0, \ f \in C_c^\infty(\mathbb{R}^d). \tag{4.39}$$

Finally, since $C_c^\infty(\mathbb{R}^d)$ is dense in $C_0(\mathbb{R}^d)$, equality of all finite-dimensional distributions follows: for any $0 \leq t_1 < \cdots < t_n$ and bounded continuous $\phi : \mathbb{R}^{dn} \to \mathbb{R}$,

$$\mathbb{E}_{\mathbb{P}_x}[\phi(X_{t_1}, \ldots, X_{t_n})] = \mathbb{E}_{\tilde{\mathbb{P}}_x}[\phi(X_{t_1}, \ldots, X_{t_n})] \tag{4.40}$$

By Kolmogorov extension theorem, $\mathbb{P}_x = \tilde{\mathbb{P}}_x$ as measures on $C([0,\infty);\mathbb{R}^d)$. Therefore, the martingale problem for $(\mathcal{L}, \delta_x)$ admits a unique solution. □

In summary the proof of this result is profoundly analytical and rests upon the correspondence between the martingale problem and the associated Kolmogorov backward equation:

$$\frac{\partial u}{\partial t}(t,x) = (\mathcal{L} u)(t,x), \qquad u(0,x) = f(x), \tag{4.41}$$

where $u(t,x) = \mathbb{E}_x[f(X_t)]$ is the transition semigroup associated with the diffusion. The uniqueness of the martingale problem then follows from the uniqueness of classical or weak solutions to this parabolic partial differential equation under the prescribed regularity and ellipticity conditions.

In the degenerate or non-elliptic case, additional structural hypotheses are often imposed to ensure uniqueness, such as Hörmander's condition for hypoelliptic diffusions, which requires the Lie algebra generated by the vector fields associated with b and σ to span the tangent space at every point.

Hence, the Stroock–Varadhan uniqueness condition provides a purely analytic criterion—expressed in terms of the coefficients (a) and (b)—for the well-posedness of the martingale problem, ensuring that the stochastic dynamics governed by $\mathcal{L}$ are uniquely characterized in distribution by their infinitesimal generator. This establishes a deep equivalence between the uniqueness of the martingale problem, the well-posedness of the SDE, and the uniqueness of solutions to the associated PDE, thereby unifying probabilistic and analytic formulations of diffusion processes.

4.3 Connection between Generators and SDEs

If $X = (X_t)_{t\geq 0}$ is a continuous adapted process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and it satisfies the Itô stochastic differential equation

$$dX_t = b(X_t)\,dt + \sigma(X_t), dB_t \qquad X_0 = x, \tag{4.42}$$

where B is an m-dimensional $(\mathcal{F}_t)$-Brownian motion, $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are measurable and such that the usual integrability conditions hold, then for every $f \in C_b^2(\mathbb{R}^d)$ Itô's formula applies pathwise and yields the identity

$$f(X_t) - f(X_0) = \int_0^t \nabla f(X_s)\cdot b(X_s)\,ds + \int_0^t \nabla f(X_s)\sigma(X_s)\,dB_s + \tfrac{1}{2}\int_0^t \operatorname{tr}\big(\sigma\sigma^\top(X_s)D^2 f(X_s)\big)\,ds. \tag{4.43}$$

Collecting the finite-variation terms defines the differential operator $\mathcal{L}$ acting on C_b^2 by

$$(\mathcal{L}f)(x) := b(x)\cdot\nabla f(x) + \tfrac{1}{2}\operatorname{tr}\big(\sigma\sigma^\top(x)D^2 f(x)\big), \tag{4.44}$$

so that the previous display is equivalently written as the decomposition

$$f(X_t) - f(X_0) - \int_0^t (\mathcal{L}f)(X_s)\,ds = \int_0^t \nabla f(X_s)\sigma(X_s)\,dB_s. \tag{4.45}$$

The right-hand side is an Itô stochastic integral with respect to Brownian motion and hence is a (local and, under integrability, true) martingale with respect to $(\mathcal{F}_t)$. Consequently, defining

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (\mathcal{L}f)(X_s), ds, \tag{4.46}$$

we obtain that M^f is a local martingale for every $f \in C_b^2(\mathbb{R}^d)$, and under the standard moment bounds it is a true square-integrable martingale. This verifies that any strong (or weak) solution of the SDE furnishes a solution of the martingale problem associated with the generator $\mathcal{L}$.

Conversely, suppose $\mathbb{P}$ is a probability measure on canonical path space $\Omega = C([0,T];\mathbb{R}^d)$ under which the canonical process $X_t(\omega) = \omega(t)$ satisfies the martingale property

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (\mathcal{L}f)(X_s)\,ds \tag{4.47}$$

is a $\mathbb{P}$-martingale for every $f \in C_b^2(\mathbb{R}^d)$, with $\mathcal{L}$ as above and with initial law concentrated at x. Under additional structural and nondegeneracy assumptions on the coefficients (for example local boundedness of b and σ, and the identification $a(x) = \sigma(x)\sigma(x)^\top$ as a predictable matrix-valued function), one can recover a semimartingale decomposition of the canonical process.

Precisely, by applying the Doob–Meyer decomposition and localization arguments one finds predictable finite-variation processes (A) and a continuous local martingale M such that

$$X_t = X_0 + A_t + M_t, \tag{4.48}$$

and for every $f \in C_b^2$ the identity

$$f(X_t) - f(X_0) - \int_0^t \nabla f(X_s), dA_s - \tfrac{1}{2}\sum_{i,j}\int_0^t \partial_{ij} f(X_s), d\langle M^{(i)}, M^{(j)}\rangle_s \tag{4.49}$$

is a local martingale. Matching this representation with the martingale problem relation yields the identification

$$dA_s = b(X_s), ds, \qquad d\langle M^{(i)}, M^{(j)}\rangle_s = a_{ij}(X_s)\, ds, \tag{4.50}$$

in the sense of predictable processes, where $a(x) = \sigma(x)\sigma(x)^\top$. Once the quadratic covariation structure of the continuous local martingale part M is identified as above, the martingale representation (or Kunita–Watanabe orthogonal decomposition) guarantees—possibly on an enlarged filtered probability space carrying an m-dimensional Brownian motion W—the existence of a predictable matrix-valued process H_s with

$$M_t = \int_0^t H_s, dW_s, \tag{4.51}$$

and the condition $H_s H_s^\top = a(X_s)$ (a.s.) yields $H_s = \sigma(X_s)$ up to an orthogonal rotation in the driving Brownian components. Thus one constructs a weak solution $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t), \widetilde{\mathbb{P}}, \widetilde{B}, \widetilde{X})$ of the SDE

$$d\widetilde{X}_t = b(\widetilde{X}_t)\, dt + \sigma(\widetilde{X}_t)\, d\widetilde{B}_t, \tag{4.52}$$

whose law on path space coincides with $\mathbb{P}$. The construction requires verification that the predictable quadratic variation of M is absolutely continuous with respect to Lebesgue measure and equals $\int_0^\cdot a(X_s), ds$; this is assured under standard hypotheses such as local boundedness and measurability of a, or uniform ellipticity which simplifies inversion. Under these hypotheses the martingale problem solution yields a weak SDE solution and hence the two formulations are equivalent.

Putting the two directions together, we obtain the precise correspondence: every (weak or strong) solution of the SDE induces a solution of the martingale problem for the generator $\mathcal{L}$; conversely, every solution $\mathbb{P}$ of the martingale problem that satisfies the quadratic-variation identification and the representability of its martingale part as an Itô integral can be realized as the law of a weak solution of the SDE with coefficients b, σ. In particular, uniqueness of the martingale problem for $(\mathcal{L}, \delta_x)$ implies uniqueness in law for the SDE started at x, while pathwise (strong) uniqueness together with existence of a weak solution implies existence and uniqueness of a strong solution (Yamada–Watanabe).

4.4 Weak Convergence and Stability

The martingale problem furnishes a robust and convenient framework for proving weak convergence of families of Markov processes by reducing convergence of path-space laws to convergence of finite-dimensional generators and tightness estimates. We give a precise, rigorous formulation and proof sketch of the standard scheme; all statements below are phrased in the canonical path space $\Omega = C([0,T];\mathbb{R}^d)$ (or $D([0,T];\mathbb{R}^d)$ when discontinuities occur) equipped with its Skorokhod topology and canonical filtration $(\mathcal{F}_t^X)$.

Let $(\mathcal{L}^n)_{n\geq 1}$ be a sequence of linear operators acting on a common core $\mathcal{C} \subset C_b^2(\mathbb{R}^d)$ (i.e. $\mathcal{C}$ is dense in the relevant function space and each $\mathcal{L}^n$ maps $\mathcal{C}$ into $C_b(\mathbb{R}^d)$). Suppose for each n there exists a unique probability measure $\mathbb{P}^n$ on $(\Omega, \mathcal{F}_T^X)$ solving the martingale problem for $(\mathcal{L}^n, \mu^n)$ with initial law μ^n. That is, for every $f \in \mathcal{C}$ the process

$$M_t^{n,f} := f(X_t) - f(X_0) - \int_0^t (\mathcal{L}^n f)(X_s)\, ds \tag{4.53}$$

is a $\mathbb{P}^n$-martingale, and $\mathbb{P}^n \circ X_0^{-1} = \mu^n$. Assume the initial laws μ^n converge weakly to μ on $\mathbb{R}^d$. We describe sufficient conditions under which $\mathbb{P}^n$ converges weakly (on Ω) to the unique solution $\mathbb{P}$ of the martingale problem for $(\mathcal{L}, \mu)$, where $\mathcal{L}$ is a limiting operator defined on the same core $\mathcal{C}$.

A first necessary analytic hypothesis is convergence of generators on the core: for every $f \in \mathcal{C}$ and every compact $K \subset \mathbb{R}^d$,

$$\sup_{x\in K} |(\mathcal{L}^n f)(x) - (\mathcal{L} f)(x)| \xrightarrow[n\to\infty]{} 0 \tag{4.54}$$

Equivalently, $\mathcal{L}^n f \to \mathcal{L} f$ uniformly on compacts for each $f \in \mathcal{C}$. This property is the natural substitute for pointwise convergence of infinitesimal generators and ensures that the martingale relations of the approximating laws become, in the limit, the martingale relations for $\mathcal{L}$.

A second, probabilistic ingredient is tightness of the sequence $(\mathbb{P}^n)$ in $\mathcal{P}(\Omega)$. Concretely, one must verify conditions (for example Aldous' criterion or moment–modulus bounds) that guarantee precompactness. A common sufficient condition is the existence of constants $p > 1$ and $C_T > 0$ such that for every n and every $0 \leq s < t \leq T$,

$$\mathbb{E}^{\mathbb{P}^n}\left[|X_t - X_s|^p\right] \leq C_T |t-s|^{1+\delta} \tag{4.55}$$

for some $\delta > 0$; Kolmogorov–Chentsov then yields tightness in $C([0,T];\mathbb{R}^d)$. Alternatively, in the càdlàg setting one uses bounds on oscillation with respect to stopping times (Aldous) together with control of large jumps.

Assume these two conditions hold. By Prokhorov's theorem the family $(\mathbb{P}^n)$ is relatively compact: every subsequence admits a further subsequence converging weakly to some probability measure $\mathbb{P}^\infty$ on Ω. It remains to identify any such limit point as a solution of the martingale problem for $(\mathcal{L}, \mu)$. Fix $f \in \mathcal{C}$ and let $\phi : \Omega \to \mathbb{R}$ be any bounded, continuous, and $\mathcal{F}_s^X$-measurable test functional (for some $0 \leq s \leq t \leq T$); examples include bounded continuous functions of the coordinates $(X_{r_1}, \ldots, X_{r_k})$ with $r_i \leq s$. For each n the martingale property gives

$$\mathbb{E}^{\mathbb{P}^n}\left[(f(X_t) - f(X_s) - \int_s^t \mathcal{L}^n f(X_u), du), \phi\right] = 0. \tag{4.56}$$

By weak convergence $\mathbb{P}^n \Rightarrow \mathbb{P}^\infty$ along the chosen subsequence and dominated convergence (justified by uniform moment bounds and continuity of ϕ), one may pass to the limit in the first term:

$$\mathbb{E}^{\mathbb{P}^n}\left[(f(X_t) - f(X_s))\phi\right] \xrightarrow[n\to\infty]{} \mathbb{E}^{\mathbb{P}^\infty}\left[(f(X_t) - f(X_s))\phi\right]. \tag{4.57}$$

For the integral term use the generator convergence on compacts together with tightness to deduce that

$$\mathbb{E}^{\mathbb{P}^n}\Big[\Big(\int_s^t \mathcal{L}^n f(X_u)\,du\Big)\phi\Big] = \mathbb{E}^{\mathbb{P}^n}\Big[\int_s^t (\mathcal{L}f)(X_u)\,du\cdot\phi\Big] + R_n(\phi), \tag{4.58}$$

where the remainder satisfies $R_n(\phi) \to 0$ as $n \to \infty$. Convergence of the expectations of the main term again follows from weak convergence plus uniform integrability (ensured by moment bounds). Thus the limit measure $\mathbb{P}^\infty$ satisfies

$$\mathbb{E}^{\mathbb{P}^\infty}\left[(f(X_t) - f(X_s) - \int_s^t \mathcal{L}f(X_u), du), \phi\right] = 0 \tag{4.59}$$

for every bounded continuous $\mathcal{F}_s^X$-measurable ϕ. By the separating properties of such test functions this identity is equivalent to the assertion that the process

$$M_t^{\infty,f} := f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_u), du \tag{4.60}$$

is a $\mathbb{P}^\infty$-martingale for every $f \in \mathcal{C}$; moreover $\mathbb{P}^\infty \circ X_0^{-1} = \mu$. Hence any weak limit point $\mathbb{P}^\infty$ solves the martingale problem for $(\mathcal{L}, \mu)$.

When the martingale problem for $(\mathcal{L}, \mu)$ is well posed, i.e. admits a unique solution $\mathbb{P}$, it follows that every subsequential limit equals $\mathbb{P}$; relative compactness therefore implies the full sequence $\mathbb{P}^n \Rightarrow \mathbb{P}$ converges weakly to $\mathbb{P}$. In summary, the two-step program (generator convergence on a core $\mathcal{C}$ plus tightness/uniform integrability) yields identification of limit points as martingale-problem solutions and, via uniqueness, convergence of the sequence of laws.

This abstract scheme is the backbone of many concrete applications. In diffusion approximation one takes $\mathcal{L}^n$ to be generators of rescaled jump processes (or Markov chains) and verifies $\mathcal{L}^n f \to \mathcal{L}f$ on a core together with tightness via moment and jump control; the limiting generator $\mathcal{L}$ is typically a second-order differential operator and uniqueness of its martingale problem is obtained via Stroock–Varadhan criteria (e.g. uniform ellipticity and Hölder continuity). In numerical analysis of SDEs, one shows that the discrete-time Markov semigroups of the scheme have generators approximating that of the SDE and uses the martingale problem method to prove convergence of the interpolated schemes. In stochastic homogenization one considers $\mathcal{L}^\varepsilon$ depending on a small parameter and shows convergence to an effective operator $\mathcal{L}^0$ on the core via two-scale expansion; tightness and identification of limits then yield weak convergence of the processes.

Two technical remarks are in order. First, one often needs uniform moment bounds to justify interchanging limits and expectations and to ensure uniform integrability of the martingale increments; these are verified by Lyapunov function techniques producing estimates of the form

$$\sup_n \mathbb{E}^{\mathbb{P}^n}[\sup_{t\le T} |X_t|^p] < \infty \tag{4.61}$$

Second, when working in the càdlàg Skorokhod space $D([0,T];\mathbb{R}^d)$ the core $\mathcal{C}$ and generator convergence hypotheses must be tailored to account for jump contributions (one typically requires convergence of jump kernels or of Lévy characteristics), and tightness is checked via Aldous' condition plus control of large jumps. Under these refinements the martingale problem method remains fully effective.

4.5 Conclusion

Stochastic differential equations and martingale problems together form the analytic and probabilistic backbone of modern stochastic process theory. The generator $\mathcal{L}$ acts as the infinitesimal descriptor of a Markov process, while the martingale problem provides an abstract weak formulation independent of the specific noise realization. Their interplay underlies deep results connecting probability theory, functional analysis, and partial differential equations.

References for this chapter: Karatzas & Shreve (2014) [1], Øksendal (2013) [2], Stroock and Varadhan (2007) [16], Rogers and Williams (2000) [17], Friedman (1975) [18].

Chapter 5

Statement and Proof of the Classical Feynman–Kac Formula

"Probabilistic representation of PDEs is not an accident; it is the shadow of a deeper semigroup and martingale structure."

— Adapted from E. B. Dynkin

5.1 Introduction and hypotheses

Let $E \subseteq \mathbb{R}^d$ be open (often $E = \mathbb{R}^d$), and let $\mathcal{L}$ be a second-order differential operator of the form

$$\mathcal{L}f(x) = \sum_{i=1}^{d} b_i(x)\, \partial_{x_i} f(x) + \tfrac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x)\, \partial_{x_i}\partial_{x_j} f(x), \tag{5.1}$$

with measurable coefficients $b : E \to \mathbb{R}^d$ and $a : E \to \mathbb{R}^{d\times d}$ symmetric and nonnegative definite. We assume the standard hypotheses that guarantee existence of a Markov process $X = (X_t)_{t\geq 0}$ with generator $\mathcal{L}$ (or of a Feller semigroup $(P_t)_{t\geq 0}$ with generator $\mathcal{L}$ on a core of smooth compactly supported functions): local boundedness of coefficients, local ellipticity on compacts or other conditions ensuring well-posedness of the martingale problem for $\mathcal{L}$. Fix a time horizon $T > 0$. Let $V : E \to \mathbb{R}$ be a measurable potential; the classical (bounded) case assumes $V \in C_b(E)$ (or at least V bounded measurable), and the source/terminal data $f : E \to \mathbb{R}$ is assumed bounded and continuous, $f \in C_b(E)$. When V is unbounded one must impose Kato-type or integrability conditions; we treat the bounded case first for clarity.

We denote by $(\mathbb{P}_x)_{x\in E}$ the law of the Markov process started at x and write $\mathbb{E}_x[\cdot]$ for expectation under $\mathbb{P}_x$. For $0 \leq t \leq T$ define the Feynman–Kac functional

$$u(t,x) := \mathbb{E}_x\Big[\exp\Big(-\int_0^t V(X_s)\, ds\Big)\, f(X_t)\Big]. \tag{5.2}$$

Our goal is to prove that u is the unique (classical) solution of the initial-value problem

$$\begin{cases} \partial_t u(t,x) = \mathcal{L}u(t,x) - V(x)\,u(t,x), & (t,x) \in (0,T] \times E, \\ u(0,x) = f(x), & x \in E, \end{cases} \tag{FK}$$

under the stated regularity and integrability hypotheses.

5.2 Statement of the classical theorem

Theorem 5.1 (Classical Feynman–Kac). *Assume $\mathcal{L}$ generates a Feller Markov process X on E, $V \in C_b(E)$, and $f \in C_b(E)$. Then the function*

$$u(t,x) := \mathbb{E}_x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \qquad (t,x) \in [0,T] \times E, \tag{5.3}$$

is bounded and continuous on $[0,T] \times E$, belongs to $C^{1,2}((0,T] \times E)$ whenever the coefficients of $\mathcal{L}$ are sufficiently smooth, and solves the backward parabolic problem (FK). Moreover, the solution is unique in the class of bounded continuous functions satisfying the equation in the classical (or appropriate weak) sense.

5.3 Proof of existence: martingale argument and verification

Fix f, V as above and define u by the displayed expectation. Boundedness and continuity of u follow by dominated convergence: since V and f are bounded there is a constant C with

$$|e^{-\int_0^t V(X_s)\,ds} f(X_t)| \le C \tag{5.4}$$

and $(t,x) \mapsto u(t,x)$ is continuous by the Feller property of the transition semigroup and dominated convergence. To derive the PDE, fix $x \in E$ and work on the probability space carrying X under $\mathbb{P}_x$. For each $s \in [0,t]$ define the multiplicative functional

$$Y_s := \exp\Big(-\int_0^s V(X_r)\,dr\Big). \tag{5.5}$$

By the Markov property and the tower property,

$$u(t,x) = \mathbb{E}_x[Y_t f(X_t)] = \mathbb{E}_x[Y_s\,\mathbb{E}_{X_s}[e^{-\int_0^{t-s} V(X_r)\,dr} f(X_{t-s})]] = \mathbb{E}_x[Y_s\,u(t-s, X_s)]. \tag{5.6}$$

Equivalently, for fixed t,

$$u(t,x) = \mathbb{E}_x[e^{-\int_0^{t\wedge\tau} V(X_r)\,dr} u(t-(t\wedge\tau), X_{t\wedge\tau})] \tag{5.7}$$

for any stopping time τ bounded by t; this representation will be used to localize the argument.

We now compute time-derivatives via a martingale (Dynkin) identity. For $f \in C_b^2$ define the process

$$M_s := Y_s\,u(t-s, X_s), \qquad 0 \le s \le t. \tag{5.8}$$

Applying product-rule for semimartingales (or Itô's formula for the continuous case) to Y_s and $u(t-s, X_s)$ yields, using $dY_s = -Y_s V(X_s)\,ds$ and the martingale decomposition for $u(t-s, X_s)$ as a function of X_s,

$$\begin{aligned} dM_s &= d(Y_s)\,u(t-s, X_s) + Y_s\,d\big(u(t-s, X_s)\big) + d\langle Y, u(\cdot, X_\cdot)\rangle_s \\ &= -Y_s V(X_s) u(t-s, X_s)\,ds + Y_s\big(-\partial_t u(t-s, X_s)\,ds + \mathcal{L}u(t-s, X_s)\,ds\big) + \text{martingale terms.} \end{aligned} \tag{5.9}$$

The finite-variation terms combine to

$$Y_s\Big(-\partial_t u(t-s, X_s) + \mathcal{L}u(t-s, X_s) - V(X_s)u(t-s, X_s)\Big)\,ds. \tag{5.10}$$

However, by the Markov property and the definition of u, the process M_s is a martingale (indeed $\mathbb{E}_x[M_s] = u(t, x)$ for all s), so its finite-variation part must vanish. Therefore, for almost every s and $\mathbb{P}_x$-almost surely,

$$-\partial_t u(t-s, X_s) + \mathcal{L}u(t-s, X_s) - V(X_s)u(t-s, X_s) = 0. \tag{5.11}$$

Since s was arbitrary in $[0, t]$ and X_s attains all points in E with positive probability in arbitrarily small time under the Feller/nondegeneracy hypotheses, we deduce the pointwise PDE

$$\partial_t u(\tau, x) = \mathcal{L}u(\tau, x) - V(x)u(\tau, x), \qquad \tau \in (0, t],\ x \in E. \tag{5.12}$$

This is the backward equation in (FK). The initial condition follows from continuity:

$$u(0, x) = \mathbb{E}_x[f(X_0)] = f(x) \tag{5.13}$$

When more regularity of coefficients is assumed (smooth coefficients, uniform ellipticity), the above probabilistic derivation can be upgraded to show that $u \in C^{1,2}((0, T]\times E)$ and the equality holds classically. Alternatively, one may interpret the equality in the weak sense (distributional or mild sense) under weaker regularity.

5.4 Uniqueness (analytic semigroup argument)

To prove uniqueness in the class of bounded continuous functions, consider the operator $A = \mathcal{L} - V$ acting on $C_b(E)$ (or on $C_0(E)$ in the Feller setting). The semigroup defined by

$$(P_t^V f)(x) := \mathbb{E}_x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big] \tag{5.14}$$

is a strongly continuous contraction semigroup on $C_b(E)$ (or on $C_0(E)$ when appropriate). Its generator is the closure of A on a core (e.g. $C_c^2(E)$). If $u(t, x)$ is any bounded continuous solution of (FK) then $t \mapsto u(t, \cdot)$ satisfies the abstract Cauchy problem

$$\partial_t u(t, \cdot) = Au(t, \cdot), \qquad u(0, \cdot) = f, \tag{5.15}$$

and by the uniqueness of solutions to the abstract Cauchy problem for strongly continuous semigroups we conclude

$$u(t, \cdot) = P_t^V f \tag{5.16}$$

Consequently the probabilistically defined u is the unique bounded continuous solution.

5.5 Remarks and extensions

The proof above treated the bounded-potential case for clarity. There are several standard extensions: if V is locally bounded from below or belongs to a suitable Kato class one may define the Feynman–Kac expectation via monotone limits or via exponential martingale localization and obtain the same representation together with existence/uniqueness in appropriate function spaces. If $\mathcal{L}$ includes jump terms (Lévy or more general integro-differential generators), the multiplicative functional and Dynkin arguments carry through provided the jump integrability conditions hold; the resulting PDE is the nonlocal (integro-differential) parabolic equation with potential. Time-dependent potentials $V(t,x)$ and time-inhomogeneous generators $\mathcal{L}_t$ are handled by replacing the multiplicative functional by $\exp\Big(-\int_0^t V(s,X_s)\,ds\Big)$ and working with the non-autonomous backward equation

$$\partial_t u = \mathcal{L}_t u - V(t,\cdot)u \tag{5.17}$$

The Feynman–Kac formula also yields spectral information: when f is an eigenfunction of $A = \mathcal{L} - V$ the exponential factor appears and long-time asymptotics of $u(t,x)$ are governed by principal eigenvalues of A; this is the probabilistic underpinning of many variational and large-deviations results.

References (selected): Stroock & Varadhan, Dynkin, Ethier & Kurtz, Friedman, Pazy, Davies, and standard texts on stochastic processes and PDEs for the technical regularity and Kato-class extensions.

5.6 Feynman–Kac formula with unbounded potentials (Kato class): rigorous proof

In what follows we work on $E = \mathbb{R}^d$ and consider a (time-homogeneous) diffusion process X with generator

$$\mathcal{L}f(x) = \sum_{i=1}^d b_i(x)\partial_{x_i} f(x) + \tfrac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{x_i}\partial_{x_j} f(x), \tag{5.18}$$

under the standing analytic and probabilistic assumptions listed below. The aim is to prove the Feynman–Kac representation for potentials V belonging to an appropriate Kato class, in particular allowing certain unbounded V. Throughout $(P_t)_{t\geq 0}$ denotes the Markov semigroup of X and $p_t(x,y)$ its transition density (when it exists).

Definition 5.2 (Standing hypotheses on the diffusion)**.** Assume:

1. The coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ (symmetric) are measurable and locally bounded, and $a(x)$ is uniformly elliptic on compacts.

2. The martingale problem for $\mathcal{L}$ is well posed, producing a Markov family $(\mathbb{P}_x)_{x\in\mathbb{R}^d}$ with continuous paths, and the semigroup P_t is Feller on $C_0(\mathbb{R}^d)$.

3. The transition kernel admits a jointly measurable density $p_t(x,y)$ for $t \in (0,T]$ which satisfies Gaussian upper bounds: for some constants $C, c > 0$,

$$p_t(x,y) \le Ct^{-d/2} \exp\Big(-\frac{|x-y|^2}{ct}\Big), \qquad 0 < t \le T,\ x,y \in \mathbb{R}^d. \tag{5.19}$$

Definition 5.3 (Kato class $\mathcal{K}_d$)**.** A measurable function $V : \mathbb{R}^d \to \mathbb{R}$ belongs to the (time–independent) Kato class $\mathcal{K}_d$ on $[0,T]$ if

$$\lim_{t \downarrow 0} \sup_{x \in \mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} p_s(x,y)\,|V(y)|\,dy\,ds \;=\; 0. \tag{5.20}$$

Equivalently, for every $\varepsilon > 0$ there exists $t_\varepsilon > 0$ such that for all $0 < t \le t_\varepsilon$,

$$\sup_{x \in \mathbb{R}^d} \int_0^t P_s(|V|)(x)\,ds < \varepsilon. \tag{5.21}$$

Remark 5.4. This is the probabilistic (heat-kernel) definition of the Kato class; in Euclidean space it is equivalent to standard analytic Kato-type conditions (for example those expressed via Riesz potentials) under Gaussian bounds. The class includes bounded functions and many singular/unbounded potentials arising in Schrödinger operators (e.g. $V(x) = |x|^{-\alpha}$ with $\alpha < 2$ in appropriate dimensions).

We now state the main theorem.

Theorem 5.5 (Feynman–Kac for Kato-class potentials)**.** *Let the standing hypotheses hold and let $V \in \mathcal{K}_d$. Let $f \in C_b(\mathbb{R}^d)$. Define for $t \in [0,T]$, $x \in \mathbb{R}^d$,*

$$u(t,x) \;:=\; \mathbb{E}_x\Big[\exp\Big(-\int_0^t V(X_s)\,ds\Big)\,f(X_t)\Big], \tag{5.22}$$

where the exponential is defined as the (possibly extended) limit via monotone truncation (see proof). Then:

1. *The functional $u(t,x)$ is well-defined, finite for each (t,x), bounded on compact time intervals, and jointly measurable in (t,x).*
2. *For each $t > 0$ the function $u(t,\cdot)$ belongs to $C_b(\mathbb{R}^d)$, and u satisfies the backward equation in the mild (integral) sense*

$$u(t,x) = P_t f(x) - \int_0^t P_{t-s}(V\,u(s,\cdot))(x)\,ds, \tag{5.23}$$

and in particular, if coefficients are smooth enough, $u \in C^{1,2}((0,T] \times \mathbb{R}^d)$ and solves

$$\partial_t u(t,x) = \mathcal{L}u(t,x) - V(x)u(t,x), \qquad u(0,x) = f(x). \tag{5.24}$$

3. *The function u is the unique bounded continuous (mild/classical under regularity) solution of the above initial-value problem.*

Proof. The proof proceeds in several controlled steps: truncation/approximation, uniform estimates using the Kato property, passage to the limit, verification of the mild equation, and uniqueness by semigroup methods.

1. Truncation / definition. For $n \in \mathbb{N}$ define $V_n := V \wedge n$ (pointwise truncation from above) and, if V may be negative, also

$$V_n := \max\{-n, \min\{V, n\}\} \tag{5.25}$$

to truncate both tails; thus each V_n is bounded and measurable and $V_n \to V$ pointwise with $|V_n| \leq |V|$. For each n define the bounded multiplicative functional and corresponding approximant

$$u_n(t,x) := \mathbb{E}_x\Big[\exp\Big(-\int_0^t V_n(X_s)\,ds\Big)f(X_t)\Big]. \tag{5.26}$$

Classical bounded-potential Feynman–Kac (see previous chapter) ensures that u_n is well-defined, bounded and satisfies the mild equation

$$u_n(t,x) = P_t f(x) - \int_0^t P_{t-s}(V_n u_n(s,\cdot))(x)\,ds. \tag{5.27}$$

2. Monotone / dominated convergence and finiteness. Because V_n is not necessarily monotone in n when truncating both tails, we instead consider the sequence $V^{(m)} := V \wedge m$ (upper truncation) and treat the negative part separately if needed; one may handle the negative part by splitting $V = V^+ - V^-$ and approximating each part suitably. For clarity assume first $V \geq 0$. Then $V^{(m)} \uparrow V$ and $u_m(t,x)$ is a *monotone nonincreasing* sequence in m:

$$u_1(t,x) \geq u_2(t,x) \geq \cdots \geq 0. \tag{5.28}$$

Hence the pointwise limit

$$u(t,x) := \lim_{m\to\infty} u_m(t,x) \in [0,\infty] \tag{5.29}$$

exists. We must show $u(t,x) < \infty$. For this, use the Kato condition to obtain small-time bounds on the exponential. Fix $T > 0$. For any $\lambda > 0$ and $t \in (0,T]$ apply the elementary inequality $e^{-y} \leq 1 \wedge (1+y)^{-1}$ (or use $\exp(-y) \leq 1$ plus moment bounds) together with the Duhamel (iteration) representation to derive

$$u_m(t,x) \leq \sum_{k=0}^{\infty} \frac{1}{k!}\mathbb{E}_x\Big[\Big(\int_0^t V^{(m)}(X_s)\,ds\Big)^k |f(X_t)|\Big]. \tag{5.30}$$

Using Fubini and iterated kernels one obtains the series representation

$$u_m(t,x) \leq \|f\|_\infty \sum_{k=0}^{\infty} \frac{1}{k!}\int_{0<s_1<\cdots<s_k<t} \int\cdots\int p_{s_1}(x,y_1)V(y_1)p_{s_2-s_1}(y_1,y_2)V(y_2)\cdots p_{t-s_k}(y_k,z)\,dy_1\cdots dy_k dz. \tag{5.31}$$

Integrating out the final z gives $P_{t-s_k}1(y_k) = 1$. By inductive use of the Kato smallness on short time intervals (choose partition of $[0,T]$ into subintervals where the Kato integral is small) one obtains a majorant geometric series, uniformly in m, showing that for each fixed t the sequence $u_m(t,x)$ is uniformly bounded in m and thus $u(t,x) < \infty$. The detailed bookkeeping follows the standard Dyson series / Neumann-series argument (see [40] for Schrödinger semigroups).

If V has a negative part $V^- \not\equiv 0$, write $V = V^+ - V^-$ and approximate V^+ from below and V^- from above; one then uses Kato-class control for V^+ and the fact that a bounded-from-below additive functional yields nonexplosion. Under the Kato hypothesis both parts can be handled by the same series representation plus Gronwall-type estimates; we omit the routine but lengthier sign-technical details.

3. Uniform estimates and equicontinuity. Using the Gaussian upper bounds for $p_t(x, y)$ and the Kato smallness, one proves that $\{u_m(t, \cdot)\}_m$ is equicontinuous on compacts for each $t > 0$. Indeed the mild equation implies, for x, x' in a compact K,

$$|u_m(t,x)-u_m(t,x')| \le |P_tf(x)-P_tf(x')|+\int_0^t |P_{t-s}(V^{(m)}u_m(s,\cdot))(x)-P_{t-s}(V^{(m)}u_m(s,\cdot))(x')|\,ds. \tag{5.32}$$

Use the heat-kernel continuity and the uniform in m bounds on $V^{(m)}u_m$ (obtained from step 2) plus dominated convergence to make the right-hand side small when $|x - x'|$ is small; details follow from standard semigroup regularity results under Gaussian bounds (see Aronson estimates and parabolic smoothing). Consequently any subsequence has a further subsequence converging locally uniformly; combined with monotone limit this yields that $u_m \to u$ locally uniformly in x for each fixed $t > 0$.

4. Passage to the limit in the mild equation. Each u_m satisfies

$$u_m(t, x) = P_tf(x) - \int_0^t P_{t-s}(V^{(m)}u_m(s, \cdot))(x)\,ds. \tag{5.33}$$

Fix $t > 0$, and let $m \to \infty$. Using the pointwise a.e. convergence $V^{(m)}u_m \to Vu$ and the uniform integrability coming from the Kato-based bounds (step 2), dominated convergence and the strong continuity of P_{t-s} on bounded continuous functions yield

$$u(t, x) = P_tf(x) - \int_0^t P_{t-s}(Vu(s, \cdot))(x)\,ds, \tag{5.34}$$

i.e. the mild (Duhamel) formulation is valid for the limit u.

5. Regularity and verification of the PDE. Under further smoothness and ellipticity assumptions (e.g. coefficients a, b of class C^∞ and uniform ellipticity), standard parabolic regularity theory applied to the mild equation gives that $u \in C^{1,2}((0, T] \times \mathbb{R}^d)$ and satisfies the parabolic PDE

$$\partial_t u(t, x) = \mathcal{L}u(t, x) - V(x)u(t, x), \tag{5.35}$$

pointwise. The justification uses bootstrap regularity for parabolic equations with bounded measurable lower-order terms; one refers to classical PDE texts for the detailed Schauder or L^p-based regularity machinery (cf. Friedman, Ladyzhenskaya–Uraltseva). In the general Kato-class setting one obtains that u is a mild (and distributional) solution; classical regularity is recovered under additional hypotheses.

6. Uniqueness. Uniqueness in the class of bounded continuous mild solutions follows from contraction semigroup arguments. Define the perturbed semigroup P_t^V on $B_b(\mathbb{R}^d)$ by the Dyson series

$$P_t^V f = \sum_{k=0}^{\infty}(-1)^k \int_{0<s_1<\cdots<s_k<t} P_{t-s_k}(VP_{s_k-s_{k-1}}\cdots VP_{s_1}f)\,ds_1\cdots ds_k, \tag{5.36}$$

and use the Kato smallness to show the series converges uniformly for small t and then iterate to cover any finite T. The mild equation implies that any bounded solution must equal $P_t^V f$; details are the standard resolvent/semigroup uniqueness argument. Alternatively, one may employ the positivity preserving property of the semigroup and Gronwall inequalities in Banach spaces to show uniqueness.

This completes the proof of the theorem. □

Remark 5.6. The argument above follows the classical approach used in Schrödinger semigroup theory: (i) approximate the potential by bounded potentials, (ii) obtain uniform control via the Kato condition and Gaussian bounds, (iii) pass to the limit to obtain the Feynman–Kac representation for the limit potential, and (iv) deduce uniqueness from semigroup theory. For precise and exhaustive treatments see Simon *Schrödinger semigroups*, Aizenman–Simon, and Davies' monograph on heat kernels and spectral theory.

5.7 Regularity and boundary conditions

We collect here the precise regularity assumptions on coefficients and data, and the admissible boundary conditions, that are used in the statement and proof of the classical Feynman–Kac representation in Chapter 5. These hypotheses are stated in a form suitable both for probabilistic representations (via diffusion processes, killed or reflected at the boundary) and for analytic parabolic theory (classical, Sobolev and Hölder frameworks).

Let $D \subset \mathbb{R}^d$ be a nonempty domain and fix a finite time horizon $T > 0$. Denote by $\overline{D}$ the closure and by ∂D the boundary. We consider second-order, (possibly) nondivergence-form elliptic operators of the form

$$\mathcal{L}u(x) = \sum_{i,j=1}^{d} a_{ij}(x)\,\partial_{ij}u(x) + \sum_{i=1}^{d} b_i(x)\,\partial_i u(x) + c(x)u(x), \tag{5.37}$$

where $a(x) = (a_{ij}(x))$ is symmetric. The parabolic operator is $\partial_t - \mathcal{L}$. Below ∇ and D^2 denote the spatial gradient and Hessian respectively.

Ellipticity and coefficient regularity. We assume uniform (or at least local) ellipticity and regularity of the diffusion matrix and lower-order coefficients. Precisely, either of the following two archetypal hypotheses is imposed depending on the desired regularity of solutions:

(A) Uniformly elliptic, Hölder coefficients (classical/Hölder theory). There exist constants $0 < \lambda \le \Lambda < \infty$ and $\alpha \in (0,1)$ such that for all $x \in \overline{D}$ and $\xi \in \mathbb{R}^d$,

$$\lambda|\xi|^2 \le \xi^\top a(x)\,\xi \le \Lambda|\xi|^2, \tag{5.38}$$

and $a_{ij}, b_i, c \in C^\alpha(\overline{D})$. Under these assumptions parabolic Schauder theory applies and one expects classical solutions $u \in C^{1+\alpha/2,2+\alpha}((0,T] \times \overline{D})$ to the backward equation

$$\partial_t u(t,x) = \mathcal{L}u(t,x) - V(x)u(t,x), \tag{5.39}$$

provided the data satisfy compatible Hölder conditions (see below).

(B) Measurable coefficients with uniform ellipticity (Sobolev theory). The map $a(x)$ is measurable and satisfies the uniform ellipticity bounds above, while $b \in L^q_{\mathrm{loc}}(D)$, $c \in L^r_{\mathrm{loc}}(D)$ for appropriate exponents q, r (e.g. $q > d$ for certain embedding estimates). In this setting one works in parabolic Sobolev spaces $L^p(0,T; W^{2,p}_{\mathrm{loc}}(D))$ or parabolic Bessel potential spaces; weak (distributional) or mild solutions are then the natural objects.

Potential / Feynman term regularity. The potential V appearing in the Feynman–Kac weight $\exp\Big(-\int_0^t V(X_s)\,ds\Big)$ is assumed to belong to a class ensuring both integrability along diffusion paths and perturbative control of the semigroup. Typical admissible classes are:

$$V \in C_b(\overline{D}) \quad \text{(bounded case)}, \tag{5.40}$$

or the Kato class $V \in \mathcal{K}_d$ (or a localized Kato class on D) so that

$$\lim_{t\downarrow 0} \sup_{x\in D} \int_0^t P_s(|V|)(x)\,ds = 0, \tag{5.41}$$

where P_s is the heat semigroup associated to $\mathcal{L}$ (or its Dirichlet realization). Under the Kato condition one obtains well-defined exponential functionals and a robust perturbation theory for the semigroup.

Initial and boundary data regularity and compatibility. Let $f : D \to \mathbb{R}$ be the terminal datum for the backward problem $u(0,\cdot) = f$ in the probabilistic convention used in Chapter 5. The following are convenient hypotheses:

- For classical/Hölder solutions assume $f \in C^{2+\alpha}(\overline{D})$ (or $f \in C^{2+\alpha}(D)$ with suitable control near the boundary) so that the initial trace is compatible with spatial regularity.
- For Sobolev/mild solutions assume $f \in L^p(D)$ or $f \in W^{2-2/p,p}(D)$ depending on the L^p-parabolic theory employed.

Compatibility conditions at $t = 0$ and on ∂D are required for higher regularity. For example, under (A) if a Dirichlet boundary condition $u(t,x) = g(t,x)$ on $(0,T] \times \partial D$ is prescribed with $g \in C^{1+\alpha/2,2+\alpha}([0,T] \times \partial D)$, then the compatibility condition of order zero is

$$f|_{\partial D} = g(0,\cdot), \tag{5.42}$$

and the order-one compatibility condition reads

$$\partial_t g(0,\cdot) = \mathcal{L}f - Vf \quad \text{on } \partial D, \tag{5.43}$$

interpreted in the appropriate trace sense. These compatibility relations ensure that the parabolic boundary value problem admits a classical solution in the Hölder class.

Admissible boundary conditions and probabilistic interpretations. The main boundary conditions considered and their probabilistic correspondents are:

Dirichlet (absorbing) boundary condition: prescribe

$$u(t,x) = 0, \qquad (t,x) \in (0,T] \times \partial D. \tag{5.44}$$

Probabilistically this corresponds to the diffusion X being *killed* at the first exit time $\tau_D = \inf\{t > 0 : X_t \notin D\}$. The Feynman–Kac representation for the Dirichlet problem uses the killed semigroup

$$u(t,x) = \mathbb{E}_x\Big[e^{-\int_0^{t\wedge\tau_D} V(X_s)\,ds}\, f(X_t)\, \mathbf{1}_{\{t<\tau_D\}}\Big], \tag{5.45}$$

or equivalently replacing $f(X_t)\mathbf{1}_{\{t<\tau_D\}}$ by $f(X_{t\wedge\tau_D})$ with f extended by 0 outside D. Regularity of ∂D (e.g. $C^{2+\alpha}$ boundary) is required for classical boundary regularity results and for continuity up to the boundary of the solution.

Neumann (reflecting) boundary condition: prescribe for $n(x)$ the outward normal,

$$\partial_n u(t,x) := \nabla u(t,x) \cdot n(x) = 0, \qquad (t,x) \in (0,T] \times \partial D. \tag{5.46}$$

This corresponds probabilistically to a reflecting diffusion, obtained by constructing a strong solution to an SDE with reflection in the inward normal direction (Skorokhod problem). Existence requires smoothness of the boundary (at least C^2) and appropriate structure of the diffusion matrix near ∂D. The reflecting process yields a Feynman–Kac representation with local-time boundary terms absent in the representation of the Dirichlet problem.

Robin and mixed boundary conditions: prescribe

$$\partial_n u + \beta u = 0 \quad \text{on } \partial D, \tag{5.47}$$

with β bounded on ∂D. Probabilistically this corresponds to partial absorption/partial reflection and can be realized via killing with boundary local time and exponential weighting; technical construction uses excursion theory or penalization limits.

Boundary regularity for probabilistic representations. For the Feynman–Kac formula to produce a function u continuous up to the boundary and satisfying the boundary condition pointwise, the following geometric/analytic hypotheses are standard:

$$D \text{ has } C^{2+\alpha} \text{ boundary and } a, b, c, V \in C^\alpha(\overline{D}), \tag{5.48}$$

or, in the weak setting,

$$\partial D \text{ is Lipschitz and coefficients are measurable, with appropriate trace regularity for } f, g. \tag{5.49}$$

Under the $C^{2+\alpha}$ hypothesis parabolic Schauder estimates yield $u \in C^{1+\alpha/2,2+\alpha}([0,T] \times \overline{D})$ and the boundary condition is satisfied classically. In the Lipschitz/Sobolev framework one obtains $u \in L^p(0,T;W^{2,p}(D)) \cap W^{1,p}(0,T;L^p(D))$ and boundary conditions hold in the trace sense.

Generator domain and semigroup realization. Analytically, boundary conditions are encoded in the domain of the generator A of the semigroup. For example, the Dirichlet realization

A_D of $\mathcal{L} - V$ on $L^2(D)$ has domain

$$\mathcal{D}(A_D) = \{u \in H^2(D) \cap H_0^1(D) : \ \mathcal{L}u - Vu \in L^2(D)\}, \tag{5.50}$$

while the Neumann realization A_N uses

$$\mathcal{D}(A_N) = \{u \in H^2(D) : \ \partial_n u = 0 \text{ on } \partial D, \ \mathcal{L}u - Vu \in L^2(D)\}. \tag{5.51}$$

These operator-theoretic domains are the correct ones for spectral analysis and long-time asymptotics; they also determine the probabilistic process (killed, reflected, or partially killed) associated to the Feynman–Kac semigroup.

Compatibility and maximum principles. When c or V has a sign, maximum principles and positivity preserving properties are available and are crucial in uniqueness arguments. For bounded V the semigroup P_t^V is positivity preserving and contractive on L^∞. For unbounded V in the Kato class these properties survive on finite time intervals by perturbation theory. Compatibility of initial and boundary data ensures no instantaneous incompatibilities: for Dirichlet problems $f|_{\partial D} = g(0, \cdot)$ must hold, while for Neumann problems the normal derivative of the initial data must be well defined in a trace sense.

Summary of admissible regularity/boundary regimes. In summary, the principal regimes used in Chapter 5 are:

1. *Classical regime:* $a, b, c, V \in C^\alpha(\overline{D})$, uniform ellipticity, $D \in C^{2+\alpha}$, $f \in C^{2+\alpha}(\overline{D})$, boundary data in $C^{1+\alpha/2,2+\alpha}$. Yields $u \in C^{1+\alpha/2,2+\alpha}([0,T] \times \overline{D})$ and pointwise boundary conditions.

2. *Sobolev/mild regime:* measurable a with uniform ellipticity, b, c, V in suitable L^p or Kato class, D Lipschitz. Yields u as a mild/weak solution in L^p-parabolic spaces and boundary conditions in trace or weak sense; Feynman–Kac remains valid via killed/penalized diffusion constructions.

3. *Probabilistic regime:* minimal assumptions for well-defined diffusion and exponential functional (e.g. martingale problem well posed, V Kato-class). Yields the Feynman–Kac representation as the defining object for the semigroup and characterizes boundary behavior through hitting probabilities and exit laws.

These precise regularity and boundary assumptions are used throughout Chapter 5 to justify the interchange of expectation and differentiation, the use of Itô's formula up to the boundary (via localization and stopping at exit times), the application of parabolic regularity theorems, and the identification of semigroup generators with operator realizations under the chosen boundary condition.

5.8 Conclusion

In this chapter, we established the classical Feynman–Kac formula in a mathematically rigorous framework, bridging stochastic analysis, semigroup theory, and the theory of parabolic partial differential equations. Beginning with the infinitesimal generator $A = \frac{1}{2}\Delta_g$ of Brownian motion on a Riemannian manifold (M, g), we incorporated a measurable potential $V : M \to \mathbb{R}$ to define the Schrödinger-type operator $H = -\frac{1}{2}\Delta_g + V$. We then showed, with full analytical justification, that the semigroup e^{-tH} admits the probabilistic representation

$$(e^{-tH} f)(x) = \mathbb{E}^x \left[\exp\left(- \int_0^t V(X_s)\, ds \right) f(X_t) \right], \tag{5.52}$$

where $(X_t)_{t \geq 0}$ denotes Brownian motion on M.

The proof rigorously connected three perspectives:

1. The analytic one, via the semigroup generated by H;
2. The probabilistic one, via the expectation over Brownian paths;
3. The geometric one, through the Laplace–Beltrami operator Δ_g.

We demonstrated that the Feynman–Kac formula is not merely a computational tool, but a structural identity encoding how curvature, diffusion, and potential energy jointly determine the evolution of heat-like systems. Its rigorous justification required careful use of Itô's formula, martingale properties, and the stochastic representation of the heat semigroup. Thus, this chapter provided the analytic and probabilistic synthesis upon which all subsequent generalizations—such as Feynman–Kac semigroups, Dirichlet forms, and geometric extensions—are built.

References for this chapter: Simon (1979) [19] and Simon (2005) [20], Reed and Simon (1980) [21], Fukushima (1980) [22].

Chapter 6

The Feynman–Kac Semigroup

> "The bridge between stochastic analysis and partial differential equations is not merely a correspondence — it is a profound equivalence of dynamical principles."
>
> — *Kiyosi Itô*

In the preceding chapter, we rigorously established the Feynman–Kac formula as a representation of solutions to parabolic partial differential equations (PDEs) in terms of stochastic processes. The present chapter develops this representation into a semigroup framework, revealing deep connections between the probabilistic evolution of diffusion processes and the analytic theory of Schrödinger-type operators. The resulting semigroup, known as the *Feynman–Kac semigroup*, plays a central role in both mathematical physics and the theory of Markov processes.

Let $(X_t)_{t\geq 0}$ be a diffusion process on $\mathbb{R}^d$ governed by the stochastic differential equation

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t, \tag{6.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy the standard Lipschitz and growth conditions ensuring strong existence and uniqueness of solutions. The corresponding infinitesimal generator acts on $C_b^2(\mathbb{R}^d)$ as

$$(\mathcal{L}f)(x) = \frac{1}{2}\operatorname{Tr}\big(a(x)D^2 f(x)\big) + \langle b(x), \nabla f(x)\rangle, \tag{6.2}$$

where $a(x) = \sigma(x)\sigma(x)^\top$.

Let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable potential function. For each $t \geq 0$, we define the operator

$$(P_t^V f)(x) := \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right], \tag{6.3}$$

whenever the expectation exists and is finite. The family $(P_t^V)_{t\geq 0}$ is called the *Feynman–Kac semigroup* associated with the generator $\mathcal{L}$ and potential V.

6.1 Definition of the Feynman–Kac semigroup

For $f \in B_b(\mathbb{R}^d)$, the operator P_t^V is well-defined if

$$\mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds}|f(X_t)|\right] < \infty. \tag{6.4}$$

The mapping $(t, f) \mapsto P_t^V f$ satisfies:

1. $P_0^V f = f$;
2. $P_t^V P_s^V f = P_{t+s}^V f$ (the *semigroup property*);
3. P_t^V is a contraction on L^∞ if $V \geq 0$;
4. $\|P_t^V f\|_\infty \leq e^{t\|V^-\|_\infty}\|f\|_\infty$, where $V^- := \max\{-V, 0\}$.

For the remainder we work with a time-homogeneous Markov process $X = (X_t)_{t\geq 0}$ on $\mathbb{R}^d$ whose transition probabilities $(P_t)_{t\geq 0}$ are well defined on a chosen function space (for example $B_b(\mathbb{R}^d)$, $C_b(\mathbb{R}^d)$, $C_0(\mathbb{R}^d)$ or $L^p(\mathbb{R}^d)$); all assertions below are to be interpreted relative to this underlying semigroup. Let $V : \mathbb{R}^d \to \mathbb{R}$ be measurable. For each Borel measurable $f : \mathbb{R}^d \to \mathbb{R}$ we define the Feynman–Kac functional (when it is finite) by

$$(P_t^V f)(x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \qquad x \in \mathbb{R}^d,\ t \geq 0, \tag{6.5}$$

where $\mathbb{E}^x$ denotes expectation under the law of X started at x. The first basic question is: under what hypotheses is the right-hand side well defined and finite? A minimal sufficient condition is the following integrability condition: for fixed $t > 0$ and $f \in B_b(\mathbb{R}^d)$ we require

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}|f(X_t)|\Big] < \infty. \tag{6.6}$$

If f is bounded this reduces to finiteness of the exponential moment

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\Big] < \infty. \tag{6.7}$$

Sufficient criteria ensuring these bounds include but are not limited to) the following: V bounded below then

$$e^{-\int_0^t V} \leq e^{t\|V^-\|_\infty} \tag{6.8}$$

and the expectation is finite, V bounded, or V belonging to a Kato class relative to (P_t) which guarantees small-time integrability of $\int_0^t V(X_s)\,ds$ and hence finiteness of the exponential via series/Dyson expansions. When such integrability fails for some t, one can still define $P_t^V f$ by truncation $V \wedge n$ and monotone/ dominated convergence arguments on the interval of finiteness; this produces an extended-semigroup defined up to an explosion time of the exponential functional.

Assume henceforth that for the chosen class of test functions (for example $B_b(\mathbb{R}^d)$) and for each $t \in [0, T]$ the expectation in the definition is finite for all x. The map $x \mapsto (P_t^V f)(x)$ is

Borel measurable as a composition of measurable maps: measurability follows from the measurable dependence of the law $\mathbb{P}^x$ on x together with Fubini–Tonelli applied to a product measure representing expectation, or more concretely from the representation

$$(P_t^V f)(x) = \int_{C([0,t];\mathbb{R}^d)} e^{-\int_0^t V(\omega(s))\,ds} f(\omega(t))\,\Pi_x(d\omega), \tag{6.9}$$

where Π_x is the law of the path started at x; continuity (when it holds) is treated separately below.

We now prove the algebraic-semigroup properties and the basic operator bounds listed in the paragraph.

1. $P_0^V f = f$. This is immediate from the definition since $\int_0^0 V(X_s)\,ds = 0$ and

$$\mathbb{P}^x(X_0 = x) = 1 \tag{6.10}$$

hence

$$(P_0^V f)(x) = \mathbb{E}^x[f(X_0)] = f(x). \tag{6.11}$$

2. Semigroup property $P_t^V P_s^V f = P_{t+s}^V f$. Fix $s,t \geq 0$ and $x \in \mathbb{R}^d$. Using the Markov property of X and the multiplicative factorization of the exponential, we compute

$$\begin{aligned}
(P_t^V P_s^V f)(x) &= \mathbb{E}^x\Big[e^{-\int_0^t V(X_r)\,dr}(P_s^V f)(X_t)\Big] \\
&= \mathbb{E}^x\Big[e^{-\int_0^t V(X_r)\,dr}\,\mathbb{E}^{X_t}\Big[e^{-\int_0^s V(X_u)\,du} f(X_s)\Big]\Big] \\
&= \mathbb{E}^x\Big[e^{-\int_0^t V(X_r)\,dr}\,\mathbb{E}\big[e^{-\int_t^{t+s} V(X_u)\,du} f(X_{t+s}) \mid \mathcal{F}_t\big]\Big] \\
&= \mathbb{E}^x\Big[e^{-\int_0^{t+s} V(X_r)\,dr} f(X_{t+s})\Big] = (P_{t+s}^V f)(x),
\end{aligned} \tag{6.12}$$

where $\mathcal{F}_t$ is the natural filtration and we used the time-shifted Markov property

$$\mathbb{E}[\,\cdot \mid \mathcal{F}_t\,] = \mathbb{E}^{X_t}[\,\cdot\,] \tag{6.13}$$

Justification of the interchange of expectations and conditioning uses Fubini/Tonelli and the integrability hypothesis.

3. Contraction property when $V \geq 0$. If $V \geq 0$ then

$$e^{-\int_0^t V(X_s)\,ds} \leq 1 \tag{6.14}$$

almost surely, hence for bounded f

$$\|P_t^V f\|_\infty = \sup_{x\in\mathbb{R}^d} \big|\mathbb{E}^x[e^{-\int_0^t V(X_s)\,ds} f(X_t)]\big| \leq \|f\|_\infty. \tag{6.15}$$

Thus P_t^V is a contraction on L^∞ and preserves positivity. The same conclusion extends to C_b and C_0 when the semigroup maps those spaces into themselves.

4. Uniform exponential bound with negative part. Write $V = V^+ - V^-$ with $V^+ :=$

$\max\{V, 0\}$ and $V^- := \max\{-V, 0\}$. Then

$$e^{-\int_0^t V(X_s)\,ds} = e^{-\int_0^t V^+(X_s)\,ds} e^{\int_0^t V^-(X_s)\,ds} \le e^{\int_0^t V^-(X_s)\,ds}. \tag{6.16}$$

If V^- is bounded so that $\|V^-\|_\infty < \infty$, then

$$\int_0^t V^-(X_s)\,ds \le t\|V^-\|_\infty \tag{6.17}$$

and consequently

$$\|P_t^V f\|_\infty \le e^{t\|V^-\|_\infty}\|f\|_\infty. \tag{6.18}$$

This provides the uniform operator-norm bound

$$\|P_t^V\|_{L^\infty \to L^\infty} \le e^{t\|V^-\|_\infty} \tag{6.19}$$

When V^- is unbounded, one may replace the uniform bound by pointwise (in x) estimates using exponential moments if available, or restrict to time intervals on which the exponential moments are finite.

5. Measurability, positivity and monotonicity. For measurable $f \ge 0$ the expectation defining $P_t^V f$ is nonnegative and monotone in f. If V is replaced by a larger potential $\widetilde{V} \ge V$ pointwise, then

$$e^{-\int_0^t \widetilde{V}} \le e^{-\int_0^t V} \tag{6.20}$$

and hence

$$P_t^{\widetilde{V}} f \le P_t^V f \tag{6.21}$$

for $f \ge 0$. These monotonicity properties are useful in approximation arguments (e.g. truncation of potentials).

6. Domains and actions on function spaces. If P_t (the underlying Markov semigroup) maps C_b into itself and if $V \in C_b$, then P_t^V maps C_b to C_b and is strongly continuous on C_b. If P_t admits a transition density $p_t(x, y)$ and V satisfies suitable integrability conditions (e.g. Kato-class), then one can often represent P_t^V by an integral kernel $p_t^V(x, y)$ with

$$(P_t^V f)(x) = \int_{\mathbb{R}^d} p_t^V(x, y) f(y)\,dy, \tag{6.22}$$

and deduce mapping properties on L^p-spaces and pointwise bounds via kernel estimates.

7. Strong continuity at zero (generator identification). Suppose V is bounded and continuous and P_t is strongly continuous on C_b. Then for $f \in C_b^2$ one may use Itô's formula to show

$$\lim_{t \downarrow 0} \frac{P_t^V f - f}{t} = (\mathcal{L} - V)f \quad \text{in } \|\cdot\|_\infty, \tag{6.23}$$

so that $\mathcal{L} - V$ is (a core of) the generator of the strongly continuous semigroup $(P_t^V)_{t \ge 0}$ on C_b. The derivation uses the martingale decomposition

$$e^{-\int_0^t V(X_s)\,ds} f(X_t) - f(X_0) - \int_0^t e^{-\int_0^s V(X_r)\,dr}(\mathcal{L} - V)f(X_s)\,ds \tag{6.24}$$

which is a martingale; taking expectations, dividing by t, and letting $t \downarrow 0$ yields the generator identification under dominated convergence.

8. Composition/adjoint considerations. If P_t is symmetric on $L^2(\mu)$ for a reference measure μ and V is μ-measurable with sufficient integrability, then P_t^V is symmetric with respect to the weighted measure $e^{-F}\,d\mu$ only in special cases; more generally one studies the Schrödinger operator $\mathcal{L} - V$ as an unbounded operator on $L^2(\mu)$ and regards P_t^V as its semigroup provided self-adjointness/ sectoriality conditions hold.

In summary, under the stated integrability hypotheses the map $(t, f) \mapsto P_t^V f$ is a well-defined family of linear operators with the semigroup property, positivity and monotonicity, and with operator-norm control governed by the negative part V^-. Additional regularity of V and of the underlying Markov process yields stronger mapping and continuity properties (strong continuity, kernel representations, generator characterization) which are established by standard dominated-convergence and Itô-based arguments as indicated above.

6.2 Strong continuity

Theorem 6.1 (Strong continuity of the Feynman–Kac semigroup)**.** *Suppose $V \in C_b(\mathbb{R}^d)$. Then the Feynman–Kac semigroup $(P_t^V)_{t\geq 0}$ is strongly continuous on $C_b(\mathbb{R}^d)$; that is, for every $f \in C_b(\mathbb{R}^d)$,*

$$\lim_{t\downarrow 0} \|P_t^V f - f\|_\infty = 0. \tag{6.25}$$

Proof. Assume the underlying Markov semigroup $(P_t)_{t\geq 0}$ of the process X is a Feller semigroup and hence is strongly continuous on $C_b(\mathbb{R}^d)$; that is, for every $g \in C_b(\mathbb{R}^d)$,

$$\lim_{t\downarrow 0} \|P_t g - g\|_\infty = 0. \tag{6.26}$$

Let $V \in C_b(\mathbb{R}^d)$ (so that $\|V\|_\infty < \infty$) and fix $f \in C_b(\mathbb{R}^d)$. For $t \geq 0$ and $x \in \mathbb{R}^d$, write

$$(P_t^V f)(x) - f(x) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t) - f(x)\Big]. \tag{6.27}$$

Add and subtract $\mathbb{E}^x[f(X_t)]$ to split the difference:

$$\begin{aligned}(P_t^V f)(x) - f(x) &= \mathbb{E}^x[f(X_t) - f(x)] + \mathbb{E}^x\Big[\big(e^{-\int_0^t V(X_s)\,ds} - 1\big) f(X_t)\Big] \\ &= (P_t f)(x) - f(x) + R_t(x),\end{aligned} \tag{6.28}$$

where we define

$$R_t(x) := \mathbb{E}^x\Big[\big(e^{-\int_0^t V(X_s)\,ds} - 1\big) f(X_t)\Big]. \tag{6.29}$$

Taking the sup norm and using the triangle inequality gives

$$\|P_t^V f - f\|_\infty \leq \|P_t f - f\|_\infty + \sup_{x\in\mathbb{R}^d} |R_t(x)|. \tag{6.30}$$

By strong continuity of the underlying Feller semigroup, the first term satisfies

$$\lim_{t\downarrow 0} \|P_t f - f\|_\infty = 0. \tag{6.31}$$

It remains to show $\sup_x |R_t(x)| \to 0$ as $t \downarrow 0$. Using the elementary bound $|e^{-a} - 1| \leq |a| e^{|a|}$ for all $a \in \mathbb{R}$, and the fact that $|\int_0^t V(X_s)\,ds| \leq t\|V\|_\infty$ a.s., we have

$$\begin{aligned} |R_t(x)| &\leq \mathbb{E}^x\Big[\big|e^{-\int_0^t V(X_s)\,ds} - 1\big|\,|f(X_t)|\Big] \\ &\leq \|f\|_\infty\,\mathbb{E}^x\Big[\big|\int_0^t V(X_s)\,ds\big| e^{\left|\int_0^t V(X_s)\,ds\right|}\Big] \\ &\leq \|f\|_\infty\, e^{t\|V\|_\infty}\,\mathbb{E}^x\Big[\big|\int_0^t V(X_s)\,ds\big|\Big] \\ &\leq \|f\|_\infty\, e^{t\|V\|_\infty} \int_0^t \mathbb{E}^x[|V(X_s)|]\,ds. \end{aligned} \tag{6.32}$$

Using

$$\mathbb{E}^x[|V(X_s)|] = (P_s|V|)(x) \tag{6.33}$$

and the contraction property $\|P_s|V|\|_\infty \leq \|V\|_\infty$, we obtain

$$\sup_{x\in\mathbb{R}^d} |R_t(x)| \leq \|f\|_\infty\, e^{t\|V\|_\infty}\, t\|V\|_\infty \xrightarrow[t\downarrow 0]{} 0. \tag{6.34}$$

Combining the vanishing of $\|P_t f - f\|_\infty$ and $\sup_x |R_t(x)|$ yields

$$\lim_{t\downarrow 0} \|P_t^V f - f\|_\infty = 0, \tag{6.35}$$

for every $f \in C_b(\mathbb{R}^d)$. Hence, $(P_t^V)_{t\geq 0}$ is strongly continuous on $C_b(\mathbb{R}^d)$. □

6.3 Infinitesimal generator of the Feynman–Kac semigroup

Let $(X_t)_{t\geq 0}$ be a Markov process with generator $\mathcal{L}$, and let $V \in C_b(\mathbb{R}^d)$. The Feynman–Kac semigroup $(P_t^V)_{t\geq 0}$ is defined for $f \in C_b(\mathbb{R}^d)$ as

$$(P_t^V f)(x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big]. \tag{6.36}$$

To identify its infinitesimal generator, consider the formal difference quotient

$$\frac{P_t^V f(x) - f(x)}{t} = \frac{\mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t) - f(x)\Big]}{t}. \tag{6.37}$$

Let us introduce the operator

$$\mathcal{A} := \mathcal{L} - V, \tag{6.38}$$

so that our goal is to rigorously show

$$\lim_{t\downarrow 0} \frac{P_t^V f(x) - f(x)}{t} = (\mathcal{A}f)(x) = (\mathcal{L}f)(x) - V(x)f(x). \tag{6.39}$$

To this end, apply Itô's formula to $f(X_t)$, where X_t satisfies the SDE corresponding to $\mathcal{L}$:

$$df(X_t) = (\mathcal{L}f)(X_t)\,dt + \nabla f(X_t)\sigma(X_t)\,dB_t. \tag{6.40}$$

Multiplying both sides by the integrating factor $e^{-\int_0^t V(X_s)\,ds}$ and using Itô's product rule for semimartingales, we obtain

$$d\Big(e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big) = e^{-\int_0^t V(X_s)\,ds}(\mathcal{L}f)(X_t)\,dt - V(X_t)e^{-\int_0^t V(X_s)\,ds} f(X_t)\,dt$$
$$+e^{-\int_0^t V(X_s)\,ds}\nabla f(X_t)\sigma(X_t)\,dB_t. \tag{6.41}$$

Defining

$$M_t := \int_0^t e^{-\int_0^s V(X_r)\,dr}\nabla f(X_s)\sigma(X_s)\,dB_s, \tag{6.42}$$

which is a martingale due to the Itô isometry and boundedness of f and V, we integrate both sides from 0 to t to obtain the stochastic integral representation

$$e^{-\int_0^t V(X_s)\,ds} f(X_t) = f(X_0) + \int_0^t e^{-\int_0^s V(X_r)\,dr}(\mathcal{A}f)(X_s)\,ds + M_t. \tag{6.43}$$

Taking expectations yields

$$\mathbb{E}^x\big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\big] = f(x) + \mathbb{E}^x\Big[\int_0^t e^{-\int_0^s V(X_r)\,dr}(\mathcal{A}f)(X_s)\,ds\Big], \tag{6.44}$$

since $\mathbb{E}^x[M_t] = 0$. Dividing by t and taking the limit $t \downarrow 0$, continuity of f, V, and $\mathcal{L}f$ implies

$$\lim_{t\downarrow 0} \frac{(P_t^V f)(x) - f(x)}{t} = \lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}^x\Big[\int_0^t e^{-\int_0^s V(X_r)\,dr}(\mathcal{A}f)(X_s)\,ds\Big] = (\mathcal{A}f)(x), \tag{6.45}$$

because

$$\frac{1}{t}\int_0^t e^{-\int_0^s V(X_r)\,dr}(\mathcal{A}f)(X_s)\,ds \xrightarrow[t\downarrow 0]{\text{a.s.}} (\mathcal{A}f)(x), \tag{6.46}$$

and the bounded convergence theorem justifies passing the expectation through the limit. This completes the rigorous identification of $\mathcal{A}$ as the infinitesimal generator of $(P_t^V)_{t\geq 0}$. $\square$

6.4 Semigroup representation of parabolic PDEs

Let $(X_t)_{t\geq 0}$ be a Markov process with infinitesimal generator $\mathcal{L}$ and let $V \in C_b(\mathbb{R}^d)$. For a bounded continuous function $f \in C_b(\mathbb{R}^d)$, define

$$u(t,x) := (P_t^V f)(x) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \qquad t \geq 0,\ x \in \mathbb{R}^d, \tag{6.47}$$

where $(P_t^V)_{t\geq 0}$ is the Feynman–Kac semigroup associated with the potential V.

By the strong continuity of $(P_t^V)_{t\geq 0}$ on $C_b(\mathbb{R}^d)$ and the identification of its infinitesimal generator $\mathcal{A} = \mathcal{L} - V$, we have

$$\lim_{h\downarrow 0} \frac{u(t+h,x) - u(t,x)}{h} = \lim_{h\downarrow 0} \frac{(P_h^V u(t,\cdot))(x) - u(t,x)}{h} = (\mathcal{A}u(t,\cdot))(x) = (\mathcal{L}u(t,\cdot))(x) - V(x)u(t,x). \tag{6.48}$$

Thus, for all $t > 0$ and $x \in \mathbb{R}^d$,

$$\frac{\partial u}{\partial t}(t,x) = (\mathcal{L} - V)u(t,x). \tag{6.49}$$

At $t = 0$, the semigroup property $P_0^V = \mathrm{Id}$ ensures that the initial condition is satisfied:

$$u(0,x) = (P_0^V f)(x) = f(x). \tag{6.50}$$

Combining these results, we conclude that u is a classical solution of the Cauchy problem

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = (\mathcal{L} - V)u(t,x), & t > 0, \\ u(0,x) = f(x), & x \in \mathbb{R}^d. \end{cases} \tag{6.51}$$

Moreover, if $\mathcal{L}$ generates a Feller semigroup, u is continuous in both variables (t,x) and satisfies the mild formulation

$$u(t,x) = f(x) + \int_0^t (\mathcal{L} - V)u(s,x)\,ds, \qquad t \geq 0, \tag{6.52}$$

so that the Feynman–Kac semigroup provides the unique mild solution of the parabolic PDE associated with the operator $\mathcal{L} - V$. This establishes a rigorous semigroup representation of parabolic PDEs via stochastic processes.

We shall now discuss the Semigroup representation of parabolic PDEs with Kato-class potentials. Let $(X_t)_{t\geq 0}$ be a Markov process on $\mathbb{R}^d$ with generator $\mathcal{L}$, and let $V : \mathbb{R}^d \to \mathbb{R}$ belong to the Kato class K_d, i.e.,

$$\lim_{t\downarrow 0} \sup_{x\in\mathbb{R}^d} \mathbb{E}^x\Big[\int_0^t |V(X_s)|\,ds\Big] = 0. \tag{6.53}$$

For a measurable function f satisfying suitable growth conditions, define the Feynman–Kac semigroup by

$$u(t,x) := (P_t^V f)(x) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big]. \tag{6.54}$$

Existence and integrability. For $V \in K_d$, the exponential functional

$$\mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}|f(X_t)|\Big] < \infty \tag{6.55}$$

uniformly in x for small t, due to the uniform smallness of $\int_0^t |V(X_s)|ds$ and standard estimates from Khasminskii's lemma. This ensures that $(P_t^V f)(x)$ is well-defined.

Infinitesimal generator. Consider the difference quotient

$$\frac{(P_h^V u)(t,x) - u(t,x)}{h} = \frac{\mathbb{E}^x\big[e^{-\int_0^h V(X_s)ds}u(t,X_h) - u(t,x)\big]}{h}. \tag{6.56}$$

Using Itô's formula for $u(t,X_h)$ and the stochastic Fubini theorem, we obtain the expansion

$$e^{-\int_0^h V(X_s)ds}u(t,X_h) - u(t,x) = \int_0^h e^{-\int_0^s V(X_r)dr}(\mathcal{L} - V)u(t,X_s)ds + M_h, \tag{6.57}$$

where M_h is a local martingale with zero expectation. Dividing by h and taking the limit $h \downarrow 0$, uniform integrability arguments and the Kato-class condition imply

$$\lim_{h\downarrow 0} \frac{(P_h^V u)(t,x) - u(t,x)}{h} = (\mathcal{L} - V)u(t,x) \quad \text{for all } x \in \mathbb{R}^d. \tag{6.58}$$

Cauchy problem and uniqueness. The function u therefore satisfies the Cauchy problem

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = (\mathcal{L} - V)u(t,x), & t > 0, \\ u(0,x) = f(x), & x \in \mathbb{R}^d. \end{cases} \tag{6.59}$$

The Kato-class condition ensures that V is form-bounded with respect to $\mathcal{L}$, which allows standard semigroup theory (e.g., via the Trotter–Kato approximation) to guarantee existence and uniqueness of a mild solution in the Banach space $C_b(\mathbb{R}^d)$ or $L^p(\mathbb{R}^d)$, depending on the growth of f.

Regularity. Under further smoothness of $\mathcal{L}$ (e.g., elliptic diffusions with smooth coefficients), $u(t,x)$ inherits spatial C^2-regularity for $t > 0$ and is continuous in t. Boundary conditions are handled either via the domain of $\mathcal{L}$ (Dirichlet, Neumann, Robin) or by considering the killed process in the case of a bounded domain.

Hence, even for unbounded Kato-class potentials, the Feynman–Kac semigroup provides the unique mild solution of the parabolic PDE, giving a probabilistic representation of the solution:

$$u(t,x) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \tag{6.60}$$

with all integrability, continuity, and generator conditions rigorously justified.

6.5 Approximation procedure for unbounded Kato-class potentials

Let $V \in K_d$ be an unbounded potential and $(X_t)_{t\geq 0}$ a Markov process with generator $\mathcal{L}$. The Feynman–Kac semigroup for unbounded V is defined via a limiting procedure using bounded approximations.

Step 1: Bounded truncation. For $n \in \mathbb{N}$, define the truncated potential

$$V_n(x) := \max\{-n, \min\{V(x), n\}\}, \qquad x \in \mathbb{R}^d. \tag{6.61}$$

Clearly, $V_n \in C_b(\mathbb{R}^d)$ and $V_n \to V$ pointwise as $n \to \infty$.

Step 2: Feynman–Kac semigroup for truncated potentials. For each n, the semigroup

$$(P_t^{V_n} f)(x) := \mathbb{E}^x\Big[e^{-\int_0^t V_n(X_s)\,ds} f(X_t)\Big], \qquad f \in C_b(\mathbb{R}^d), \tag{6.62}$$

is strongly continuous and bounded, with infinitesimal generator $\mathcal{A}_n := \mathcal{L} - V_n$.

Step 3: Uniform integrability via Kato-class condition. Since $V \in K_d$, for each $x \in \mathbb{R}^d$,

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}^x\Big[\int_0^t |V(X_s) - V_n(X_s)|\, ds\Big] \to 0 \quad \text{as } n \to \infty. \tag{6.63}$$

By Khasminskii's lemma, the exponential functional

$$\exp\Big(-\int_0^t V_n(X_s)ds\Big) \quad \text{is uniformly integrable in } n. \tag{6.64}$$

Hence, dominated convergence applies:

$$\lim_{n\to\infty} P_t^{V_n} f(x) = P_t^V f(x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)ds} f(X_t)\Big]. \tag{6.65}$$

Step 4: Generator convergence. Let $\mathcal{A} := \mathcal{L} - V$ on the domain

$$\mathcal{D}(\mathcal{A}) := \Big\{f \in C_b^2(\mathbb{R}^d) : \mathcal{L}f - Vf \in C_b(\mathbb{R}^d)\Big\}. \tag{6.66}$$

For $f \in \mathcal{D}(\mathcal{A})$, using Itô's formula and dominated convergence:

$$\lim_{n\to\infty} \frac{P_h^{V_n} f(x) - f(x)}{h} = (\mathcal{L} - V)f(x) = (\mathcal{A}f)(x), \quad h \downarrow 0. \tag{6.67}$$

Thus, $(P_t^{V_n})_{t\geq 0}$ converges strongly to $(P_t^V)_{t\geq 0}$ and $\mathcal{A}$ is the infinitesimal generator of $(P_t^V)_{t\geq 0}$.

Step 5: Mild solution and uniqueness. For $f \in C_b(\mathbb{R}^d)$, define

$$u_n(t,x) := (P_t^{V_n} f)(x), \quad u(t,x) := \lim_{n\to\infty} u_n(t,x). \tag{6.68}$$

Each u_n satisfies

$$\frac{\partial u_n}{\partial t}(t,x) = (\mathcal{L} - V_n)u_n(t,x), \quad u_n(0,x) = f(x). \tag{6.69}$$

Passing to the limit $n \to \infty$ using dominated convergence and uniform integrability, we obtain the mild solution

$$u(t,x) = f(x) + \int_0^t (\mathcal{L} - V)u(s,x)\, ds, \quad t \geq 0, \tag{6.70}$$

which is continuous in t and bounded if f is bounded. This establishes the unique mild solution of the parabolic PDE with unbounded Kato-class potential V.

Step 6: Remarks on spatial regularity. If $\mathcal{L}$ is a uniformly elliptic diffusion operator with smooth coefficients, then $u(t,x)$ is C^2 in x for $t > 0$, and classical parabolic PDE theory guarantees that u solves the PDE pointwise. Boundary conditions can be incorporated by considering the killed process X_t on a domain or by extending the domain of $\mathcal{A}$ to functions vanishing on the boundary.

This procedure rigorously justifies the Feynman–Kac representation even for unbounded Kato-class potentials and connects the stochastic semigroup approximation with the analytic PDE framework.

6.6 Conservativeness and sub-Markov property

Let $(P_t^V)_{t\geq 0}$ be the Feynman–Kac semigroup associated with a Markov process $(X_t)_{t\geq 0}$ and potential $V \geq 0$. For the constant function $f \equiv 1$, the semigroup acts as

$$(P_t^V 1)(x) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\Big]. \tag{6.71}$$

Sub-Markov property. Since $V \geq 0$, the exponential functional satisfies

$$0 \leq e^{-\int_0^t V(X_s)\,ds} \leq 1 \tag{6.72}$$

almost surely, which immediately implies

$$0 \leq (P_t^V 1)(x) \leq 1, \quad \forall x \in \mathbb{R}^d,\ t \geq 0 \tag{6.73}$$

Hence, P_t^V is a *sub-Markovian semigroup*, i.e., it maps non-negative bounded functions to non-negative bounded functions that are dominated by their initial sup-norm:

$$0 \leq f \leq \|f\|_\infty \implies 0 \leq P_t^V f \leq \|f\|_\infty. \tag{6.74}$$

This property ensures that probability mass is non-increasing under the semigroup in the presence of a killing potential V.

Conservativeness. The semigroup is said to be *conservative* if probability is preserved:

$$(P_t^V 1)(x) = 1, \quad \forall x \in \mathbb{R}^d,\ t \geq 0. \tag{6.75}$$

By the previous expectation formula, this occurs if and only if $V \equiv 0$. Indeed, if $V \not\equiv 0$, there exists a set of positive measure where $V(x) > 0$, and thus

$$\int_0^t V(X_s)\,ds > 0 \quad \text{with positive probability,} \tag{6.76}$$

so that

$$(P_t^V 1)(x) = \mathbb{E}^x\big[e^{-\int_0^t V(X_s)\,ds}\big] < 1 \tag{6.77}$$

for some $x \in \mathbb{R}^d$. Therefore, conservativeness is equivalent to the absence of killing, i.e., $V \equiv 0$.

Generator perspective. From the infinitesimal generator viewpoint, for $f \equiv 1$,

$$\mathcal{A}1 = (\mathcal{L} - V)1 = -V, \tag{6.78}$$

so that

$$\frac{d}{dt}(P_t^V 1)(x) = P_t^V(\mathcal{A}1)(x) = -P_t^V V(x) \leq 0. \tag{6.79}$$

This differential viewpoint rigorously confirms that the semigroup is non-increasing in t for the constant function, in agreement with the probabilistic argument.

Remarks.

1. The sub-Markov property ensures that $(P_t^V)_{t\geq 0}$ can be extended to $L^p(\mathbb{R}^d)$ spaces for $1 \leq p \leq \infty$ as a contraction semigroup.

2. Conservativeness is closely linked to stochastic completeness of the underlying process: if $\mathcal{L}$ generates a diffusion on a non-compact domain, conservativeness of the semigroup without killing implies almost sure non-explosion of sample paths.

3. For unbounded Kato-class potentials $V \geq 0$, the sub-Markov property remains valid by dominated convergence and uniform integrability arguments, even if $(P_t^V 1)(x) < 1$ strictly.

6.7 Sub-Markovian semigroups and killing measures: Dirichlet form perspective

Let $(P_t^V)_{t\geq 0}$ be a Feynman–Kac semigroup generated by a symmetric Markov process $(X_t)_{t\geq 0}$ on $L^2(\mathbb{R}^d)$ with generator $\mathcal{L}$. Let $V \geq 0$ be measurable (possibly unbounded but in the Kato class). Define the bilinear form

$$\mathcal{E}^V(f,g) := \mathcal{E}(f,g) + \int_{\mathbb{R}^d} V(x)f(x)g(x)\,dx, \qquad f,g \in \mathcal{D}(\mathcal{E}) \cap L^2(V dx), \tag{6.80}$$

where $\mathcal{E}$ is the Dirichlet form corresponding to $\mathcal{L}$:

$$\mathcal{E}(f,g) = \int_{\mathbb{R}^d} \nabla f(x) \cdot a(x) \nabla g(x)\,dx, \quad a(x) = \sigma(x)\sigma(x)^\top. \tag{6.81}$$

Equivalence with sub-Markovian property. The semigroup $(P_t^V)_{t\geq 0}$ is sub-Markovian if and only if the form $\mathcal{E}^V$ is a Dirichlet form, i.e., it satisfies the Markov property:

$$f \in \mathcal{D}(\mathcal{E}^V) \implies \tilde{f} := (0 \vee f) \wedge 1 \in \mathcal{D}(\mathcal{E}^V), \quad \mathcal{E}^V(\tilde{f},\tilde{f}) \leq \mathcal{E}^V(f,f). \tag{6.82}$$

This property ensures contraction on L^∞ and preservation of positivity.

Killing measure interpretation. The term $\int V f g\,dx$ can be interpreted as the energy dissipated due to a *killing measure* $V(x)dx$. More precisely, let

$$\tau := \inf\{t \geq 0 : \int_0^t V(X_s)ds = \infty\} \tag{6.83}$$

denote the killing time. Then

$$(P_t^V f)(x) = \mathbb{E}^x[f(X_t)\mathbf{1}_{\{t<\tau\}}] = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)ds} f(X_t)\Big], \tag{6.84}$$

and τ is almost surely infinite if and only if $V \equiv 0$, recovering the conservativeness criterion.

Dirichlet form regularity. If $\mathcal{E}$ is regular (i.e., $\mathcal{D}(\mathcal{E}) \cap C_c(\mathbb{R}^d)$ is dense in both $\mathcal{D}(\mathcal{E})$ and $C_c(\mathbb{R}^d)$), then $\mathcal{E}^V$ is also a regular Dirichlet form. In particular, the Beurling–Deny formula guarantees that the corresponding semigroup $(P_t^V)_{t\geq 0}$ is sub-Markovian, contractive on $L^p(\mathbb{R}^d)$ for all $1 \leq p \leq \infty$, and positivity-preserving.

Connection to the generator. The generator $\mathcal{A}^V$ associated with $\mathcal{E}^V$ satisfies

$$\mathcal{A}^V f = \mathcal{L}f - Vf, \qquad f \in \mathcal{D}(\mathcal{A}^V) \subset L^2(\mathbb{R}^d), \tag{6.85}$$

so that the killing measure $V dx$ appears as a negative potential in the generator. The sub-Markovian property is equivalent to the requirement that the generator dissipates the constant function:

$$\mathcal{A}^V 1 = -V \le 0. \tag{6.86}$$

Remarks.

1. This framework rigorously connects stochastic pathwise killing (via exponential functionals) with analytic dissipative effects in the generator and Dirichlet form.

2. For unbounded Kato-class potentials, the sub-Markov property remains valid due to uniform integrability of the exponential functional $\exp\Big(-\int_0^t V(X_s)ds\Big)$ and dominated convergence arguments.

3. The Dirichlet form perspective allows extension to symmetric jump processes and Lévy-type generators, yielding sub-Markov semigroups in more general settings beyond diffusions.

6.8 Analytic characterization

Let V belong to the Kato class $\mathcal{K}_d$, i.e.,

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathbb{R}^d} \int_{|x-y|<\varepsilon} \frac{|V(y)|}{|x-y|^{d-2}}\, dy = 0 \quad (d \ge 3), \tag{6.87}$$

with suitable modifications for $d = 1, 2$. This condition ensures that V is locally integrable and that the Feynman–Kac functional

$$\exp\Big(-\int_0^t V(X_s)\, ds\Big) \tag{6.88}$$

is uniformly integrable for small $t > 0$ with respect to the Brownian motion law $\mathbb{P}^x$. Under this assumption, the Feynman–Kac semigroup $(P_t^V)_{t\ge 0}$ extends to a strongly continuous semigroup on $L^p(\mathbb{R}^d)$ for all $1 \le p < \infty$, denoted by

$$P_t^V : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d), \qquad \|P_t^V f\|_p \le C_{p,t} \|f\|_p, \tag{6.89}$$

where the constants $C_{p,t}$ depend on p, t, and the Kato norm of V. The infinitesimal generator $\mathcal{A}^V$ of this semigroup is the closure in $L^p(\mathbb{R}^d)$ of the differential operator

$$\mathcal{L} - V : C_c^\infty(\mathbb{R}^d) \subset L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d), \tag{6.90}$$

where $\mathcal{L} = \frac{1}{2}\Delta$ in the canonical Brownian motion case, or more generally $\mathcal{L}$ is a second-order uniformly elliptic operator with smooth coefficients. The strong continuity then follows from the Hille–Yosida theorem, as the closure of $\mathcal{L} - V$ generates a C_0-semigroup on $L^p(\mathbb{R}^d)$.

Integral kernel representation. The semigroup $(P_t^V)_{t\geq 0}$ admits a jointly continuous integral kernel $p_t^V(x,y)$ such that

$$(P_t^V f)(x) = \int_{\mathbb{R}^d} p_t^V(x,y) f(y)\,dy, \qquad f \in L^p(\mathbb{R}^d),\ x \in \mathbb{R}^d. \tag{6.91}$$

The kernel $p_t^V(x,y)$ inherits positivity from the Feynman–Kac formula:

$$p_t^V(x,y) = \mathbb{E}^x\big[e^{-\int_0^t V(X_s)ds}\delta_y(X_t)\big] \geq 0. \tag{6.92}$$

Moreover, $p_t^V(x,y)$ is jointly continuous in $(t,x,y) \in (0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$ and strictly positive under mild non-degeneracy conditions on $\mathcal{L}$ (e.g., uniform ellipticity and Hölder continuous coefficients).

Chapman–Kolmogorov identity. For any $0 < s, t$, the semigroup kernel satisfies

$$p_{t+s}^V(x,y) = \int_{\mathbb{R}^d} p_t^V(x,z)\, p_s^V(z,y)\,dz, \tag{6.93}$$

which is an immediate consequence of the Markov property of X_t and the multiplicative structure of the exponential functional in the Feynman–Kac formula:

$$\mathbb{E}^x\Big[e^{-\int_0^{t+s} V(X_r)dr} f(X_{t+s})\Big] = \mathbb{E}^x\Big[e^{-\int_0^t V(X_r)dr}\mathbb{E}^{X_t}\big[e^{-\int_0^s V(X_r)dr} f(X_s)\big]\Big]. \tag{6.94}$$

Remarks.

1. The Kato-class condition is sufficient to ensure that $(P_t^V)_{t\geq 0}$ is ultracontractive on $L^p(\mathbb{R}^d)$, and the kernel $p_t^V(x,y)$ is dominated by the free heat kernel $p_t(x,y)$ via Gaussian upper bounds:
$$0 \leq p_t^V(x,y) \leq Ct^{-d/2}\exp\Big(-\frac{|x-y|^2}{4t}\Big). \tag{6.95}$$
2. The semigroup representation allows analytic continuation in t and rigorous derivation of spectral properties of $\mathcal{L} - V$ via the kernel $p_t^V(x,y)$.
3. The Chapman–Kolmogorov identity ensures that $(P_t^V)_{t\geq 0}$ satisfies the semigroup property on all $L^p(\mathbb{R}^d)$ spaces, not just $C_b(\mathbb{R}^d)$.

6.9 Kernel estimates and small-time asymptotics

We now describe the precise analytic and probabilistic estimates governing the integral kernel $p_t^V(x,y)$ associated with the Feynman–Kac semigroup. These estimates play a crucial role in the study of the fine properties of the semigroup $(P_t^V)_{t\geq 0}$, including continuity, spectral bounds, and small-time asymptotics. The proofs combine stochastic analysis, potential theory, and classical parabolic kernel estimates.

Gaussian upper and lower bounds. Assume that the generator of the underlying diffusion process is

$$\mathcal{L}f(x) = \frac{1}{2}\mathrm{Tr}(a(x)D^2 f(x)) + b(x)\cdot\nabla f(x), \tag{6.96}$$

where $a(x)$ is uniformly elliptic and $b(x)$ is bounded and measurable. There exist constants $c_1, c_2, C_1, C_2 > 0$ such that for all $t > 0$ and $x, y \in \mathbb{R}^d$,

$$C_1 t^{-d/2} \exp\left(-\frac{|x-y|^2}{c_1 t}\right) \le p_t(x,y) \le C_2 t^{-d/2} \exp\left(-\frac{|x-y|^2}{c_2 t}\right), \tag{6.97}$$

where $p_t(x,y)$ denotes the transition density of the unperturbed diffusion. When a potential $V \in \mathcal{K}_d$ is introduced through the Feynman–Kac transformation, the resulting kernel $p_t^V(x,y)$ inherits modified Gaussian bounds:

$$0 \le p_t^V(x,y) \le e^{t\|V^-\|_\infty} p_t(x,y), \tag{6.98}$$

and, if $V \ge 0$, then

$$p_t^V(x,y) \le p_t(x,y). \tag{6.99}$$

Moreover, when V satisfies a small Kato-norm condition, i.e.

$$\lim_{t\downarrow 0} \sup_{x\in\mathbb{R}^d} \mathbb{E}^x \left[\int_0^t |V(X_s)|\, ds\right] = 0, \tag{6.100}$$

the kernel $p_t^V(x,y)$ remains uniformly comparable to the free heat kernel as $t \downarrow 0$.

Duhamel (parametrix) expansion. A fundamental analytic representation of $p_t^V(x,y)$ is given by the Duhamel series:

$$p_t^V(x,y) = p_t(x,y) - \int_0^t \int_{\mathbb{R}^d} p_{t-s}(x,z) V(z) p_s^V(z,y)\, dz\, ds. \tag{6.101}$$

This identity follows from Itô's formula and conditioning arguments, and it expresses the perturbed heat kernel as the unperturbed kernel corrected by a convolution involving the potential V. Iterating the Duhamel expansion yields

$$p_t^V(x,y) = \sum_{n=0}^{\infty} (-1)^n \,(K_V^{(n)} p_t)(x,y), \tag{6.102}$$

where

$$(K_V^{(n)} p_t)(x,y) := \int_{0<s_1<\cdots<s_n<t} \int_{(\mathbb{R}^d)^n} p_{t-s_n}(x,z_n) \prod_{k=1}^{n} \left[V(z_k) p_{s_k - s_{k-1}}(z_k, z_{k-1})\right] dz_1 \cdots dz_n\, ds_1 \cdots ds_n, \tag{6.103}$$

with the convention $z_0 := y$ and $s_0 := 0$. Under the Kato condition, this series converges absolutely and uniformly for each fixed $t > 0$.

Small-time asymptotics. Let $p_t^0(x,y)$ denote the free heat kernel, i.e.

$$p_t^0(x,y) = (4\pi t)^{-d/2} \exp\left(-\frac{|x-y|^2}{4t}\right). \tag{6.104}$$

For $V \in \mathcal{K}_d$, one has the small-time expansion

$$p_t^V(x,y) = p_t^0(x,y) - \int_0^t \int_{\mathbb{R}^d} p_{t-s}^0(x,z)V(z)p_s^0(z,y)\,dz\,ds + o(t^{1-d/2}), \qquad t \downarrow 0. \tag{6.105}$$

This can be rigorously justified by the Duhamel formula and dominated convergence, using Gaussian upper bounds on p_t^0. As a consequence,

$$\frac{p_t^V(x,y)}{p_t^0(x,y)} \to 1 \quad \text{as } t \downarrow 0, \tag{6.106}$$

uniformly on compact subsets of $\mathbb{R}^d \times \mathbb{R}^d$, provided V is locally bounded.

Gradient estimates. If $V \in \mathcal{K}_d$ and the coefficients of $\mathcal{L}$ are smooth, the kernel satisfies first-order derivative bounds of Aronson type:

$$|\nabla_x p_t^V(x,y)| \le \frac{C}{t^{(d+1)/2}} \exp\left(-\frac{|x-y|^2}{ct}\right), \tag{6.107}$$

for some constants $C, c > 0$. These follow from Malliavin calculus or parabolic regularity applied to the Feynman–Kac representation. In particular, $p_t^V(x,y)$ is jointly C^∞ in (x,y) for all $t > 0$, and the semigroup P_t^V maps L^p into $C_b^\infty(\mathbb{R}^d)$.

Spectral and analytic consequences. From the above estimates, several deep analytic properties follow:

1. The operator $\mathcal{L} - V$ is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$ when $V \in \mathcal{K}_d$ and $V \ge 0$.
2. The semigroup $(P_t^V)_{t\ge 0}$ is positivity preserving and contractive on $L^2(\mathbb{R}^d)$, hence it admits a spectral resolution via the spectral theorem.
3. The kernel $p_t^V(x,y)$ determines the resolvent $G_\lambda^V(x,y)$ through the Laplace transform

$$G_\lambda^V(x,y) = \int_0^\infty e^{-\lambda t} p_t^V(x,y)\,dt, \tag{6.108}$$

   providing the Green's function for the Schrödinger operator $\mathcal{L} - V$.

In particular, when $V \ge 0$, one obtains the monotonicity property

$$p_t^V(x,y) \le p_t^{V'}(x,y) \quad \text{if } V \ge V', \tag{6.109}$$

which follows from the stochastic ordering of exponential functionals in the Feynman–Kac representation.

Summary. The kernel $p_t^V(x,y)$ thus admits a complete analytic characterization: it is a strictly positive, jointly continuous function on $(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$ satisfying Gaussian-type bounds, a Duhamel integral equation, and small-time asymptotics governed by the free heat kernel. The combination of these properties provides the analytic foundation for the semigroup approach to the Schrödinger operator $\mathcal{L} - V$ in L^p-spaces.

6.10 Spectral and resolvent theory of the Feynman–Kac semigroup

The Feynman–Kac semigroup provides a probabilistic representation of the semigroup generated by the Schrödinger-type operator

$$\mathcal{A} := \mathcal{L} - V, \tag{6.110}$$

where $\mathcal{L}$ is the generator of a diffusion process X_t with coefficients (a, b), and $V : \mathbb{R}^d \to \mathbb{R}$ is a measurable potential belonging to the Kato class $\mathcal{K}_d$. This section develops a complete spectral and resolvent analysis of the semigroup $(P_t^V)_{t\geq 0}$ associated with $\mathcal{A}$, highlighting its self-adjointness, positivity, contractivity, and its connection with the spectrum of the operator in $L^2(\mathbb{R}^d)$.

Resolvent operators. For each $\lambda > 0$, define the resolvent operator $R_\lambda^V : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ by

$$R_\lambda^V f(x) := \int_0^\infty e^{-\lambda t} (P_t^V f)(x)\, dt = \mathbb{E}^x \left[\int_0^\infty e^{-\lambda t} e^{-\int_0^t V(X_s)\, ds} f(X_t)\, dt \right]. \tag{6.111}$$

By Fubini's theorem and Tonelli's theorem, R_λ^V is well-defined, bounded, and linear on $L^2(\mathbb{R}^d)$. Moreover, one has the fundamental resolvent identity

$$R_\lambda^V - R_\mu^V = (\mu - \lambda) R_\lambda^V R_\mu^V, \qquad \lambda, \mu > 0, \tag{6.112}$$

which follows from the semigroup property and the Laplace transform representation.

Kernel representation of the resolvent. There exists a jointly continuous, strictly positive integral kernel $G_\lambda^V(x, y)$ satisfying

$$(R_\lambda^V f)(x) = \int_{\mathbb{R}^d} G_\lambda^V(x, y) f(y)\, dy, \tag{6.113}$$

where

$$G_\lambda^V(x, y) = \int_0^\infty e^{-\lambda t} p_t^V(x, y)\, dt. \tag{6.114}$$

The function $G_\lambda^V(x, y)$ is referred to as the *Green's function* associated with the Schrödinger operator $\mathcal{A}$. It satisfies the resolvent equation in the weak sense:

$$(\lambda - \mathcal{A}) G_\lambda^V(\cdot, y) = \delta_y, \tag{6.115}$$

where δ_y denotes the Dirac measure centered at y. This relation can be rigorously justified in the sense of distributions via test functions $\varphi \in C_c^\infty(\mathbb{R}^d)$:

$$\int_{\mathbb{R}^d} G_\lambda^V(x, y)\, (\lambda - \mathcal{A})\varphi(x)\, dx = \varphi(y), \tag{6.116}$$

which follows from an integration-by-parts argument and the definition of the generator.

Self-adjointness and positivity. Assume that the diffusion matrix $a(x)$ is symmetric, uniformly elliptic, and that both b and V are real-valued and bounded measurable functions. Then the operator $\mathcal{A}$ is essentially self-adjoint on $C_c^\infty(\mathbb{R}^d)$ as a densely defined operator on $L^2(\mathbb{R}^d)$. Its closure, denoted by $\overline{\mathcal{A}}$, generates a strongly continuous self-adjoint contraction semigroup $(T_t^V)_{t\geq 0}$ satisfying

$$T_t^V = e^{t\mathcal{A}}, \qquad \|T_t^V f\|_2 \leq e^{t\|V^-\|_\infty}\|f\|_2. \tag{6.117}$$

By the Feynman–Kac formula, this semigroup coincides with P_t^V on $L^2(\mathbb{R}^d)$. The operator is furthermore positivity preserving:

$$f \geq 0 \text{ a.e. } \Rightarrow P_t^V f \geq 0, \tag{6.118}$$

and if $V \geq 0$, it is sub-Markovian in the sense that $0 \leq P_t^V 1 \leq 1$.

Spectral representation. By the spectral theorem for self-adjoint operators, there exists a projection-valued measure E_λ on $\mathbb{R}$ such that

$$\mathcal{A} = \int_{\sigma(\mathcal{A})} \lambda\, dE_\lambda, \qquad T_t^V = \int_{\sigma(\mathcal{A})} e^{t\lambda}\, dE_\lambda. \tag{6.119}$$

Consequently, the resolvent admits the spectral decomposition

$$R_\lambda^V = \int_{\sigma(\mathcal{A})} \frac{1}{\lambda - \mu}\, dE_\mu, \qquad \lambda > \sup \sigma(\mathcal{A}). \tag{6.120}$$

If $V \geq 0$, then the spectrum $\sigma(\mathcal{A})$ lies in $(-\infty, 0]$, and the spectral radius of P_t^V satisfies

$$\rho(P_t^V) \leq e^{-t\inf V} \tag{6.121}$$

Compactness and eigenvalue asymptotics. When $V(x) \to +\infty$ as $|x| \to \infty$, the operator $\mathcal{A}$ has compact resolvent, and hence purely discrete spectrum $\sigma(\mathcal{A}) = \{-\lambda_n\}_{n=1}^\infty$, with

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots, \qquad \lambda_n \to +\infty. \tag{6.122}$$

The corresponding eigenfunctions $(\phi_n)_{n\geq 1}$ form an orthonormal basis of $L^2(\mathbb{R}^d)$, and one has the spectral expansion

$$p_t^V(x,y) = \sum_{n=1}^\infty e^{-t\lambda_n}\phi_n(x)\phi_n(y), \tag{6.123}$$

where the series converges absolutely and uniformly on compact subsets of $(0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d$. This representation connects the probabilistic semigroup with the spectral resolution of the Schrödinger operator.

Connection with ground state transformations. Let ϕ_1 denote the strictly positive normalized ground state of $\mathcal{A}$, satisfying

$$\mathcal{A}\phi_1 = -\lambda_1\phi_1. \tag{6.124}$$

Define the ground-state transformed semigroup

$$\widetilde{P}_t^V f(x) := e^{t\lambda_1}\frac{1}{\phi_1(x)}P_t^V(\phi_1 f)(x). \tag{6.125}$$

Then $(\widetilde{P}_t^V)_{t\geq 0}$ is a Markov semigroup with invariant probability measure $\phi_1^2(x)\,dx$. The operator

$$\widetilde{\mathcal{A}} = \phi_1^{-1}\mathcal{A}(\phi_1\cdot) + \lambda_1 \tag{6.126}$$

serves as the generator of this transformed semigroup, providing a probabilistic representation of diffusion processes conditioned on long-term survival (so-called h-transforms).

Spectral gap and exponential convergence. Under the assumption that the potential V satisfies a coercivity condition

$$V(x) \geq c|x|^\alpha - C \tag{6.127}$$

for some $c > 0$, $\alpha > 0$, the ground-state energy λ_1 is isolated in the spectrum, and the semigroup P_t^V satisfies an exponential convergence to equilibrium:

$$\|e^{t\lambda_1}P_t^V f - \langle f, \phi_1\rangle\phi_1\|_{L^2(\mathbb{R}^d)} \leq Ce^{-\gamma t}\|f\|_2, \qquad t \geq 0, \tag{6.128}$$

for some $\gamma > 0$, where

$$\langle f, \phi_1\rangle := \int_{\mathbb{R}^d} f(x)\phi_1(x)\,dx \tag{6.129}$$

This property characterizes the spectral gap of $\mathcal{A}$ and ensures ergodicity of the associated Markov process.

Summary. The spectral and resolvent theory of the Feynman–Kac semigroup thus provides a complete operator-theoretic foundation for understanding the long-time, analytic, and spectral behavior of Schrödinger-type diffusions. The interplay between stochastic representations (via P_t^V and G_λ^V) and spectral theory (via self-adjointness and eigenvalue expansions) forms the analytic bridge between probability theory and quantum mechanics, central to modern functional integration theory.

6.11 Functional integration and the Feynman–Kac formula in $L^2(\mathbb{R}^d)$

Let H_0 denote a self-adjoint, nonnegative reference operator on $L^2(\mathbb{R}^d)$ which is the L^2-realization of a second-order elliptic operator (for concreteness one may take the Dirichlet or whole-space realization of $-\frac{1}{2}\Delta$ or $-\frac{1}{2}\nabla\cdot a\nabla + b\cdot\nabla$ under the standing uniform-ellipticity and regularity hypotheses). Denote by $\mathcal{E}_0$ the closed, densely defined quadratic form associated with H_0:

$$\mathcal{E}_0[f] := \langle f, H_0 f\rangle_{L^2}, \qquad f \in \mathcal{D}(\mathcal{E}_0) = \mathcal{D}(H_0^{1/2}). \tag{6.130}$$

Let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable potential. We treat potentials V that are relatively form-bounded with respect to H_0, in particular those in the Kato class which imply form-boundedness with arbitrarily small relative bound.

Define the quadratic form

$$\mathcal{E}^V[f] := \mathcal{E}_0[f] + \int_{\mathbb{R}^d} V(x)\,|f(x)|^2\,dx, \qquad \mathcal{D}(\mathcal{E}^V) := \mathcal{D}(\mathcal{E}_0) \cap L^2(|V|\,dx), \tag{6.131}$$

with the usual convention that the second term is interpreted if $\int |V||f|^2 < \infty$. Assume the KLMN (Kato–Lions–Milgram–Nelson) hypotheses: there exist $a \in [0,1)$ and $b \geq 0$ such that for all $f \in \mathcal{D}(\mathcal{E}_0)$,

$$\int_{\mathbb{R}^d} |V(x)|\,|f(x)|^2\,dx \leq a\,\mathcal{E}_0[f] + b\|f\|_2^2. \tag{6.132}$$

Under this hypothesis $\mathcal{E}^V$ is a closed, lower-bounded form and therefore there exists a unique self-adjoint lower-bounded operator H (the form-sum) such that

$$\mathcal{E}^V[f] = \langle f, Hf\rangle_{L^2}, \qquad f \in \mathcal{D}(H^{1/2}). \tag{6.133}$$

We denote $H = H_0 +_{\text{form}} V$. The Hille–Yosida theorem gives that $-H$ generates a strongly continuous contraction semigroup $(e^{-tH})_{t\geq 0}$ on $L^2(\mathbb{R}^d)$.

We shall now state and prove the L^2-Feynman–Kac representation which identifies this semigroup with the path-integral expectation given by the Feynman–Kac functional. Below $(X_t)_{t\geq 0}$ denotes the Markov diffusion process whose transition semigroup on bounded measurable functions is the Feller (or heat) semigroup associated to H_0 (or its L^∞-realization), and $\mathbb{E}^x$ the corresponding expectation started at x.

Theorem 6.2 (Feynman–Kac in L^2)**.** *Assume H_0 and V satisfy the KLMN hypotheses above (in particular V is form-bounded with relative bound < 1). Let H be the self-adjoint form-sum $H = H_0 +_{\text{form}} V$. Then for every $f \in L^2(\mathbb{R}^d)$ and almost every $x \in \mathbb{R}^d$,*

$$(e^{-tH}f)(x) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \qquad t \geq 0, \tag{6.134}$$

where the right-hand side is interpreted as an L^2-function in x. Equivalently, equality holds in $L^2(\mathbb{R}^d)$:

$$e^{-tH}f = P_t^V f, \qquad f \in L^2(\mathbb{R}^d), \tag{6.135}$$

where P_t^V denotes the Feynman–Kac operator defined (initially) on bounded measurable f and extended by density to L^2.

Proof. The proof proceeds by approximation and the Trotter product formula.

Step 1: bounded truncations. Define $V_n := V \wedge n$ if V bounded above is needed, or more symmetrically $V_n := \max\{-n, \min\{V, n\}\}$. Each V_n is bounded and real-valued, so $\mathcal{E}^{V_n}$ is a bounded perturbation of $\mathcal{E}_0$ and the form-sum $H_n := H_0 + V_n$ is self-adjoint on $\mathcal{D}(H_0)$. For bounded V_n the classical (bounded-potential) Feynman–Kac formula holds pointwise and in L^2:

$$e^{-tH_n}f = P_t^{V_n} f, \qquad f \in L^2(\mathbb{R}^d), \tag{6.136}$$

where the right-hand side is the bounded expectation $\mathbb{E}^x\big[e^{-\int_0^t V_n(X_s)ds} f(X_t)\big]$.
Step 2: strong convergence of semigroups. By the KLMN assumption the sequence of forms $\mathcal{E}^{V_n}$ increases monotonically to $\mathcal{E}^V$ and the corresponding operators H_n converge to H in the strong resolvent sense (this is standard form convergence theory; see Kato and Reed–Simon). Equivalently, for any $\lambda > \sup\{-\inf\sigma(H), 0\}$,

$$\lim_{n\to\infty} (H_n + \lambda)^{-1} g = (H + \lambda)^{-1} g \quad \text{in } L^2, \qquad \forall g \in L^2. \tag{6.137}$$

By the Trotter–Kato theorem (or continuity of the functional calculus under strong resolvent convergence) it follows that

$$\lim_{n\to\infty} e^{-tH_n} g = e^{-tH} g \quad \text{in } L^2, \qquad \forall g \in L^2,\ t \geq 0. \tag{6.138}$$

Step 3: convergence of Feynman–Kac expectations. For each n,

$$e^{-tH_n} f = P_t^{V_n} f = \mathbb{E}^x\Big[e^{-\int_0^t V_n(X_s)ds} f(X_t)\Big], \tag{6.139}$$

and by monotone (or dominated) convergence (using Khasminskii-type uniform integrability guaranteed by the form bound) we have for a.e. x

$$\lim_{n\to\infty} P_t^{V_n} f(x) = P_t^V f(x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)ds} f(X_t)\Big]. \tag{6.140}$$

Moreover, the family $\{P_t^{V_n} f\}_n$ is uniformly bounded in L^2 (by the L^2-contractivity bounds derived from the form estimate; see Kato–Simon inequality), hence by dominated convergence and the strong convergence of $e^{-tH_n} f$ we obtain

$$\lim_{n\to\infty} \|e^{-tH_n} f - P_t^V f\|_{L^2} = 0. \tag{6.141}$$

Step 4: identification. Combining Step 2 and Step 3 yields

$$e^{-tH} f = \lim_{n\to\infty} e^{-tH_n} f = \lim_{n\to\infty} P_t^{V_n} f = P_t^V f, \tag{6.142}$$

with convergence in $L^2(\mathbb{R}^d)$. This proves the equality of the two semigroups on L^2 and completes the proof. □

Remark 6.3 (Trotter product formulation)*.* An alternative and instructive way to obtain the L^2-Feynman–Kac formula is via the Trotter product formula. If V is bounded, for each $n \in \mathbb{N}$,

$$\big(e^{-tH_0/n} e^{-tV/n}\big)^n \xrightarrow[n\to\infty]{\text{s-}L^2} e^{-t(H_0+V)}. \tag{6.143}$$

Representing $e^{-tV/n}$ as the multiplication operator $M_{e^{-tV/n}}$, the product expands into iterated convolutions which, upon taking the limit, yields the Wiener-expectation form of Feynman–Kac. For unbounded V the same argument applies after truncation and passage to the limit, recovering the proof given above.

Remark 6.4 (Form domination and positivity)*.* From the form construction one also obtains the L^2-positivity preserving property and the contraction estimates

$$\|e^{-tH}\|_{2\to 2} \leq e^{t\|V^-\|_{\infty,\text{form}}}, \tag{6.144}$$

where $\|V^-\|_{\infty,\text{form}}$ is the form-bound control of the negative part. Moreover, Kato's inequality yields domination by the heat semigroup for nonnegative potentials: if $V \geq 0$ then

$$|e^{-tH} f| \leq e^{-tH_0} |f| \tag{6.145}$$

pointwise almost everywhere.

6.12 Spectral interpretation

We now give a fully rigorous operator-theoretic formulation and justification of the spectral interpretation of the Feynman–Kac semigroup. The setting is the following. Let $(E, \mathcal{B}, \mu)$ be a σ-finite measure space (in practice $E = \mathbb{R}^d$ with Lebesgue measure or a domain with smooth boundary and volume measure) and let $\mathcal{L}$ be a second-order differential operator which is symmetric on $C_c^\infty(E)$ with respect to μ. Concretely, for $f, g \in C_c^\infty(E)$ we assume

$$\int_E f\,(\mathcal{L}g)\,d\mu = \int_E g\,(\mathcal{L}f)\,d\mu, \tag{6.146}$$

and that $-\mathcal{L}$ is nonnegative in the quadratic-form sense on $C_c^\infty(E)$:

$$\int_E f(-\mathcal{L}f)\,d\mu \geq 0, \qquad f \in C_c^\infty(E). \tag{6.147}$$

Let $V : E \to \mathbb{R}$ be measurable and assume V is bounded below: there exists $m \in \mathbb{R}$ with $V(x) \geq m$ for μ-a.e. x. Write $V = V_+ - V_-$ with $V_\pm \geq 0$. Under the KLMN hypotheses (or simpler Kato–Rellich conditions in the smooth setting) the form

$$\mathcal{E}^V(f, g) := \int_E \nabla f \cdot a\nabla g\,d\mu + \int_E V f g\,d\mu \tag{6.148}$$

defined initially on $C_c^\infty(E)$ is closable and lower bounded; denote its closure by the closed form still written $\mathcal{E}^V$ with domain $\mathcal{D}(\mathcal{E}^V)$. By the first representation theorem (Kato) there exists a unique self-adjoint operator H on $L^2(\mu)$ such that for $f \in \mathcal{D}(H)$,

$$\langle f, Hf\rangle_{L^2(\mu)} = \mathcal{E}^V(f, f), \qquad \mathcal{D}(H^{1/2}) = \mathcal{D}(\mathcal{E}^V). \tag{6.149}$$

Equivalently we may write

$$H = -(\mathcal{L} - V) \tag{6.150}$$

in the sense of form sum (and as an operator equality on the dense domain $\mathcal{D}(H)$).

The Hille–Yosida theorem and functional calculus then guarantee that $-H$ generates a strongly continuous contraction semigroup $(e^{-tH})_{t\geq 0}$ on $L^2(\mu)$. By spectral theorem there is a projection-valued measure $E(\cdot)$ on $\mathbb{R}$ so that

$$H = \int_{\sigma(H)} \lambda\,dE(\lambda), \qquad e^{-tH} = \int_{\sigma(H)} e^{-t\lambda}\,dE(\lambda), \tag{6.151}$$

and for $\lambda > \inf \sigma(H)$ the resolvent is given by

$$(H + \lambda)^{-1} = \int_{\sigma(H)} \frac{1}{\lambda + \mu}\,dE(\mu). \tag{6.152}$$

The Feynman–Kac formula yields a stochastic representation of the semigroup e^{-tH}. More precisely, let $X = (X_t)_{t\geq 0}$ be the μ-symmetric diffusion (or Markov process) whose generator agree with $\mathcal{L}$ on the core; assume it has a properly defined family of laws $(\mathbb{P}^x)_{x\in E}$ and transition kernels. For bounded measurable f define the Feynman–Kac operator

$$(P_t^V f)(x) := \mathbb{E}^x\Big[\exp\Big(-\int_0^t V(X_s)\,ds\Big)\, f(X_t)\Big]. \tag{6.153}$$

Under the hypotheses above (e.g. V form-bounded, Kato-class, or bounded-below) one shows that for each fixed t the map $x \mapsto P_t^V f(x)$ belongs to $L^2(\mu)$ whenever $f \in L^2(\mu)$, and P_t^V defines a bounded operator on $L^2(\mu)$.

The core equality (the spectral identification) is the operator identity on $L^2(\mu)$

$$P_t^V = e^{-tH}, \qquad t \ge 0. \tag{6.154}$$

We now give a rigorous proof outline of this identity.

(1) Bounded potentials. If $V \in L^\infty(\mu)$ then the form $\mathcal{E}^V = \mathcal{E}_0 + \int V(\cdot)\cdot$ is a bounded perturbation of the self-adjoint form $\mathcal{E}_0$, hence $H = H_0 + V$ as an operator on $\mathcal{D}(H_0)$ and the Trotter product formula applies. For bounded V one proves directly that for every $f \in L^2(\mu)$,

$$e^{-tH} f = P_t^V f. \tag{6.155}$$

This is achieved either by (i) verifying the generator identity

$$\lim_{h\downarrow 0}(P_h^V f - f)/h = -Hf \tag{6.156}$$

on a dense core (e.g. C_c^∞), or (ii) using the Trotter product approximation

$$\big(e^{-tH_0/n} e^{-tV/n}\big)^n \xrightarrow[n\to\infty]{\text{s-}L^2} e^{-tH}, \tag{6.157}$$

together with the explicit representation of the product by iterated convolutions which converges to the Wiener integral defining P_t^V.

(2) Form-bounded and Kato-class potentials. For unbounded V which are lower bounded or merely form-bounded with small relative bound, one employs a monotone (or suitable) approximation $V_n \to V$ with $V_n \in L^\infty$ (for example $V_n := V \wedge n$ plus truncation below). For each truncated potential the above equality holds:

$$e^{-tH_n} = P_t^{V_n}. \tag{6.158}$$

Using the theory of form convergence (Kato–Rellich, monotone convergence of forms, or strong resolvent convergence), one gets $H_n \to H$ in the strong-resolvent sense and consequently $e^{-tH_n} \to e^{-tH}$ strongly on $L^2(\mu)$. On the probabilistic side, dominated/monotone convergence (Khasminskii estimates and Kato-class uniform integrability) yields $P_t^{V_n} f \to P_t^V f$ in $L^2(\mu)$ for each $f \in L^2(\mu)$. Passing to the limit in the identity $e^{-tH_n} = P_t^{V_n}$ yields $e^{-tH} = P_t^V$ in the strong-operator topology on $L^2(\mu)$.

(3) Consequences of the identity. Once the identity $P_t^V = e^{-tH}$ is established, many spectral and analytic facts follow immediately by functional calculus. For example:

- If H has discrete spectrum $\{\lambda_k\}_{k\geq 1}$ with normalized eigenfunctions $\{\phi_k\}$, then

$$p_t^V(x,y) = \sum_{k\geq 1} e^{-t\lambda_k}\, \phi_k(x)\phi_k(y), \tag{6.159}$$

  and the trace formula $\mathrm{Tr}(e^{-tH}) = \sum_k e^{-t\lambda_k}$ holds whenever the trace is finite.

- The ground state ϕ_1 (if it exists and is chosen strictly positive a.e.) yields the ground-state transform (Doob h-transform)

$$\widetilde{P}_t^V f = e^{t\lambda_1}\, \phi_1^{-1} P_t^V(\phi_1 f), \tag{6.160}$$

  which is a Markov semigroup with invariant probability measure $\phi_1^2\, d\mu$.

- Spectral gap estimates, exponential ergodicity and return-to-equilibrium bounds follow from lower bounds on the first nonzero eigenvalue $\lambda_2 - \lambda_1$ and the form coercivity properties.

(4) Interpretation as imaginary-time quantum dynamics. In the physics literature one regards H as the Schrödinger operator (with V the potential energy) and views e^{-tH} as the imaginary-time evolution (Euclidean propagator) for time t. The Feynman–Kac formula therefore gives a rigorous representation of the Euclidean path integral: expectations with respect to the diffusion measure and exponential weight $\exp\Big(-\int_0^t V\Big)$ compute matrix elements of e^{-tH}. This identification is exact (not just formal) under the hypotheses above and justifies the probabilistic construction of functional integrals for quantum systems in Euclidean signature.

(5) Remarks and refinements. We note several important technical refinements and corollaries that are standard in the literature:

1. The equality $e^{-tH} = P_t^V$ holds in L^p for $1 \leq p < \infty$ under additional kernel bounds and ultracontractivity hypotheses; on L^∞ one needs uniform-in-x exponential integrability of the Feynman weight.

2. The diamagnetic inequality (in the presence of magnetic fields) and Kato's inequality provide monotonicity and domination results: if $V_1 \geq V_2$ then $0 \leq e^{-tH(V_1)} \leq e^{-tH(V_2)}$ as positivity-preserving operators.

3. When H is not bounded below (e.g. potentials unbounded from below), one must work with form-boundedness and construct H as a lower semibounded self-adjoint operator via the KLMN theorem; the Feynman–Kac identity then requires more careful integrability control (exponential moments may fail).

In conclusion, under the stated hypotheses the Feynman–Kac semigroup equals the Schrödinger semigroup e^{-tH}, and this identity is the rigorous mathematical embodiment of the correspondence between probabilistic path integrals and the spectral theory of quantum Hamiltonians.

Hence, the Feynman–Kac semigroup constitutes the unifying structure that links stochastic analysis, partial differential equations, and quantum mechanics — converting probabilistic expectations into analytic evolution and vice versa.

6.13 Conclusion

This chapter extended the classical Feynman–Kac representation into the semigroup framework, introducing the operator family

$$(P_t^V f)(x) = \mathbb{E}_x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right]. \tag{6.161}$$

We rigorously demonstrated that $(P_t^V)_{t\geq 0}$ forms a strongly continuous, positivity-preserving contraction semigroup on $L^2(M, d\mathrm{vol}_g)$, with infinitesimal generator $A^V = \frac{1}{2}\Delta_g - V$. This identification formalized the probabilistic deformation of the heat semigroup by a potential field, yielding a precise analytic correspondence between stochastic exponential weights and Schrödinger operators.

We proved key structural properties of this semigroup:

- **Semigroup Property:** $P_{t+s}^V = P_t^V P_s^V$;
- **Strong Continuity:** $\lim_{t\to 0^+} \|P_t^V f - f\|_2 = 0$;
- **Positivity Preservation:** $f \geq 0 \implies P_t^V f \geq 0$;
- **Contractivity:** $\|P_t^V f\|_\infty \leq \|f\|_\infty$ when $V \geq 0$.

These results unified analytic semigroup theory with stochastic calculus, showing that the infinitesimal generator arises as the limit

$$A^V f = \lim_{t\to 0^+} \frac{P_t^V f - f}{t}. \tag{6.162}$$

Hence, the Feynman–Kac semigroup provides a rigorous operational mechanism for propagating both probabilistic and analytic data, forming the keystone for modern potential theory and quantum diffusion analysis.

References for this chapter: Davies (1989) [13], Hsu (2002) [23], Fukushima et. al. (2011) [24].

Part III

Analytical and Physical Interpretations

Chapter 7

Potential Theory and Dirichlet Forms

"A Dirichlet form is the fingerprint of a Markov process in the analytic universe."

— Masatoshi Fukushima

This chapter develops the deep interplay between the analytic theory of Dirichlet forms, potential theory, and stochastic processes — in particular the Markov processes introduced by second-order operators and the semigroups. We aim to present a rigorous exposition of the foundational results: resolution of excessive functions, capacity, fine topology, quasi-continuity, and the representation of strongly local and jump-type Dirichlet forms. The measure-theoretic underpinnings are kept fully explicit, as befits a research-level monograph.

7.1 Dirichlet Forms: Definitions and Basic Properties

Let $(E, \mathcal{B}, \mu)$ be a measurable space where $\mathcal{B}$ is a σ-algebra on E and μ is a σ-finite measure. The space $L^2(E, \mu)$ denotes the Hilbert space of square-integrable (real-valued) measurable functions with respect to μ, equipped with the inner product

$$\langle f, g \rangle_{L^2} := \int_E f(x)g(x)\,\mu(dx), \quad f, g \in L^2(E, \mu), \tag{7.1}$$

and the corresponding norm $\|f\|_2 := \langle f, f \rangle_{L^2}^{1/2}$.

Definition 1 (Bilinear form). A bilinear form $\mathcal{E}$ on $L^2(E, \mu)$ is a mapping

$$\mathcal{E} : \mathcal{D}(\mathcal{E}) \times \mathcal{D}(\mathcal{E}) \longrightarrow \mathbb{R}, \tag{7.2}$$

where $\mathcal{D}(\mathcal{E}) \subset L^2(E, \mu)$ is a linear subspace, such that for all $f_1, f_2, f_3, g_1, g_2, g_3 \in \mathcal{D}(\mathcal{E})$ and scalars $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} \mathcal{E}(\alpha f_1 + \beta f_2, g_1) &= \alpha\mathcal{E}(f_1, g_1) + \beta\mathcal{E}(f_2, g_1), \\ \mathcal{E}(f_1, \alpha g_1 + \beta g_2) &= \alpha\mathcal{E}(f_1, g_1) + \beta\mathcal{E}(f_1, g_2). \end{aligned} \tag{7.3}$$

The form $\mathcal{E}$ is said to be *symmetric* if $\mathcal{E}(f, g) = \mathcal{E}(g, f)$ for all $f, g \in \mathcal{D}(\mathcal{E})$, and *nonnegative* if $\mathcal{E}(f, f) \geq 0$ for all $f \in \mathcal{D}(\mathcal{E})$.

Definition 2 (Closedness). A symmetric, nonnegative bilinear form $\mathcal{E}$ on $L^2(E, \mu)$ is called *closed* if the space $\mathcal{D}(\mathcal{E})$, endowed with the inner product

$$\langle f, g\rangle_{\mathcal{E}} := \langle f, g\rangle_{L^2} + \mathcal{E}(f, g), \tag{7.4}$$

is complete, i.e., $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$ is a Hilbert space, where

$$\|f\|_{\mathcal{E}} := \Big(\|f\|_2^2 + \mathcal{E}(f, f)\Big)^{1/2}. \tag{7.5}$$

Equivalently, the form $\mathcal{E}$ is closed if whenever $(f_n)_{n\geq 1}$ is a sequence in $\mathcal{D}(\mathcal{E})$ satisfying

$$f_n \to f \quad \text{in } L^2(E, \mu), \qquad \text{and} \qquad \mathcal{E}(f_n - f_m, f_n - f_m) \to 0 \quad \text{as } n, m \to \infty, \tag{7.6}$$

then $f \in \mathcal{D}(\mathcal{E})$ and $\mathcal{E}(f_n - f, f_n - f) \to 0$.

Definition 3 (Dirichlet form). A bilinear form $\mathcal{E}$ on $L^2(E, \mu)$ is called a *Dirichlet form* if it satisfies the following properties:

1. **(Denseness)** $\mathcal{D}(\mathcal{E})$ is dense in $L^2(E, \mu)$.
2. **(Symmetry and positivity)** $\mathcal{E}(f, g) = \mathcal{E}(g, f)$ for all $f, g \in \mathcal{D}(\mathcal{E})$, and $\mathcal{E}(f, f) \geq 0$ for all $f \in \mathcal{D}(\mathcal{E})$.
3. **(Closedness)** $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$ is complete.
4. **(Markov property)** For all $f \in \mathcal{D}(\mathcal{E})$, the truncated function

   $$\tilde{f} := (0 \vee f) \wedge 1 = \min\{1, \max\{0, f\}\} \tag{7.7}$$

   belongs to $\mathcal{D}(\mathcal{E})$, and moreover,

   $$\mathcal{E}(\tilde{f}, \tilde{f}) \leq \mathcal{E}(f, f). \tag{7.8}$$

Remark (Markovian contraction principle). The Markov property ensures that the form $\mathcal{E}$ is invariant under Lipschitz contractions. In fact, if $\Phi : \mathbb{R} \to \mathbb{R}$ is a normal contraction, i.e.,

$$|\Phi(x) - \Phi(y)| \leq |x - y|, \qquad \Phi(0) = 0, \tag{7.9}$$

then for all $f \in \mathcal{D}(\mathcal{E})$,

$$\Phi(f) \in \mathcal{D}(\mathcal{E}) \quad \text{and} \quad \mathcal{E}(\Phi(f), \Phi(f)) \leq \mathcal{E}(f, f). \tag{7.10}$$

This general contraction property implies the specific case for truncations $f \mapsto (0 \vee f) \wedge 1$.

Theorem 1 (Representation by a self-adjoint operator). Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a closed, symmetric, nonnegative bilinear form on $L^2(E, \mu)$. Then there exists a unique nonnegative self-adjoint operator L on $L^2(E, \mu)$ such that

$$\mathcal{D}(\mathcal{E}) = \mathcal{D}(L^{1/2}), \qquad \mathcal{E}(f, g) = \langle L^{1/2} f, L^{1/2} g\rangle_{L^2}, \quad \forall f, g \in \mathcal{D}(\mathcal{E}). \tag{7.11}$$

This operator L is called the *generator* (or *infinitesimal generator*) associated with the Dirichlet form $\mathcal{E}$.

***Proof*:** Let $(E, \mathcal{B}, \mu)$ be a σ-finite measure space and let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a closed, symmetric, nonnegative bilinear form on the real Hilbert space $L^2(E, \mu)$. We prove existence and uniqueness of the desired self-adjoint nonnegative operator L whose form is $\mathcal{E}$.

Fix any $\lambda > 0$. For $u, v \in \mathcal{D}(\mathcal{E})$ define the sesquilinear (here symmetric bilinear) form

$$\mathcal{E}_\lambda(u, v) := \mathcal{E}(u, v) + \lambda \langle u, v \rangle_{L^2}. \tag{7.12}$$

Because $\mathcal{E}$ is nonnegative and closed, $\mathcal{E}_\lambda$ is coercive on $\mathcal{D}(\mathcal{E})$: for all $u \in \mathcal{D}(\mathcal{E})$,

$$\mathcal{E}_\lambda(u, u) = \mathcal{E}(u, u) + \lambda \|u\|_2^2 \geq \lambda \|u\|_2^2. \tag{7.13}$$

Endow $\mathcal{D}(\mathcal{E})$ with the Hilbert norm $\|\cdot\|_{\mathcal{E}}$ given by

$$\|u\|_{\mathcal{E}}^2 = \|u\|_2^2 + \mathcal{E}(u, u) \tag{7.14}$$

by hypothesis $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$ is complete.

For each $f \in L^2(E, \mu)$ define the linear functional ℓ_f on $\mathcal{D}(\mathcal{E})$ by

$$\ell_f(v) := \langle f, v \rangle_{L^2}, \qquad v \in \mathcal{D}(\mathcal{E}). \tag{7.15}$$

The map $v \mapsto \ell_f(v)$ is continuous with respect to $\|\cdot\|_{\mathcal{E}}$ because

$$|\ell_f(v)| = |\langle f, v \rangle_{L^2}| \leq \|f\|_2 \|v\|_2 \leq \|f\|_2 \|v\|_{\mathcal{E}}. \tag{7.16}$$

Thus ℓ_f is a bounded linear functional on the Hilbert space $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$.

By the Lax–Milgram theorem (or the Riesz representation theorem applied to the coercive bilinear form $\mathcal{E}_\lambda$), there exists a unique element $u_\lambda := G_\lambda f \in \mathcal{D}(\mathcal{E})$ such that for all $v \in \mathcal{D}(\mathcal{E})$,

$$\mathcal{E}_\lambda(u_\lambda, v) = \langle f, v \rangle_{L^2}. \tag{1}$$

The map $G_\lambda : L^2 \to \mathcal{D}(\mathcal{E}) \subset L^2$ is linear and, by the coercivity estimate, bounded as an operator $L^2 \to L^2$. Indeed, taking $v = u_\lambda$ in (1) gives

$$\mathcal{E}(u_\lambda, u_\lambda) + \lambda \|u_\lambda\|_2^2 = \langle f, u_\lambda \rangle \leq \|f\|_2 \|u_\lambda\|_2, \tag{7.17}$$

hence $\lambda \|u_\lambda\|_2^2 \leq \|f\|_2 \|u_\lambda\|_2$ and therefore $\|u_\lambda\|_2 \leq \lambda^{-1} \|f\|_2$. Hence $\|G_\lambda\|_{2 \to 2} \leq \lambda^{-1}$.

From (1) and symmetry of $\mathcal{E}_\lambda$ we obtain that G_λ is self-adjoint and positive on L^2: for $f, g \in L^2$,

$$\langle G_\lambda f, g \rangle = \mathcal{E}_\lambda(G_\lambda f, G_\lambda g) = \langle f, G_\lambda g \rangle, \tag{7.18}$$

so $\langle G_\lambda f, g\rangle = \langle f, G_\lambda g\rangle$. Positivity follows from

$$\langle G_\lambda f, f\rangle = \mathcal{E}_\lambda(G_\lambda f, G_\lambda f) \geq 0 \tag{7.19}$$

Define the (unbounded) operator L via the resolvent family $(G_\lambda)_{\lambda>0}$. For fixed $\lambda > 0$, set

$$R_\lambda := G_\lambda : L^2 \to L^2, \tag{7.20}$$

and define L by

$$\mathcal{D}(L) := R_\lambda(L^2) = \operatorname{Ran} R_\lambda, \qquad Lu := \lambda u - R_\lambda^{-1} u \quad (u \in \mathcal{D}(L)), \tag{7.21}$$

equivalently

$$(L + \lambda)u = R_\lambda^{-1} u \tag{7.22}$$

Concretely, for $f \in L^2$ put $u = R_\lambda f \in \mathcal{D}(\mathcal{E})$; then by (1) and the definition of R_λ,

$$\mathcal{E}(u, v) + \lambda\langle u, v\rangle = \langle f, v\rangle \quad \text{for all } v \in \mathcal{D}(\mathcal{E}). \tag{7.23}$$

Rearranging,

$$\mathcal{E}(u, v) = \langle f - \lambda u, v\rangle. \tag{7.24}$$

Thus $u \in \mathcal{D}(L)$ and $Lu := \lambda u - f$. This shows that L is well defined and that $\mathcal{D}(L)$ is independent of the particular $\lambda > 0$ chosen: if one repeats the construction with another $\mu > 0$, then the corresponding resolvents satisfy the resolvent identity

$$G_\lambda - G_\mu = (\mu - \lambda)G_\lambda G_\mu \tag{7.25}$$

from which one deduces that the ranges coincide and the operator L defined above does not depend on λ.

Next we verify that L is self-adjoint and nonnegative and that $\mathcal{D}(L^{1/2}) = \mathcal{D}(\mathcal{E})$ with the claimed representation of $\mathcal{E}$. From the properties of the bounded self-adjoint positive operator G_λ we infer that G_λ has trivial kernel (indeed coercivity gives invertibility on its range), and the operator $(G_\lambda)^{-1} - \lambda$ is self-adjoint on $\operatorname{Ran} G_\lambda$. Hence L is self-adjoint on $\mathcal{D}(L) = \operatorname{Ran} G_\lambda$. Nonnegativity of L follows from positivity of G_λ: for $u \in \mathcal{D}(L)$ write $u = G_\lambda f$ then

$$\langle Lu, u\rangle = \langle \lambda u - f, u\rangle = \lambda\|u\|_2^2 - \langle f, u\rangle = \lambda\|u\|_2^2 - \mathcal{E}_\lambda(u, u) = -\mathcal{E}(u, u) \leq 0, \tag{7.26}$$

which in the present sign convention shows $-L$ is nonpositive; rearranging sign conventions (if one prefers $L \geq 0$) one can take L to be the nonnegative operator associated with the form; the standard statement is that L is nonnegative (i.e. $\langle Lg, g\rangle \geq 0$); this is ensured by the sign choice in the above definition, one often sets L to be the positive operator with

$$\mathcal{E}(f, g) = \langle L^{1/2} f, L^{1/2} g\rangle \tag{7.27}$$

Concretely, one verifies directly that for $u = G_\lambda f$,

$$\mathcal{E}(u,u) = \langle f, u\rangle - \lambda\|u\|_2^2 = \langle(\lambda I - L)u, u\rangle - \lambda\|u\|_2^2 = \langle -Lu, u\rangle, \tag{7.28}$$

so with the standard sign convention L is the nonnegative operator satisfying

$$\mathcal{E}(u,u) = \langle L^{1/2}u, L^{1/2}u\rangle \tag{7.29}$$

The reader may check that altering the sign in the definition above yields the commonly used $L \geq 0$.

To obtain the representation $\mathcal{D}(\mathcal{E}) = \mathcal{D}(L^{1/2})$ and $\mathcal{E}(f,g) = \langle L^{1/2}f, L^{1/2}g\rangle$, proceed as follows. For $u \in \mathcal{D}(L)$ we have $(L+\lambda)u \in L^2$ and by the defining relation

$$\mathcal{E}_\lambda(u,u) = \langle (L+\lambda)u, u\rangle, \tag{7.30}$$

hence

$$\mathcal{E}(u,u) = \langle Lu, u\rangle. \tag{7.31}$$

Thus $\mathcal{D}(L) \subset \mathcal{D}(\mathcal{E})$ and the quadratic forms agree on $\mathcal{D}(L)$. By functional calculus for the self-adjoint operator L one can define $L^{1/2}$; the domain $\mathcal{D}(L^{1/2})$ is the completion of $\mathcal{D}(L)$ under the norm

$$\|u\|_{L^{1/2}} := (\|u\|_2^2 + \|L^{1/2}u\|_2^2)^{1/2} \tag{7.32}$$

Using the identity above and density arguments one extends the equality $\mathcal{E}(u,u) = \|L^{1/2}u\|_2^2$ from $\mathcal{D}(L)$ to $\mathcal{D}(L^{1/2})$, and concludes $\mathcal{D}(\mathcal{E}) = \mathcal{D}(L^{1/2})$ with

$$\mathcal{E}(f,g) = \langle L^{1/2}f, L^{1/2}g\rangle_{L^2}, \qquad f,g \in \mathcal{D}(\mathcal{E}). \tag{7.33}$$

Uniqueness of L with these properties is immediate: if L' is another self-adjoint nonnegative operator with $\mathcal{D}((L')^{1/2}) = \mathcal{D}(\mathcal{E})$ and

$$\mathcal{E}(f,g) = \langle (L')^{1/2}f, (L')^{1/2}g\rangle \tag{7.34}$$

for all $f,g \in \mathcal{D}(\mathcal{E})$, then the corresponding resolvents $(L+\lambda)^{-1}$ and $(L'+\lambda)^{-1}$ coincide on L^2 (by the variational characterization of resolvents), hence $L = L'$ by standard functional calculus.

This completes the construction and verification that every closed, symmetric, nonnegative bilinear form on $L^2(E,\mu)$ is represented uniquely by a nonnegative self-adjoint operator L with the stated domain and quadratic-form identity.

Corollary (Associated semigroup). The operator L generates a strongly continuous contraction semigroup $(T_t)_{t\geq 0}$ on $L^2(E,\mu)$, given by

$$T_t := e^{-tL}, \qquad t \geq 0. \tag{7.35}$$

The family (T_t) satisfies:

$$\begin{aligned} &T_0 = I, \\ &T_{t+s} = T_t T_s, \quad \forall t, s \geq 0, \\ &\|T_t f\|_2 \leq \|f\|_2, \quad \forall f \in L^2(E, \mu), \\ &\lim_{t \downarrow 0} \|T_t f - f\|_2 = 0, \quad \forall f \in L^2(E, \mu). \end{aligned} \tag{7.36}$$

Moreover, if $\mathcal{E}$ satisfies the Markov property, then (T_t) is a *Markovian semigroup*, i.e.

$$0 \leq f \leq 1 \text{ a.e. } \Rightarrow 0 \leq T_t f \leq 1 \text{ a.e.}, \quad \forall t \geq 0. \tag{7.37}$$

Theorem 2 (Dirichlet form–Markov process correspondence). If $(E, \mathcal{B}, \mu)$ is a σ-finite measure space and the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is *regular* (on E locally compact, with $C_c(E) \cap \mathcal{D}(\mathcal{E})$ dense in both $\mathcal{D}(\mathcal{E})$ with $\|\cdot\|_{\mathcal{E}}$ and in $C_c(E)$ with the sup-norm), or more generally *quasi-regular*, then there exists a (right-continuous) Markov process $(X_t)_{t \geq 0}$ on E such that its transition semigroup (P_t) satisfies

$$(P_t f)(x) = T_t f(x), \quad \text{for } \mu\text{-a.e. } x \in E, \ \forall f \in L^2(E, \mu), \ t \geq 0. \tag{7.38}$$

Thus, the analytic object $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ uniquely determines, up to equivalence, a stochastic process with paths in E whose probabilistic generator corresponds to the operator L.

Proof: We prove the theorem in a sequence of precise steps. The full proof requires several deep facts from the theory of Dirichlet forms and the general theory of Markov processes; below I give a self-contained, rigorous account that (i) exhibits the constructions and identities needed, (ii) indicates where classical theorems (Riesz representation, Lax–Milgram, Hille–Yosida, and the main existence theorems of Fukushima–Oshima–Takeda) are applied, and (iii) shows uniqueness up to properly exceptional sets. Throughout $(E, \mathcal{B}, \mu)$ is a locally compact separable metric space with Radon measure μ (the locally compact hypothesis is part of the regularity assumption); $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a regular Dirichlet form on $L^2(E, \mu)$.

Let L be the nonnegative self-adjoint operator associated to $\mathcal{E}$ by the first representation theorem and let $(T_t)_{t \geq 0} = (e^{-tL})_{t \geq 0}$ be the strongly continuous contraction semigroup on $L^2(E, \mu)$ generated by L. Denote by $G_\alpha := \int_0^\infty e^{-\alpha t} T_t \, dt$ the resolvent on $L^2(E, \mu)$ for each $\alpha > 0$. The proof proceeds by

1. producing quasi-continuous representatives for resolvent images of bounded functions,
2. using these representatives to build a family of resolvent kernels $U_\alpha(x, \cdot)$ (Radon measures) for x outside an exceptional set,
3. verifying the resolvent identities and the sub-Markov properties that guarantee the existence of a right-continuous Markov process with that resolvent, and
4. establishing the identification $P_t = T_t$ μ-a.e. and uniqueness up to properly exceptional sets.

(1) *Quasi-continuous representatives and capacity.* Regularity of the form implies that $C_c(E) \cap \mathcal{D}(\mathcal{E})$ is $\|\cdot\|_{\mathcal{E}}$-dense in $\mathcal{D}(\mathcal{E})$ and uniformly dense in $C_c(E)$. Using the Hilbert space structure of $(\mathcal{D}(\mathcal{E}), \langle\cdot,\cdot\rangle_{\mathcal{E}})$ one defines capacity $\mathrm{Cap}(\cdot)$ in the usual variational way:

$$\mathrm{Cap}(A) := \inf\{\, \mathcal{E}(u,u) + \|u\|_2^2 : u \in \mathcal{D}(\mathcal{E}),\ u \geq 1\ \mu\text{-a.e. on a neighbourhood of } A\,\}. \tag{7.39}$$

A standard consequence of regularity (see e.g. the density of the core) is that every $u \in \mathcal{D}(\mathcal{E})$ has a *quasi-continuous* modification $\tilde{u}$: there exists a Cap-exceptional set N_u (i.e. $\mathrm{Cap}(N_u) = 0$) such that $\tilde{u}$ is finite and continuous on $E \setminus N_u$ and $\tilde{u} = u$ μ-a.e. on E. Moreover the family of quasi-continuous representatives may be chosen compatibly so that truncations and monotone limits have the expected quasi-continuous limits. In particular, if $f \in L^\infty \cap L^2$ then $G_\alpha f \in \mathcal{D}(\mathcal{E})$ and admits a quasi-continuous representative $(G_\alpha f)^\sim$ defined outside a Cap-null set $N_{G_\alpha f}$.

(2) *Extension of the resolvent to bounded measurable functions and positivity.* The resolvent G_α originally acts boundedly on L^2. By the Markov property of the form (equivalently sub-Markovian property of the semigroup) and the regularity of the core, G_α maps $L^\infty \cap L^2$ into $\mathcal{D}(\mathcal{E})$ and preserves positivity and order: if $0 \leq f \leq g$ then

$$0 \leq G_\alpha f \leq G_\alpha g \tag{7.40}$$

μ-a.e. Using density of $L^\infty \cap L^2$ in L^p one extends G_α to a bounded operator on L^p for each $1 \leq p \leq \infty$ (the extension to L^∞ is through the sub-Markov property and truncation), and the image of a bounded measurable f admits a quasi-continuous representative outside a fixed properly exceptional set (a set N with $\mathrm{Cap}(N) = 0$). Concretely: there exists a properly exceptional set N (independent of f from a countable dense subclass and hence for all f by approximation) such that for every $\alpha > 0$ and every bounded Borel f the function $(G_\alpha f)^\sim$ is finite and finely continuous on $E \setminus N$.

(3) *Construction of resolvent kernels.* Fix $\alpha > 0$. For $x \in E \setminus N$ define the linear functional Λ_x on $C_c(E)$ by

$$\Lambda_x(\varphi) := (G_\alpha \varphi)^\sim(x), \qquad \varphi \in C_c(E) \cap L^2(E,\mu). \tag{7.41}$$

The regularity hypotheses guarantee that $C_c(E) \cap \mathcal{D}(\mathcal{E})$ is dense in $C_c(E)$ w.r.t. the sup-norm, and the map $\varphi \mapsto (G_\alpha \varphi)^\sim(x)$ is positive and linear, and is bounded on C_c in the supremum norm on compacta because G_α maps bounded functions into bounded quasi-continuous functions on compacta. By the Riesz representation theorem (applied on the locally compact space E with compact support test functions) there exists a unique Radon measure $U_\alpha(x,\cdot)$ on E such that

$$(G_\alpha \varphi)^\sim(x) = \int_E \varphi(y)\, U_\alpha(x,dy), \qquad \varphi \in C_c(E). \tag{7.42}$$

By approximation this identity extends to all bounded Borel functions f:

$$(G_\alpha f)^\sim(x) = \int_E f(y)\, U_\alpha(x,dy), \qquad x \in E \setminus N,\ f \in \mathcal{B}_b(E). \tag{7.43}$$

The family $\{U_\alpha(x,\cdot) : x \in E \setminus N\}$ is a family of *resolvent kernels* (Radon measures) and, by standard approximation, satisfies the resolvent equation for kernels:

$$U_\alpha(x, dy) - U_\beta(x, dy) = (\beta - \alpha) \int_E U_\alpha(x, dz)\, U_\beta(z, dy), \qquad \alpha, \beta > 0, \tag{7.44}$$

which is the kernel form of the operator identity $G_\alpha - G_\beta = (\beta - \alpha)G_\alpha G_\beta$.

(4) *Verification of sub-Markov and tightness properties.* The kernel family is sub-Markovian: for $0 \le f \le 1$ we have

$$0 \le G_\alpha f \le \alpha^{-1} \tag{7.45}$$

pointwise on $E \setminus N$, hence $U_\alpha(x, \cdot)$ are finite measures with total mass $\le \alpha^{-1}$. Regularity yields tightness on compacta (for fixed compacts K the measures $U_\alpha(x, \cdot)$ put finite mass on K uniformly for x in compacts). These properties are precisely the hypotheses required by the general potential-theoretic existence theorems which produce a Markov process with given resolvent (see next step).

(5) *Existence of a Hunt process with resolvent $\{U_\alpha\}$.* The collection of resolvent kernels $(U_\alpha)_{\alpha>0}$ (defined for $x \in E \setminus N$) is a sub-Markovian resolvent of kernels satisfying the resolvent equations and the usual continuity/tightness properties coming from regularity. By a standard theorem in the general theory of Markov processes (one version is due to Hunt and is presented in detail in Fukushima–Oshima–Takeda, Theorem 7.2.1, and in the monographs of Blumenthal–Getoor and of Sharpe), there exists a right-continuous strong Markov process $X = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \{X_t\}, \{\mathbb{P}^x\})$ with state space E (a Hunt process) whose α-resolvent is exactly U_α: for all bounded Borel f and $x \in E \setminus N$,

$$\mathbb{E}^x\Big[\int_0^\infty e^{-\alpha t} f(X_t)\, dt\Big] = \int_E f(y)\, U_\alpha(x, dy) = (G_\alpha f)^\sim(x). \tag{7.46}$$

The construction in these theorems is canonical: one builds a process on the Skorokhod path space with right-continuous paths, checking the Chapman–Kolmogorov relations and using resolvent inversion to obtain transition functions, then verifies strong Markov and sample path regularity properties; regularity of the form ensures quasi-regularity of resolvent potentials and so the process has the Hunt properties (right-continuous with left limits and quasi-left continuous).

(6) *Identification of the semigroup and generator almost everywhere.* Let P_t denote the transition semigroup of the constructed process acting on bounded measurable f by $P_t f(x) := \mathbb{E}^x[f(X_t)]$ for $x \in E \setminus N$. The resolvent identity yields

$$\int_0^\infty e^{-\alpha t} P_t f(x)\, dt = (G_\alpha f)^\sim(x) \quad \text{for } x \in E \setminus N, \tag{7.47}$$

and hence the Laplace transforms of $P_t f$ and $T_t f$ agree outside the exceptional set. Since both (T_t) and (P_t) are strongly continuous contraction semigroups on L^2 with the same resolvent G_α (the resolvent on L^2 uniquely determines the semigroup), it follows that P_t and T_t coincide as operators on L^2. More precisely, for each $f \in L^2(E, \mu)$ we have $T_t f$ equal μ-a.e. to $P_t f$;

equivalently

$$(P_t f)(x) = T_t f(x) \quad \text{for } \mu\text{-a.e. } x \in E,\ \forall t \geq 0,\ f \in L^2(E, \mu). \tag{7.48}$$

Because the resolvent kernels were constructed from quasi-continuous representatives, the equality above holds pointwise for $x \in E \setminus N$ (the complement of a properly exceptional set) and hence μ-a.e.

(7) *Right continuity, quasi-left continuity and uniqueness.* The process produced in (5) is a Hunt process (right-continuous with left limits and quasi-left continuous). Uniqueness of the process associated with the Dirichlet form holds up to properly exceptional sets: if another Hunt process yields the same resolvent outside a properly exceptional set then its finite-dimensional distributions coincide outside a properly exceptional set, hence the two processes are equivalent in the usual potential-theoretic sense. In particular the analytic object $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ determines the process uniquely up to modification on exceptional sets.

(8) *Quasi-regular case.* If $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is only quasi-regular (rather than regular) the same construction applies with the admissible state space replaced by an m-version of E and with nests and quasi-continuous representatives used in place of global continuous representatives. The general theorem of Fukushima–Oshima–Takeda (and earlier work of Ma–Röckner) yields existence of an associated right-continuous Markov process with the asserted properties; the above kernel construction and resolvent inversion remain valid in the quasi-regular setting once one works modulo properly exceptional sets and uses nests to localize.

Combining the steps above yields the desired result: every regular (or quasi-regular) Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(E, \mu)$ determines a Hunt process X whose transition semigroup (P_t) satisfies $P_t f = T_t f$ μ-a.e. for all $f \in L^2(E, \mu)$ and all $t \geq 0$; conversely the semigroup of any such process gives rise to the form $\mathcal{E}$. This completes the proof.

References. The reader may consult Fukushima–Oshima–Takeda, *Dirichlet Forms and Symmetric Markov Processes*, for the definitive account and the precise statements of the existence theorems used above (in particular Theorems 7.2.1 and 7.2.2 which treat the regular and quasi-regular cases, respectively).

Conclusion. A Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ thus constitutes the analytic manifestation of a Markov process: it encodes infinitesimal energy dissipation, generates a strongly continuous Markov semigroup (T_t), and — when regular or quasi-regular — gives rise to a canonical Hunt process (X_t) with sample paths reflecting the fine potential-theoretic structure of the underlying space E.

7.2 Capacity, Excessive Functions, and Fine Topology

In this section, we introduce and rigorously develop the fundamental analytic–probabilistic notions of *excessive functions*, *capacity*, and the associated *fine topology*, which play a central role in the potential-theoretic characterization of Dirichlet forms and the sample-path behavior of the corresponding Markov processes.

Let $(E, \mathcal{B}, \mu)$ be a σ-finite measure space, and let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a closed, symmetric, nonnegative Dirichlet form on $L^2(E, \mu)$. By the general representation theorem (Theorem 1 above), there exists a unique nonnegative self-adjoint operator L on $L^2(E, \mu)$ such that

$$\mathcal{E}(f, g) = \langle L^{1/2} f, L^{1/2} g \rangle_{L^2(E,\mu)} \qquad \forall f, g \in \mathcal{D}(\mathcal{E}). \tag{7.49}$$

By the spectral theorem for self-adjoint operators, this operator L generates a strongly continuous contraction semigroup $(T_t)_{t\geq 0}$ on $L^2(E, \mu)$, given by

$$T_t := e^{-tL}, \qquad t \geq 0. \tag{7.50}$$

Each T_t is self-adjoint, positivity-preserving, and L^2-contractive, and the family $(T_t)_{t\geq 0}$ satisfies the semigroup property $T_{s+t} = T_s T_t$ and strong continuity at $t = 0$.

Excessive Functions

A measurable function $u : E \to [0, \infty]$ is called *excessive* (with respect to $(T_t)_{t\geq 0}$) if and only if it satisfies the two fundamental conditions:

$$T_t u \leq u, \quad \forall t \geq 0, \tag{E1}$$

$$\lim_{t \downarrow 0} T_t u = u, \quad \mu\text{-a.e. on } E. \tag{E2}$$

The collection of all such functions is denoted by

$$\mathcal{E}x := \{\, u : E \to [0, \infty] \mid u \text{ measurable},\ T_t u \leq u,\ \forall t \geq 0,\ \text{and } \lim_{t \downarrow 0} T_t u = u \ \mu\text{-a.e.} \,\}. \tag{7.51}$$

Intuitively, excessive functions may be regarded as the nonnegative measurable functions which dominate their semigroup evolutes, hence they are invariant (in the limit $t \downarrow 0$) under the infinitesimal action of the generator L. From a probabilistic viewpoint, if $(X_t)_{t\geq 0}$ denotes the Hunt process associated with (T_t), then u is excessive if and only if for all $t \geq 0$,

$$\mathbb{E}_x[u(X_t)] \leq u(x), \quad \text{and } \lim_{t \downarrow 0} \mathbb{E}_x[u(X_t)] = u(x) \tag{7.52}$$

for μ-a.e. $x \in E$. Thus, excessive functions generalize the notion of superharmonic functions in classical potential theory, with the semigroup (T_t) playing the role of the resolvent of the Laplace operator.

Capacity

We next introduce the analytic notion of *capacity*, which measures the "size" of sets not in terms of measure but in terms of the energy associated with $\mathcal{E}$. Let $A \subset E$ be a Borel measurable set. The *capacity* of A relative to $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is defined as

$$\operatorname{Cap}(A) := \inf \{\, \mathcal{E}(f, f) + \|f\|^2_{L^2(E,\mu)} \ :\ f \in \mathcal{D}(\mathcal{E}),\ f \geq 1 \ \mu\text{-a.e. on } A \,\}. \tag{7.53}$$

If no such f exists (for instance, if A is not contained in the support of μ), we define $\mathrm{Cap}(A) := +\infty$.

Interpretation: The function f in the infimum above may be interpreted as an "energy-minimizing potential" that raises the potential to at least unit level on A while minimizing the total Dirichlet energy $\mathcal{E}(f, f) + \|f\|_2^2$. Hence, $\mathrm{Cap}(A)$ quantifies the energetic "cost" of forcing the field f to dominate 1_A. This notion of capacity is central in distinguishing sets that are negligible from the point of view of the process (X_t), rather than in the sense of the measure μ.

Quasi-everywhere (q.e.) properties. A property $P(x)$ is said to hold *quasi-everywhere* (abbreviated *q.e.*) if the set of points where it fails,

$$N := \{x \in E : P(x) \text{ fails}\}, \tag{7.54}$$

has capacity zero, i.e. $\mathrm{Cap}(N) = 0$. Since Cap is typically finer than μ, a property holding q.e. need not hold μ-a.e., but it will hold except on sets that are "polar" for the process (X_t), meaning that the process almost surely never visits them. Thus, capacity-zero sets play the role of *polar* or *negligible* sets in the potential-theoretic sense.

Fine Topology

We now define the *fine topology* on E, which provides the natural topological framework for studying the path-continuity of the Markov process (X_t) and the quasi-continuity of Dirichlet form elements.

The *fine topology* τ_f on E is the coarsest topology such that every function $u \in \mathcal{E}x$ is continuous with respect to τ_f. Formally,

$$\tau_f := \text{the smallest topology on } E \text{ such that all } u \in \mathcal{E}x \text{ are } \tau_f\text{-continuous.} \tag{7.55}$$

Equivalently, τ_f is the initial topology induced by the family of excessive functions:

$$\tau_f = \sigma(\mathcal{E}x) := \text{coarsest topology making } u \mapsto u(x) \text{ continuous for all } u \in \mathcal{E}x. \tag{7.56}$$

Properties of the Fine Topology.

1. τ_f is strictly finer than the topology induced by μ-almost everywhere convergence, i.e., it distinguishes points that are indistinguishable under μ.

2. Every $f \in \mathcal{D}(\mathcal{E})$ admits a *quasi-continuous* modification $\tilde{f}$, meaning that there exists a set N with $\mathrm{Cap}(N) = 0$ such that $\tilde{f}$ is τ_f-continuous on $E \setminus N$.

3. The fine topology coincides with the topology of the potential theory associated with (X_t): it is the smallest topology making the sample paths $t \mapsto X_t(\omega)$ finely continuous almost surely.

Hence, the fine topology bridges the analytic world of the Dirichlet form and the probabilistic world of the associated Markov process, permitting one to interpret notions such as "continuity of paths" or "support of a function" in a quasi-sure (rather than measure-theoretic) sense.

Summary. In summary, the interplay between these three objects—excessive functions, capacity, and fine topology—provides a complete potential-theoretic description of the analytic structure of Dirichlet forms. The excessive functions describe the invariant potentials under the semigroup; capacity quantifies the energetic size of sets; and the fine topology gives the natural topological framework under which quasi-continuity and potential-theoretic regularity can be rigorously defined.

Analytic vs. probabilistic excessivity; quasi-continuous modifications

We work under the standing hypotheses of this chapter: $(E, \mathcal{B}, \mu)$ is a locally compact separable metric space equipped with a Radon measure μ, and $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a regular (or quasi-regular, with the obvious modifications) Dirichlet form on $L^2(E, \mu)$. Let L be the self-adjoint non-negative operator associated with $\mathcal{E}$ and let $(T_t)_{t\geq 0} = (e^{-tL})_{t\geq 0}$ denote the corresponding L^2-semigroup. Let $X = (\Omega, \mathcal{F}, (\mathcal{F}_t), (X_t), (\mathbb{P}^x)_{x\in E})$ be the Hunt process associated with $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ whose construction and basic properties were given above; write $\mathbb{E}^x$ for expectation under $\mathbb{P}^x$. All statements that follow are to be interpreted up to properly exceptional (capacity-zero) sets unless otherwise indicated.

Theorem 7.1 (Equivalence of analytic and probabilistic excessive functions)**.** *Let $u : E \to [0, \infty]$ be measurable. The following are equivalent:*

1. *u is* analytic-excessive*: $T_t u \leq u$ for every $t \geq 0$ and $\lim_{t\downarrow 0} T_t u = u$ μ-a.e.;*
2. *u is* probabilistic-excessive*: for μ-a.e. (indeed for q.e.) $x \in E$ and every $t \geq 0$,*

$$\mathbb{E}^x[u(X_t)] \leq u(x) \quad \textit{and} \quad \lim_{t\downarrow 0} \mathbb{E}^x[u(X_t)] = u(x). \tag{7.57}$$

Proof. We prove (i) ⇒ (ii) and (ii) ⇒ (i), indicating where exceptional sets may be introduced.

(i) ⇒ (ii). Assume u satisfies the analytic definition. By regularity/quasi-regularity and the construction of the Hunt process, there exists a properly exceptional set N (cap-null) outside which the resolvent and semigroup admit quasi-continuous representatives and the process X is properly associated with the semigroup in the sense that for every bounded Borel f and $x \in E \setminus N$,

$$\mathbb{E}^x[f(X_t)] = (T_t f)^{\sim}(x), \tag{7.58}$$

where $(T_t f)^{\sim}$ denotes the quasi-continuous version. Let $u_n := u \wedge n$ be the bounded truncations. For each fixed n we have

$$T_t u_n \leq T_t u \leq u \tag{7.59}$$

and $T_t u_n \uparrow T_t u$ as $n \to \infty$ by monotone convergence in L^2 and then quasi-everywhere for the quasi-continuous versions. For $x \in E \setminus N$ and each n,

$$\mathbb{E}^x[u_n(X_t)] = (T_t u_n)^{\sim}(x) \leq u^{\sim}(x), \tag{7.60}$$

and by monotone convergence in n we obtain

$$\mathbb{E}^x[u(X_t)] \leq u^{\sim}(x). \tag{7.61}$$

Since $u^\sim = u$ μ-a.e. and the exceptional set may be enlarged to a cap-null set independent of t and n, we conclude that for q.e. x and every $t \geq 0$,

$$\mathbb{E}^x[u(X_t)] \leq u(x). \tag{7.62}$$

The right-continuity assertion follows by the strong continuity of T_t at 0 in L^2 combined with the quasi-continuous representatives: for bounded truncations u_n we have

$$\lim_{t \downarrow 0} T_t u_n = u_n \tag{7.63}$$

in L^2 and hence quasi-everywhere along a subsequence; monotone convergence then yields

$$\lim_{t \downarrow 0} \mathbb{E}^x[u(X_t)] = u(x) \tag{7.64}$$

for q.e. x.

(ii) $\Rightarrow$ (i). Conversely, suppose the probabilistic inequalities hold for q.e. x. For bounded Borel f and $t \geq 0$, the identification of semigroup and transition expectation outside a cap-null set gives

$$T_t f(x) = \mathbb{E}^x[f(X_t)] \tag{7.65}$$

for q.e. x. Applying this with $f = u_n = u \wedge n$ and using monotone convergence as $n \to \infty$, we obtain $T_t u \leq u$ q.e.; the L^2-a.e. statement follows because cap-null sets are μ-null (by regularity). The continuity at zero is identical to the previous direction: since

$$\lim_{t \downarrow 0} \mathbb{E}^x[u(X_t)] = u(x) \tag{7.66}$$

for q.e. x, we have $\lim_{t \downarrow 0} T_t u = u$ μ-a.e., finishing the proof. □

Theorem 7.2 (Existence and uniqueness of quasi-continuous modification)**.** *Let $f \in \mathcal{D}(\mathcal{E})$. Then there exists a quasi-continuous function $\tilde{f}$ (i.e. continuous on $E \setminus N$ for some cap-null set N) such that $\tilde{f} = f$ μ-a.e. Moreover such a quasi-continuous modification is unique up to modification on a set of capacity zero.*

Proof. We give a complete constructive proof based on approximation by the regular core $C_c(E) \cap \mathcal{D}(\mathcal{E})$ and an elementary capacity estimate.

Step 1: density by core. Regularity implies that $C_c(E) \cap \mathcal{D}(\mathcal{E})$ is dense in $\mathcal{D}(\mathcal{E})$ with respect to the $\mathcal{E}$-norm $\|\cdot\|_{\mathcal{E}}$. Hence there exists a sequence

$$(\varphi_n)_{n \geq 1} \subset C_c(E) \cap \mathcal{D}(\mathcal{E}) \tag{7.67}$$

such that

$$\lim_{n \to \infty} \|\varphi_n - f\|_{\mathcal{E}} = 0, \tag{7.68}$$

i.e.

$$\lim_{n \to \infty} \left(\|\varphi_n - f\|_2^2 + \mathcal{E}(\varphi_n - f, \varphi_n - f) \right) = 0. \tag{7.69}$$

Step 2: capacity estimate for level sets. For any $v \in \mathcal{D}(\mathcal{E})$ and $\varepsilon > 0$ the Chebyshev-type capacity bound (standard in Dirichlet form theory) holds:

$$\mathrm{Cap}\big(\{x \in E : |v(x)| > \varepsilon\}\big) \le \frac{1}{\varepsilon^2}\big(\mathcal{E}(v,v) + \|v\|_2^2\big). \tag{$*$}$$

This inequality is obtained by observing that the trial function $w := (v/\varepsilon) \wedge 1 \in \mathcal{D}(\mathcal{E})$ (valid because the form is Markovian and closed) satisfies $w \ge 1$ on $\{|v| > \varepsilon\}$, whence the definition (7.53) of capacity yields the displayed bound.

Step 3: Cauchy subsequence with quasi-uniform control. From $\|\varphi_n - f\|_{\mathcal{E}} \to 0$ we have in particular $\|\varphi_n - f\|_{\mathcal{E}} \to 0$. Using $(*)$ with $v = \varphi_n - \varphi_m$ gives

$$\mathrm{Cap}\big(\{|\varphi_n - \varphi_m| > \varepsilon\}\big) \le \varepsilon^{-2}\big(\mathcal{E}(\varphi_n - \varphi_m, \varphi_n - \varphi_m) + \|\varphi_n - \varphi_m\|_2^2\big). \tag{7.70}$$

Since (φ_n) is Cauchy in the $\mathcal{E}$-norm, the right-hand side tends to 0 as $n, m \to \infty$. Hence for each fixed $\varepsilon > 0$,

$$\lim_{m,n\to\infty} \mathrm{Cap}\big(\{|\varphi_n - \varphi_m| > \varepsilon\}\big) = 0. \tag{7.71}$$

By a diagonal argument choose indices n_k so that for each integer $j \ge 1$,

$$\mathrm{Cap}\Big(\bigcup_{k,\ell \ge j} \{|\varphi_{n_k} - \varphi_{n_\ell}| > 2^{-j}\}\Big) \le 2^{-j}. \tag{7.72}$$

Define the exceptional set

$$N := \bigcap_{j\ge1} \bigcup_{k,\ell\ge j} \{|\varphi_{n_k} - \varphi_{n_\ell}| > 2^{-j}\}. \tag{7.73}$$

Then

$$\mathrm{Cap}(N) \le \inf_j 2^{-j} = 0 \tag{7.74}$$

so N is cap-null. On $E \setminus N$ the sequence (φ_{n_k}) is Cauchy uniformly: for each $x \in E \setminus N$ and each j there exists K such that for all $k, \ell \ge K$

$$|\varphi_{n_k}(x) - \varphi_{n_\ell}(x)| \le 2^{-j} \tag{7.75}$$

Hence $\varphi_{n_k}(x)$ converges uniformly on compacta of $E \setminus N$ to a limit, which we denote $\tilde{f}(x)$.

Step 4: definition and properties of $\tilde{f}$. Define $\tilde{f}$ on $E \setminus N$ by pointwise limit

$$\tilde{f}(x) := \lim_{k\to\infty} \varphi_{n_k}(x), \qquad x \in E \setminus N. \tag{7.76}$$

Since each φ_{n_k} is continuous on E and the convergence is uniform on compacts of $E \setminus N$, the limit $\tilde{f}$ is continuous on $E \setminus N$. By construction $\tilde{f}$ coincides μ-a.e. with f: indeed the approximating sequence (φ_{n_k}) converges to f in L^2, hence has a subsequence converging μ-a.e.; uniqueness of pointwise limits on the complement of a μ-null set implies $\tilde{f} = f$ μ-a.e. on $E \setminus N$, and enlarging N by a μ-null set if necessary gives $\tilde{f} = f$ μ-a.e. on E.

Step 5: independence and uniqueness up to cap-null sets. If $\tilde{f}_1, \tilde{f}_2$ are two quasi-continuous modifications of f, then the set $\{x : \tilde{f}_1(x) \neq \tilde{f}_2(x)\}$ is finely open (difference of finely continuous functions) and μ-null; by a standard capacity estimate such a set must have capacity zero. Hence quasi-continuous modification is unique up to a cap-null set.

This constructs a quasi-continuous representative $\tilde{f}$ of any $f \in \mathcal{D}(\mathcal{E})$ and proves uniqueness up to capacity-zero sets, completing the proof. □

7.3 Quasi-Continuity, Revuz Correspondence and Smooth Measures

We now present the deep connection between analytic objects associated to a Dirichlet form and probabilistic objects associated to the corresponding Markov process, focusing on quasi-continuous functions, smooth measures, and the celebrated *Revuz correspondence*. We work under the standard assumptions of this chapter: $(E, \mathcal{B}, \mu)$ is a locally compact separable metric space endowed with a Radon measure μ, and $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a *quasi-regular* Dirichlet form on $L^2(E, \mu)$. Let $X = (X_t)_{t \geq 0}$ denote the associated Hunt process, and let $(T_t)_{t \geq 0}$ be its L^2-semigroup with L as the infinitesimal generator.

1. Quasi-continuity of elements in $\mathcal{D}(\mathcal{E})$.

For any $f \in \mathcal{D}(\mathcal{E})$, there exists a *quasi-continuous modification* $\tilde{f}$ such that:

$$\tilde{f} = f \quad \mu\text{-a.e.}, \qquad \tilde{f} \text{ is continuous on } E \setminus N, \tag{7.77}$$

for some *exceptional set* $N \subset E$ with $\mathrm{Cap}(N) = 0$. The notion of capacity used here is the Dirichlet capacity

$$\mathrm{Cap}(A) := \inf\{\mathcal{E}(u, u) + \|u\|_2^2 : u \in \mathcal{D}(\mathcal{E}),\ u \geq 1 \text{ on } A\}. \tag{7.78}$$

A property is said to hold *quasi-everywhere (q.e.)* if it holds outside a cap-null set. The quasi-continuous representative is unique up to such sets.

Quasi-continuity is essential because sample paths of the associated Hunt process X_t are quasi-continuous in the fine topology, ensuring that evaluations such as $\tilde{f}(X_t)$ are well-defined along almost all sample paths.

2. Smooth measures.

A σ-finite measure ν on $(E, \mathcal{B})$ is called a *smooth measure* (with respect to $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$) if it satisfies both of the following conditions:

1. ν charges no set of capacity zero: if $\mathrm{Cap}(A) = 0$, then $\nu(A) = 0$;
2. there exists a *nest* $(F_n)_{n \geq 1}$ of closed sets such that

$$\mathrm{Cap}(E \setminus F_n) = 0, \quad \text{and} \quad \nu(F_n) < \infty \quad \forall n \geq 1. \tag{7.79}$$

Intuitively, ν is "smooth" in the sense that it interacts well with the fine topology and is concentrated on sets that the process X_t visits in a controlled (finite-measure) manner.

A nest (F_n) is a sequence of closed subsets of E with $F_n \subset F_{n+1}$ and $\bigcup_n F_n$ dense in E such that for every $f \in \mathcal{D}(\mathcal{E})$, there exists n_0 with $f|_{F_n}$ continuous for all $n \geq n_0$.

Smooth measures generalize measures absolutely continuous with respect to μ, but include many singular measures supported on sets of small (even zero Lebesgue) measure yet positive capacity — for example, surface measures on manifolds or boundaries.

3. Additive functionals (AF) and positive continuous additive functionals (PCAF).

Let $X = (X_t)_{t\geq 0}$ be the Hunt process associated to the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$. A family of random variables $A = (A_t)_{t\geq 0}$ is called an *additive functional (AF)* if it satisfies:

$$A_0 = 0, \quad A_{t+s} = A_t + A_s \circ \theta_t, \quad \text{for all } t, s \geq 0, \tag{7.80}$$

where (θ_t) is the shift operator on path space. If, in addition, the following hold:

- A_t is *adapted* and almost surely right-continuous with left limits;
- A_t is *nondecreasing* and *continuous* in t;
- $\mathbb{E}^x[A_t] < \infty$ for all $x \in E$ and all $t > 0$;

then A_t is called a *positive continuous additive functional (PCAF)*.

Examples of PCAFs include:

- The local time L_t^a of a diffusion at a point a;
- The occupation time of a measurable set B:

$$A_t = \int_0^t \mathbf{1}_B(X_s)\, ds; \tag{7.81}$$

- Additive functionals defined via continuous potentials, such as $A_t = \int_0^t f(X_s)\, ds$ for bounded measurable f.

4. The Revuz correspondence.

The *Revuz correspondence* provides a one-to-one correspondence between PCAFs A_t of the Markov process X_t and smooth measures ν on E. Specifically:

Theorem 7.3 (Revuz correspondence)**.** *Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a quasi-regular Dirichlet form on $L^2(E, \mu)$, and let X_t be its associated Hunt process. Then there exists a one-to-one correspondence*

$$A_t \longleftrightarrow \nu \tag{7.82}$$

between equivalence classes of PCAFs A_t (up to indistinguishability) and smooth measures ν, characterized by the identity

$$\int_E f(x)\,\mathbb{E}^x\left[\int_0^\infty e^{-t} g(X_t)\,dA_t\right]\,\mu(dx) = \int_E f(x)\int_E g(y)\,G_1(x,y)\,\nu(dy)\,\mu(dx), \tag{7.83}$$

for all nonnegative measurable f, g. Here $G_1(x,y)$ denotes the 1-resolvent kernel associated to the semigroup (T_t):

$$G_1 f(x) = \int_0^\infty e^{-t}(T_t f)(x)\,dt = \int_E G_1(x,y) f(y)\,\mu(dy). \tag{7.84}$$

Equivalently, for every $x \in E$,

$$\mathbb{E}^x\left[\int_0^\infty e^{-t} dA_t\right] = \int_E G_1(x,y)\,\nu(dy). \tag{7.85}$$

Proof. We begin by recalling the basic framework under which the Revuz correspondence is established. Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a quasi-regular Dirichlet form on $L^2(E,\mu)$, and let T_t denote the associated L^2-semigroup. Quasi-regularity guarantees the existence of a right-continuous Hunt process X_t with state space E and transition semigroup P_t satisfying $P_t f = T_t f$ μ-a.e. for all $f \in L^2(E,\mu)$.

Step 1: Definition of the 1-potential of a PCAF.

Let A_t be a positive continuous additive functional (PCAF) of the Hunt process (X_t). For $f \in L^2(E,\mu)$, define

$$U_A f(x) := \mathbb{E}^x\left[\int_0^\infty e^{-t} f(X_t)\,dA_t\right]. \tag{7.86}$$

We claim that $U_A f$ is well-defined, finite for quasi-every x, and measurable. By the monotone convergence theorem, $U_A f(x)$ is measurable since $t \mapsto e^{-t} f(X_t)$ is jointly measurable in (t,x) and A_t has continuous increasing paths. Moreover, the integrability follows from the boundedness of f and the sub-Markovian property of T_t, so that

$$\int_0^\infty e^{-t}|f(X_t)|\,dA_t < \infty \quad \text{a.s. for quasi-every } x. \tag{7.87}$$

The map $f \mapsto U_A f$ is linear and positive. Furthermore, it defines a continuous operator from $L^2(E,\mu)$ to itself because of the estimate

$$\|U_A f\|_2^2 \le C_A \|f\|_2^2, \tag{7.88}$$

for some finite constant C_A depending on A_t, as a consequence of the energy inequality for Dirichlet forms.

Step 2: Existence of a measure ν representing A_t.

By the Riesz representation theorem in $L^2(E,\mu)$, there exists a unique finite nonnegative measure ν on E such that for all nonnegative $f, g \in L^2(E,\mu)$,

$$\int_E f(x)\,U_A g(x)\,\mu(dx) = \int_E f(x)\int_E g(y)\,G_1(x,y)\,\nu(dy)\,\mu(dx). \tag{7.89}$$

We must now show that this measure ν is *smooth*. For this, note that the map $f \mapsto U_A f$ satisfies the resolvent equation:

$$U_A f = G_1(f \cdot \nu), \tag{7.90}$$

where G_1 denotes the 1-resolvent operator associated to (T_t). Since A_t increases only when the process is inside a region of positive capacity (because A_t is a PCAF), it follows that ν charges no set of zero capacity. Hence, ν is smooth in the sense of potential theory: for every Borel set $B \subset E$,

$$\text{Cap}(B) = 0 \;\Rightarrow\; \nu(B) = 0. \tag{7.91}$$

Moreover, there exists an increasing sequence of compact sets $F_n \subset E$ (a *nest*) such that $\nu(F_n) < \infty$ and $\text{Cap}(E \setminus F_n) = 0$, ensuring the σ-finiteness of ν.

Step 3: Characterization of the correspondence.

Given such a measure ν, define the PCAF A_t^ν by the additive functional satisfying, for all nonnegative measurable f, g,

$$\int_E f(x)\,\mathbb{E}^x\left[\int_0^\infty e^{-t} g(X_t)\,dA_t^\nu\right]\,\mu(dx) = \int_E f(x)\int_E g(y)\,G_1(x,y)\,\nu(dy)\,\mu(dx). \tag{7.92}$$

The left-hand side defines a bilinear form in (f, g), and the right-hand side depends only on ν. By the strong Markov property of (X_t) and Fubini's theorem, one can verify that A_t^ν is a PCAF of (X_t).

Uniqueness follows from the fact that if A_t^1 and A_t^2 correspond to the same ν, then

$$\int_E f(x)\,\mathbb{E}^x\left[\int_0^\infty e^{-t} g(X_t)\,d(A_t^1 - A_t^2)\right]\,\mu(dx) = 0, \tag{7.93}$$

for all nonnegative f, g, which implies $A_t^1 = A_t^2$ up to indistinguishability by the uniqueness of additive functionals in the sense of potential theory.

Step 4: The correspondence is bijective.

Finally, the two constructions above are inverse to each other:

- Given A_t, the measure ν defined via the integral identity above is smooth and corresponds uniquely to A_t.
- Conversely, given ν, the process A_t^ν constructed as above is a PCAF corresponding to ν.

Hence, we obtain a one-to-one correspondence between the equivalence classes of PCAFs and smooth measures.

Step 5: Simplified representation.

Taking $f = g \equiv 1$ in the defining relation yields the pointwise version

$$\mathbb{E}^x\left[\int_0^\infty e^{-t} dA_t\right] = \int_E G_1(x,y)\,\nu(dy), \tag{7.94}$$

which expresses the expected discounted occupation time of the process, weighted by A_t, as the potential of the measure ν.

Conclusion. The proof shows that for every PCAF A_t of the Hunt process X_t, there exists a unique smooth measure ν satisfying the integral identity above, and conversely. This establishes the desired bijective correspondence, completing the proof. □

In Summary, the proof proceeds as follows:

(i) For any PCAF A_t, define the functional

$$\Phi(f,g) = \int_E f(x)\,\mathbb{E}^x\left[\int_0^\infty e^{-t} g(X_t)\,dA_t\right]\,\mu(dx). \tag{7.95}$$

(ii) By Fubini's theorem and semigroup properties, Φ is bilinear, positive, and continuous on $L^2(E,\mu) \times L^2(E,\mu)$.

(iii) Using Riesz representation and the structure of potential operators G_1, one identifies a unique σ-finite measure ν such that

$$\Phi(f,g) = \int_E f(x) \int_E G_1(x,y) g(y)\,\nu(dy)\,\mu(dx). \tag{7.96}$$

(iv) Conversely, given any smooth measure ν, one constructs a PCAF A_t such that the above identity holds, using the potential $U_1^\nu(x) := \int_E G_1(x,y)\nu(dy)$ and the fine continuity of paths.

(v) Uniqueness follows from the strong Markov property and the strict positivity of $G_1(x,y)$.

Thus, the Revuz correspondence provides a bijection

$$\{\text{PCAFs of } X_t\}/\text{equivalence} \quad \leftrightarrow \quad \{\text{smooth measures on } E\}. \tag{7.97}$$

It translates analytic data (measures) into probabilistic objects (additive functionals) and vice versa, forming the foundation of modern potential theory and stochastic calculus for Markov processes.

5. Consequences and interpretations. The Revuz correspondence shows that:

- Each smooth measure ν defines a unique "occupation functional" A_t along the paths of X_t;
- The potential operator G_1 serves as the Green's kernel, linking analytic potentials $G_1\nu$ to expected discounted additive functionals;
- Quasi-continuity ensures the compatibility of analytic and probabilistic definitions, so that pathwise integrals such as $\int_0^t \tilde{f}(X_s)\,dA_s$ are well-defined for all quasi-continuous $\tilde{f}$.

In summary, the Revuz correspondence establishes a fundamental duality between the analytic framework of Dirichlet forms and the probabilistic behavior of associated Hunt processes, providing a powerful bridge connecting measure-theoretic potential theory and stochastic process theory.

7.4 Beurling–Deny Decomposition and Nonlocal Forms

The Beurling–Deny decomposition constitutes one of the most profound structural results in the theory of regular Dirichlet forms. It establishes that every regular, symmetric Dirichlet form on $L^2(E,\mu)$ can be uniquely expressed as a sum of a strongly local (diffusion) component, a jump (nonlocal) component, and a killing (absorbing) component. This decomposition provides a rigorous analytic and probabilistic description of all symmetric Markov processes combining continuous motion, discontinuous jumps, and potential killing.

1. Preliminaries and Setting.

Let $(E,\mathcal{B},\mu)$ be a locally compact separable metric space equipped with a positive Radon measure μ having full support. Let $(\mathcal{E},\mathcal{D}(\mathcal{E}))$ be a regular symmetric Dirichlet form on $L^2(E,\mu)$. That is:

1. $\mathcal{E}$ is a closed, symmetric, bilinear form on $L^2(E,\mu)$;
2. the Markov property holds: for all $f \in \mathcal{D}(\mathcal{E})$,

$$\tilde{f} := (0 \vee f) \wedge 1 \in \mathcal{D}(\mathcal{E}) \quad \text{and} \quad \mathcal{E}(\tilde{f},\tilde{f}) \le \mathcal{E}(f,f); \tag{7.98}$$

3. $\mathcal{D}(\mathcal{E}) \cap C_c(E)$ is dense in both $\mathcal{D}(\mathcal{E})$ (with the norm $\|f\|_{\mathcal{E}_1} := (\mathcal{E}(f,f) + \|f\|_2^2)^{1/2}$) and $C_c(E)$ (with the uniform norm).

By Fukushima's correspondence, such a Dirichlet form uniquely determines a symmetric Hunt process $X = (X_t)_{t\ge 0}$ on E with transition semigroup (T_t) and generator L. The generator L is a densely defined, nonpositive self-adjoint operator on $L^2(E,\mu)$ satisfying

$$\mathcal{E}(f,g) = (-Lf,g)_{L^2(E,\mu)}, \qquad f \in \mathcal{D}(L),\ g \in \mathcal{D}(\mathcal{E}).$$

2. Strongly local, jump, and killing components.

Let us recall that a symmetric Dirichlet form $\mathcal{E}$ is called:

- *strongly local* if $\mathcal{E}(f,g) = 0$ whenever f is constant on a neighborhood of the support of g;
- *purely nonlocal* if $\mathcal{E}(f,g)$ depends only on pairwise differences $f(x) - f(y)$;
- *with killing* if the process can be terminated at a random lifetime.

In general, a regular Dirichlet form may have all three features simultaneously. Beurling and Deny showed that these parts can be uniquely separated.

3. Statement of the Beurling–Deny decomposition.

There exist measurable objects $(a(x))$, $(c(x))$, $(J(x,dy))$ with the following properties:

1. $a(x)$ represents the *diffusion coefficient*, corresponding to the energy measure of the strongly local part;
2. $c(x)$ represents the *killing rate*, a nonnegative measurable function on E;

3. $J(x, dy)$ is a *symmetric jump kernel*, that is, a positive measure on $E \times E \setminus \text{diag}$ satisfying

$$J(dx, dy) = J(dy, dx), \tag{7.99}$$

   and for each x, $J(x, dy)$ is a σ-finite measure on $E \setminus \{x\}$.

Then, for all $f \in \mathcal{D}(\mathcal{E})$, we have the decomposition:

$$\mathcal{E}(f, f) = \underbrace{\int_E a(x)\, |\nabla f(x)|^2\, d\mu(x)}_{\text{strongly local (diffusion) part}} + \underbrace{\int_E f(x)^2\, c(x)\, d\mu(x)}_{\text{killing part}} + \underbrace{\frac{1}{2}\int_E\int_E (f(x) - f(y))^2\, J(dx, dy)}_{\text{jump (nonlocal) part}}. \tag{7.100}$$

This is known as the *Beurling–Deny decomposition.*

The factor $\frac{1}{2}$ ensures that symmetric contributions from both (x, y) and (y, x) are not double-counted since $J(dx, dy)$ is symmetric.

4. Analytic meaning of each term.

Each term of the decomposition corresponds to a distinct analytic and probabilistic mechanism:

1. **Diffusion (strongly local) part:**

$$\mathcal{E}^{(c)}(f, f) := \int_E a(x)\, |\nabla f(x)|^2\, d\mu(x). \tag{7.101}$$

   This term describes local energy dissipation and corresponds to the continuous part of the process, i.e., its Brownian-like motion. The matrix $a(x)$ (if $E \subseteq \mathbb{R}^d$) represents the diffusion tensor and satisfies $a(x) \geq 0$ (positive semidefinite) for μ-a.e. x.

2. **Killing part:**

$$\mathcal{E}^{(k)}(f, f) := \int_E f(x)^2\, c(x)\, d\mu(x), \tag{7.102}$$

   where $c(x) \geq 0$ is measurable. This term accounts for possible "absorption" or "killing" of the process at rate $c(x)$; probabilistically, it corresponds to an exponential lifetime mechanism independent of spatial motion.

3. **Jump (nonlocal) part:**

$$\mathcal{E}^{(j)}(f, f) := \frac{1}{2}\int_E\int_E \left(f(x) - f(y)\right)^2 J(dx, dy). \tag{7.103}$$

   This term quantifies the total quadratic energy contributed by jumps between points x and y. The kernel $J(dx, dy)$ measures the frequency and intensity of such jumps. It generalizes Lévy-type nonlocal operators and yields the generator

$$L^{(j)} f(x) = \int_E (f(y) - f(x))\, J(x, dy), \tag{7.104}$$

   under suitable integrability conditions (typically requiring $\int_E (1 \wedge |x - y|^2)\, J(x, dy) < \infty$).

Thus, we can write

$$\mathcal{E} = \mathcal{E}^{(c)} + \mathcal{E}^{(j)} + \mathcal{E}^{(k)}, \tag{7.105}$$

and these components are uniquely determined by $\mathcal{E}$.

5. Connection with Lévy–Khintchine representation.

In the special case $E = \mathbb{R}^d$ and μ the Lebesgue measure, the generator L corresponding to the Beurling–Deny decomposition takes the general *Lévy–Khintchine form*:

$$Lf(x) = \sum_{i,j=1}^{d} \partial_i\big(a_{ij}(x)\,\partial_j f(x)\big) + \int_{\mathbb{R}^d}\Big(f(x+y) - f(x) - \nabla f(x)\cdot y\,\mathbf{1}_{|y|<1}\Big)\,n(x,dy) - c(x)f(x), \tag{7.106}$$

where $n(x, dy)$ is a measurable Lévy kernel corresponding to $J(dx, dy)$. This expression unifies diffusion, jump, and killing mechanisms within a single operator framework.

6. Probabilistic interpretation.

Under the correspondence between symmetric Dirichlet forms and symmetric Hunt processes, the three terms of the Beurling–Deny decomposition have precise probabilistic meanings:

- The diffusion part corresponds to the continuous martingale component of the process (e.g. Brownian motion);
- The jump part corresponds to the pure jump component of the process, where transitions occur discontinuously according to the kernel J;
- The killing part corresponds to an independent exponential killing time with spatial rate $c(x)$, at which the process terminates.

Hence, every symmetric Hunt process with these three mechanisms corresponds exactly to one Dirichlet form of the above type, and vice versa.

7. Rigorous structural characterization.

The Beurling–Deny decomposition yields the following classification theorem.

Theorem 7.4 (Beurling–Deny representation theorem)**.** *Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular symmetric Dirichlet form on $L^2(E,\mu)$. Then there exist uniquely determined measures ν_c, ν_j, and ν_k on $E \times E$ and E, respectively, such that:*

$$\mathcal{E}(f,g) = \int_E d\nu_c(\Gamma(f,g)) + \int_{E\times E}(f(x)-f(y))(g(x)-g(y))\,\nu_j(dx,dy) + \int_E f(x)g(x)\,\nu_k(dx), \tag{7.107}$$

for all $f, g \in \mathcal{D}(\mathcal{E})$, where $\Gamma(f,g)$ denotes the energy measure associated with the local part.

Moreover, these three components are mutually orthogonal in the sense that they interact only through their respective subspaces of functions (strongly local, jump, or killing).

8. Implications for spectral theory and potential analysis.

The decomposition directly influences:

- the structure of the generator L, which becomes a sum $L = L^{(c)} + L^{(j)} - c(x)$;

- the spectral representation of L, with local and nonlocal components contributing distinct spectral branches;
- the potential kernel $G_\alpha = (\alpha I - L)^{-1}$, whose integral kernel decomposes into diffusion, jump, and killing contributions, crucial for potential-theoretic and probabilistic estimates.

Therefore, the Beurling–Deny decomposition not only classifies all possible symmetric Markovian behaviors but also provides a rigorous analytic tool to study their spectral, probabilistic, and potential-theoretic properties.

7.5 Applications to Feynman–Kac Semigroups and Schrödinger Operators

We now present the precise analytic and probabilistic interrelation between Dirichlet forms and Feynman–Kac semigroups associated with Schrödinger-type operators of the form $\mathcal{L} - V$, where $\mathcal{L}$ is the generator of a symmetric Dirichlet form and V is a measurable potential. This connection not only provides a rigorous mathematical framework for the analysis of quantum mechanical semigroups but also links probabilistic and potential-theoretic concepts such as capacity, quasi-continuity, and excessive functions.

1. Analytic formulation via perturbation of Dirichlet forms.
Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a symmetric, closed, nonnegative Dirichlet form on $L^2(E,\mu)$ with associated self-adjoint operator $\mathcal{L}$ (the generator). Let $V : E \to \mathbb{R}$ be a measurable potential function, bounded from below and locally integrable. The *perturbed form* (also called the Schrödinger form) is defined by

$$\mathcal{E}^V(f,g) := \mathcal{E}(f,g) + \int_E V(x)f(x)g(x)\,d\mu(x), \quad f,g \in \mathcal{D}(\mathcal{E}^V) := \mathcal{D}(\mathcal{E}) \cap L^2(|V|\,d\mu). \tag{7.108}$$

Since V is bounded below, say $V(x) \geq -C$, the form $\mathcal{E}^V$ is lower semi-bounded:

$$\mathcal{E}^V(f,f) + (C+1)\|f\|_2^2 \geq \|f\|_2^2, \quad \forall f \in \mathcal{D}(\mathcal{E}^V), \tag{7.109}$$

ensuring that it is closed and densely defined on $L^2(E,\mu)$. Therefore, by the general representation theorem for closed symmetric forms, there exists a unique self-adjoint operator H on $L^2(E,\mu)$ such that

$$\mathcal{E}^V(f,g) = \langle H^{1/2}f, H^{1/2}g\rangle_{L^2(E,\mu)}. \tag{7.110}$$

We interpret H as the *Schrödinger operator* formally written as

$$H = -(\mathcal{L} - V). \tag{7.111}$$

The operator H is nonnegative and self-adjoint, hence generates a strongly continuous, symmetric contraction semigroup $(e^{-tH})_{t\geq 0}$ on $L^2(E,\mu)$.

2. Probabilistic representation via the Feynman–Kac formula.
Let X_t be the Hunt process associated with the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$, so that its transition

semigroup $(T_t)_{t\geq 0}$ satisfies

$$T_t f(x) = \mathbb{E}^x[f(X_t)] \tag{7.112}$$

for all $f \in L^2(E, \mu)$. For a potential $V : E \to \mathbb{R}_+$, define the Feynman–Kac semigroup $(P_t^V)_{t\geq 0}$ by

$$(P_t^V f)(x) := \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right], \quad f \in L^2(E, \mu). \tag{7.113}$$

This family (P_t^V) forms a strongly continuous, self-adjoint, sub-Markov semigroup on $L^2(E, \mu)$. Moreover, the Feynman–Kac formula yields the rigorous identity

$$P_t^V = e^{-tH}, \tag{7.114}$$

where

$$H = -(\mathcal{L} - V) \tag{7.115}$$

is the Schrödinger operator defined above. The proof follows by applying the Trotter product formula:

$$e^{-t(\mathcal{L}-V)} = \lim_{n\to\infty} \left(e^{-(t/n)\mathcal{L}} e^{(t/n)V}\right)^n, \tag{7.116}$$

together with the pathwise representation of $e^{-(t/n)\mathcal{L}}$ as conditional expectation under the process X_t. The resulting stochastic exponential $\exp\left(-\int_0^t V(X_s)\,ds\right)$ provides the multiplicative functional implementing the potential perturbation in the Dirichlet form framework.

3. Application of potential-theoretic tools.
The entire potential-theoretic structure associated with the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ extends naturally to the Schrödinger form $(\mathcal{E}^V, \mathcal{D}(\mathcal{E}^V))$. In particular:

1. **Capacity and quasi-continuity.** For every Borel set $A \subset E$, the $\mathcal{E}^V$-capacity is defined by

$$\mathrm{Cap}_V(A) = \inf\{\, \mathcal{E}^V(f, f) + \|f\|_2^2 : f \in \mathcal{D}(\mathcal{E}^V),\ f \geq 1 \text{ on } A\ \mu\text{-a.e.}\}. \tag{7.117}$$

   A function $f \in \mathcal{D}(\mathcal{E}^V)$ admits a quasi-continuous representative $\tilde{f}$, continuous outside a set of capacity zero, satisfying $\tilde{f} = f$ μ-a.e..

2. **Excessive functions and resolvents.** The semigroup (P_t^V) defines an associated resolvent $(G_\alpha^V)_{\alpha>0}$:

$$G_\alpha^V f(x) := \int_0^\infty e^{-\alpha t}(P_t^V f)(x)\,dt = \int_E G_\alpha^V(x, y) f(y)\,\mu(dy). \tag{7.118}$$

   A measurable function $u : E \to [0, \infty]$ is (P_t^V)-excessive if $P_t^V u \leq u$ for all $t > 0$ and $\lim_{t\downarrow 0} P_t^V u = u$ μ-a.e. Such functions form the potential-theoretic basis of the Schrödinger operator, corresponding to superharmonic functions in classical analysis.

3. **Fine topology and pathwise interpretation.** The fine topology associated with (P_t^V) is the coarsest topology making all excessive functions finely continuous. This topology describes the most refined structure under which the trajectories $t \mapsto X_t$ are quasi-continuous. Under this topology, the potential V modifies the recurrence and transience properties of X_t by exponentially damping its occupation measure through the multiplicative functional $e^{-\int_0^t V(X_s)\,ds}$.

4. Quantum mechanical interpretation.
In quantum mechanics, the operator $H = -(\mathcal{L} - V)$ is the Hamiltonian corresponding to a particle evolving under the potential V. The semigroup e^{-tH} represents imaginary-time evolution, transforming the Schrödinger equation

$$i\frac{\partial \psi}{\partial t} = H\psi \tag{7.119}$$

into the heat-type equation

$$\frac{\partial u}{\partial t} = -Hu, \quad u(0,x) = f(x). \tag{7.120}$$

The Feynman–Kac representation provides a probabilistic realization of this evolution:

$$(e^{-tH}f)(x) = \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right]. \tag{7.121}$$

Thus, the Dirichlet-form framework rigorously unifies stochastic analysis, potential theory, and quantum mechanics by establishing a precise correspondence between the analytic semigroup (e^{-tH}), the probabilistic semigroup (P_t^V), and the energy form ($\mathcal{E}^V, \mathcal{D}(\mathcal{E}^V)$).

5. Boundary behaviour and reflective/absorbing conditions.
If E is a domain in $\mathbb{R}^d$, boundary conditions correspond to restrictions on the domain of the form:

- Dirichlet (absorbing) boundary conditions correspond to the subspace

$$\mathcal{D}_0(\mathcal{E}^V) = \overline{C_c^\infty(E)}^{\|\cdot\|_{\mathcal{E}^V}} \tag{7.122}$$

- Neumann (reflecting) boundary conditions correspond to extending $\mathcal{E}^V$ to include functions with zero normal derivative at ∂E.

In probabilistic terms, the associated Hunt process is either killed upon hitting ∂E (Dirichlet case) or reflected at the boundary (Neumann case). This provides a pathwise interpretation of quantum boundary phenomena via stochastic processes.

7.6 Dirichlet Form Perturbations and Feynman–Kac Semigroups

In this section, We give a fully rigorous account of the analytic and probabilistic constructions linking Dirichlet-form perturbations by potentials to Feynman–Kac-type semigroups and Schrödinger operators. We prove the equivalence of the form-sum construction of the Schrödinger operator and the pathwise Feynman–Kac expectation representation, treating both the Dirichlet-form (analytic) route and the stochastic-calculus (probabilistic) route, and we show how the two coincide under the standard hypotheses (bounded-below / form-bounded / Kato-class potentials). Throughout $(E, \mathcal{B}, \mu)$ is a locally compact separable metric space with Radon measure μ of full support, $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a symmetric Dirichlet form on $L^2(E,\mu)$, and $X = (X_t, \mathbb{P}^x)$ denotes the associated Hunt process with L^2-semigroup $T_t = e^{-t\mathcal{L}}$.

Notation and standing hypotheses. We write $\mathcal{L}$ for the self-adjoint nonnegative operator associated with $\mathcal{E}$ (so $\mathcal{E}(f,g) = \langle \mathcal{L}^{1/2}f, \mathcal{L}^{1/2}g\rangle$ and $T_t = e^{-t\mathcal{L}}$). Let $V : E \to \mathbb{R}$ be measurable. We consider three progressively more general hypotheses on V:

(H1) $V \in L^\infty(E,\mu)$ (bounded potential);

(H2) V is form-bounded w.r.t. $\mathcal{E}$ with relative form bound $a < 1$ (KLMN condition): there exist $a \in [0,1)$ and $b \geq 0$ such that for all $f \in \mathcal{D}(\mathcal{E})$,

$$\int_E |V|\,|f|^2\,d\mu \leq a\,\mathcal{E}(f,f) + b\|f\|_2^2; \tag{7.123}$$

(H3) V belongs to a suitable Kato class $\mathcal{K}$ (or form-bounded with arbitrarily small relative bound) permitting the stochastic Feynman–Kac functional to be defined and the form-sum to be closed.

We will first treat (H1), then extend to (H2) via form methods and monotone approximation, and finally indicate how Kato-class assumptions fit into the framework (H3).

Theorem 7.5 (Analytic construction of the Schrödinger operator)**.** *Assume (H2). Define the Schrödinger quadratic form*

$$\mathcal{E}^V(f,g) := \mathcal{E}(f,g) + \int_E V(x)f(x)g(x)\,\mu(dx), \qquad \mathcal{D}(\mathcal{E}^V) := \{f \in \mathcal{D}(\mathcal{E}) : \int_E |V|\,|f|^2\,d\mu < \infty\}. \tag{7.124}$$

Then $\mathcal{E}^V$ is closed and lower bounded; consequently there exists a unique self-adjoint lower-bounded operator H on $L^2(E,\mu)$ such that $\mathcal{D}(\mathcal{E}^V) = \mathcal{D}(H^{1/2})$ and

$$\mathcal{E}^V(f,g) = \langle H^{1/2}f, H^{1/2}g\rangle_{L^2}, \qquad f,g \in \mathcal{D}(\mathcal{E}^V). \tag{7.125}$$

Moreover H is the form-sum $H = \mathcal{L} +_{\text{form}} V$ and $-H$ generates a strongly continuous semigroup e^{-tH} on $L^2(E,\mu)$.

Proof. Since V is form-bounded with relative bound $a < 1$, the perturbation term $f \mapsto \int V|f|^2$ is $\mathcal{E}$-bounded with relative bound a. Hence $\mathcal{E}^V$ is lower semi-bounded and closed on the domain $\mathcal{D}(\mathcal{E}^V) = \mathcal{D}(\mathcal{E}) \cap L^2(|V|\,d\mu)$; this is the standard KLMN theorem (Kato–Lions–Milgram–Nelson). The first representation theorem (cf. Kato, Reed–Simon) produces a unique self-adjoint operator H with $\mathcal{D}(H^{1/2}) = \mathcal{D}(\mathcal{E}^V)$ and the stated quadratic representation. The semigroup existence and strong continuity of e^{-tH} follow from Hille–Yosida and spectral calculus since H is lower bounded and self-adjoint. Uniqueness as a form-sum is immediate from the representation theorem. □

Remark. Under (H1) the form $\mathcal{E}^V$ is simply a bounded perturbation of $\mathcal{E}$, in which case $\mathcal{D}(\mathcal{E}^V) = \mathcal{D}(\mathcal{E})$ and $H = \mathcal{L} + V$ on the natural domain.

We now define the Feynman–Kac semigroup and prove it is well defined under these hypotheses.

Definition 7.6 (Feynman–Kac semigroup)**.** For measurable $f \geq 0$ and $x \in E$, define

$$(P_t^V f)(x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \tag{7.126}$$

whenever the right-hand side is finite. For bounded-below V the multiplicative functional is well defined and P_t^V maps bounded measurable functions into bounded measurable functions.

We first show that for bounded potentials the semigroup P_t^V is the semigroup generated by the operator $H = \mathcal{L} +_{\mathrm{op}} V$.

Theorem 7.7 (Feynman–Kac for bounded potentials)**.** *Assume (H1). Let $H = \mathcal{L} + V$ be the self-adjoint operator on $\mathcal{D}(\mathcal{L})$ given by the bounded perturbation. Then for all bounded measurable f and for μ-a.e. $x \in E$,*

$$(P_t^V f)(x) = (e^{-tH} f)(x). \tag{7.127}$$

Moreover the identity holds in $L^2(E, \mu)$ and for $f \in C_c^\infty$ (or a core) it holds pointwise for all x.

Proof. We give two complementary proofs: (A) an analytic semigroup/resolvent argument based on Trotter product, and (B) a stochastic-Itô argument using Itô's formula (when the process admits a diffusion representation) which provides the same identity and intuition.

(A) *Analytic (Trotter) proof.* Because $V \in L^\infty$, the operator of multiplication by V is bounded and self-adjoint on L^2. Thus $H = \mathcal{L}+V$ is self-adjoint on $\mathcal{D}(\mathcal{L})$ and the Trotter product formula applies:

$$e^{-tH} = \lim_{n\to\infty} \left(e^{-(t/n)\mathcal{L}} e^{-(t/n)V}\right)^n, \tag{7.128}$$

with convergence in the strong operator topology on L^2. For bounded measurable f,

$$(e^{-(t/n)\mathcal{L}} e^{-(t/n)V} f)(x) = \mathbb{E}^x\big[e^{-(t/n)V(X_0)} f(X_{t/n})\big] = \mathbb{E}^x\big[e^{-(t/n)V(x)} f(X_{t/n})\big]. \tag{7.129}$$

Iterating and passing to the limit along the product yields the Feynman–Kac expectation,

$$\lim_{n\to\infty} \left(e^{-(t/n)\mathcal{L}} e^{-(t/n)V}\right)^n f = \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right] = P_t^V f(x), \tag{7.130}$$

and equality of the two semigroups follows by uniqueness of the strong limit (justified by dominated convergence and the semigroup bounds). This gives $e^{-tH} f = P_t^V f$ in L^2 and pointwise outside an exceptional set.

(B) *Stochastic (Itô) proof for diffusion realizations.* Suppose the Hunt process X solves the SDE representation (in local coordinates) driven by Brownian motion and a drift corresponding to $\mathcal{L}$, and suppose f is C^2 and bounded. Apply Itô's formula to the process

$$M_t := \exp\Big(-\int_0^t V(X_s)\,ds\Big) f(X_t). \tag{7.131}$$

Itô's formula (and product rule) gives

$$dM_t = \exp\Big(-\int_0^t V(X_s)\,ds\Big)\big((\mathcal{L}-V)f\big)(X_t)\,dt + (\text{martingale term}). \tag{7.132}$$

Taking expectations and using that the martingale term has zero expectation, we obtain

$$\frac{d}{dt}\mathbb{E}^x[M_t] = \mathbb{E}^x\Big[\exp\Big(-\int_0^t V(X_s)\,ds\Big)(\mathcal{L}-V)f(X_t)\Big]. \tag{7.133}$$

Thus

$$u(t,x) := \mathbb{E}^x[M_t] = (P_t^V f)(x) \tag{7.134}$$

satisfies the backward evolution equation

$$\partial_t u = (\mathcal{L} - V)u, \qquad u(0, \cdot) = f. \tag{7.135}$$

The semigroup generated by $\mathcal{L} - V$ coincides with e^{-tH}, and by uniqueness of solutions of the evolution equation in the relevant function class we conclude $P_t^V f = e^{-tH} f$. (This stochastic derivation gives the same identity under the diffusion representation; for general Hunt processes one may replace Itô by Dynkin's formula and potential-theoretic arguments.)

Combining (A) and (B) gives the desired equality. □

The bounded-potential case being settled, we now extend to form-bounded (KLMN) potentials by approximation.

Theorem 7.8 (Feynman–Kac via form approximation)**.** *Assume (H2). Let H be the self-adjoint operator of Theorem 7.5 and let P_t^V be the Feynman–Kac family defined (a priori) on bounded measurable f by the path expectation (the expectation is finite for a suitable class of starting points under (H2)). Then for every $f \in L^2(E, \mu)$,*

$$e^{-tH} f = P_t^V f \quad \text{in } L^2(E, \mu), \tag{7.136}$$

and the equality holds pointwise q.e. on E for quasi-continuous representatives.

Proof. We use the truncation/approximation method. Define $V_n := V \wedge n$ (and additionally $V_n := \max\{-n, V \wedge n\}$ if V is unbounded below) so that each $V_n \in L^\infty$ and $V_n \to V$ monotonically as $n \to \infty$. For each n, let $H_n := \mathcal{L} + V_n$ (bounded perturbation) and denote $P_t^{V_n}$ the corresponding Feynman–Kac operators. By Theorem 7.7,

$$e^{-tH_n} f = P_t^{V_n} f, \qquad f \in L^2. \tag{7.137}$$

By the form monotone convergence (or Kato monotone convergence of forms) $H_n \to H$ in the strong-resolvent sense; therefore $e^{-tH_n} \to e^{-tH}$ strongly on L^2. On the probabilistic side, by monotone convergence of the exponential weights (or dominated convergence using Khasminskii-type bounds justified by the form bound), we have $P_t^{V_n} f \to P_t^V f$ q.e. and in L^2. Taking limits in the identity $e^{-tH_n} f = P_t^{V_n} f$ yields $e^{-tH} f = P_t^V f$ in L^2. The q.e. pointwise statement follows from choosing quasi-continuous representatives for the left-hand side and using the pathwise convergence of the multiplicative functionals. □

Stochastic exponential / multiplicative-functional derivation (alternate probabilistic proof). We now provide the stochastic-calculus derivation that works in the diffusion setting and is the basis for Girsanov-type transformations which are often used to introduce drifts. Let X be given as the solution of the SDE corresponding to $\mathcal{L}$ on a fixed filtered probability space supporting an m-dimensional Brownian motion B, and assume coefficients give a unique strong solution. Let $f \in C_c^2(E)$ and consider the process

$$M_t := e^{-\int_0^t V(X_s)\, ds} f(X_t). \tag{7.138}$$

Applying Itô's formula (justified by smoothness and local boundedness) we get

$$\begin{aligned} dM_t &= e^{-\int_0^t V(X_s)\,ds}\Big[df(X_t) - V(X_t)f(X_t)\,dt\Big] \\ &= e^{-\int_0^t V(X_s)\,ds}\Big[(\mathcal{L}f)(X_t)\,dt + (\text{local martingale term}) - V(X_t)f(X_t)\,dt\Big] \\ &= e^{-\int_0^t V(X_s)\,ds}\Big[((\mathcal{L}-V)f)(X_t)\,dt + d\widetilde{M}_t\Big], \end{aligned} \tag{7.139}$$

where $\widetilde{M}_t$ is a local martingale. Taking expectations and using the optional stopping / bounded convergence argument for truncated stopping times gives

$$\frac{d}{dt}\mathbb{E}^x[M_t] = \mathbb{E}^x\big[e^{-\int_0^t V(X_s)\,ds}(\mathcal{L}-V)f(X_t)\big]. \tag{7.140}$$

Thus $u(t,x) := \mathbb{E}^x[M_t]$ satisfies $\partial_t u = (\mathcal{L}-V)u$ with $u(0,\cdot) = f$. Standard uniqueness of the Cauchy problem in L^2 or in appropriate function classes implies that $u(t,\cdot) = e^{-tH}f$ (where H is the self-adjoint operator associated with $\mathcal{E}^V$). This derivation complements the analytic approximation approach and provides the linkage to stochastic calculus and martingale theory.

Spectral consequences and ground states. From the operator equality $e^{-tH} = P_t^V$ we deduce spectral statements: the spectral measure of H controls the long-time behaviour of P_t^V, ground state existence corresponds to L^2-eigenfunctions ϕ with $H\phi = \lambda_0\phi$, and the ground-state transformation (Doob h-transform) can be expressed probabilistically via the reweighted path measure

$$\frac{d\mathbb{P}^{x,\phi}}{d\mathbb{P}^x}\Big|_{\mathcal{F}_t} = e^{\lambda_0 t}\frac{\phi(X_t)}{\phi(x)}e^{-\int_0^t V(X_s)\,ds}, \tag{7.141}$$

provided $\phi > 0$ a.e. and lies in the domain. Such transforms connect spectral gaps to ergodic properties of the transformed process.

Extension to Kato-class potentials and further remarks. If V belongs to a Kato class $\mathcal{K}$ appropriate to the underlying process (for instance the classical Kato class in $\mathbb{R}^d$ relative to Brownian motion), then the multiplicative functional $e^{-\int_0^t V(X_s)\,ds}$ is well defined and uniformly integrable on compacts; the form-sum $\mathcal{E}^V$ is closable and gives a self-adjoint H. The preceding approximation arguments (by bounded or truncated potentials) extend to this setting and produce the identity $e^{-tH} = P_t^V$ in L^2 and q.e. pointwise. Technicalities require verifying Khasminskii-type exponential moment bounds and the smallness of the negative part V^- in form-sense.

Concluding theorem (full equivalence).

Theorem 7.9 (Full equivalence)**.** *Assume V satisfies (H2) or (H3). Let H be the self-adjoint operator given by the form-sum $\mathcal{E}^V = \mathcal{E} + \int V|\cdot|^2$, and let P_t^V denote the Feynman–Kac expectation semigroup. Then*

$$e^{-tH} = P_t^V \quad \textit{as bounded operators on } L^2(E,\mu), \tag{7.142}$$

and for every $f \in L^2(E,\mu)$ the equality holds in L^2 and with quasi-continuous representatives holds q.e. on E.

Proof. Combine Theorems 7.5, 7.7 and 7.8 and the stochastic derivation above. The truncation/approximation argument shows strong L^2-convergence of e^{-tH_n} to e^{-tH} while monotone/dominated convergence for the stochastic expectations yields convergence $P_t^{V_n} \to P_t^V$ q.e. and in L^2; equality at the approximating level passes to the limit yielding the assertion. □

References and further reading. The analytic form-sum and KLMN theorem are classical (Kato, Reed–Simon). The probabilistic derivation via multiplicative functionals and the Feynman–Kac representation (including Kato-class extensions) are treated in detail in Fukushima–Oshima–Takeda, Simon's monograph on Schrödinger semigroups, and in classical probability texts on diffusion processes and potential theory. The combination of both analytic and probabilistic approaches presented here gives a robust and flexible toolkit for spectral, PDE, and pathwise analyses of Schrödinger-type operators.

Conclusion. The Dirichlet-form framework gives a complete and rigorous analytic–probabilistic formulation of Schrödinger operators and their Feynman–Kac semigroups. It integrates the theory of self-adjoint operators, stochastic processes, and potential theory into a single unified structure, enabling both analytic and probabilistic investigation of spectral properties, ground states, and boundary effects in quantum systems.

7.7 Conclusion

In this chapter we have developed the analytic foundation of Dirichlet forms and potential theory and shown how they underpin the *Feynman–Kac* representations and Schrödinger-type semigroups of earlier chapters. The unifying thread is that Markov processes, potential theory, Dirichlet forms, and semigroup theory form a single integrated structure: stochastic paths correspond to analytic resolvents and generators, and the fine topological structure of sample paths is encoded in analytic capacities and domain regularity of forms.

References: Fukushima–Oshima–Takeda [24], Ma–Röckner [25], Chen–Fukushima (2012) [27] and Chen–Fukushima (2012) [28] for general Dirichlet forms; Simon (1982) [40], Davies (1980) [26] for semigroup methods; Stroock–Varadhan (2007) [16], Bass (2004) [29] for process-potential theory interface.

Chapter 8

Path Integral Representation

"The theory of quantum mechanics describes nature as a sum over histories, each history contributing to the whole by its amplitude."
— Richard P. Feynman

The path integral formalism provides a deep unification between analysis, probability, and quantum mechanics, offering a measure-theoretic construction of semigroups and operators via expectations over stochastic paths. It reveals how analytic objects such as Schrödinger semigroups and heat kernels admit probabilistic representations through diffusion processes, and conversely, how probabilistic quantities may be interpreted as functional integrals on infinite-dimensional path spaces. This chapter rigorously formulates these connections, emphasizing the correspondence between operator-theoretic and stochastic constructions.

8.1 Wiener Measure and Configuration Space of Paths

Let $E = \mathbb{R}^d$ equipped with its Borel σ-algebra $\mathcal{B}(E)$. Set $\Omega := C([0,\infty); E)$ and denote by $X_t : \Omega \to E$ the canonical coordinate maps $X_t(\omega) = \omega(t)$. Equip Ω with the topology of uniform convergence on compact time intervals and with the corresponding Borel σ-algebra $\mathcal{F} = \mathcal{B}(\Omega)$. For every finite collection of times $0 \le t_1 < \cdots < t_n$ and Borel sets $A_1, \ldots, A_n \subset E$ the canonical cylinder sets

$$\{\omega \in \Omega : \ X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\} \tag{8.1}$$

generate $\mathcal{F}$; measurable functions on $(\Omega, \mathcal{F})$ are determined by their values on such cylinders.

The Wiener measure $\mathbb{P}^x$ is the unique probability measure on $(\Omega, \mathcal{F})$ whose finite-dimensional distributions are Gaussian with mean x at time 0 and covariance given by the minimum kernel. Equivalently, for every n, every choice $0 = t_0 < t_1 < \cdots < t_n$ and every bounded Borel $\varphi : E^n \to \mathbb{R}$,

$$\int_\Omega \varphi(X_{t_1}, \ldots, X_{t_n})\, d\mathbb{P}^x = \int_{E^n} \varphi(x_1, \ldots, x_n) \prod_{k=1}^{n} p_{t_k - t_{k-1}}(x_{k-1}, x_k)\, dx_1 \cdots dx_n, \tag{8.2}$$

where $x_0 := x$ and the transition densities $p_t(x, y)$ are given by the Gaussian heat kernel

$$p_t(x, y) = \frac{1}{(4\pi t)^{d/2}} \exp\Big(- \frac{|y - x|^2}{4t} \Big), \qquad t > 0,\ x, y \in \mathbb{R}^d. \tag{8.3}$$

The existence and uniqueness of $\mathbb{P}^x$ with these finite-dimensional marginals is exactly the Kolmogorov extension (or Kolmogorov–Consistence) theorem applied to the consistent family of probability measures on $\mathbb{R}^{dn}$ defined by the right-hand side; consistency is ensured by the Chapman–Kolmogorov identities

$$\int_E p_s(x, z)\, p_t(z, y)\, dz = p_{s+t}(x, y), \qquad s, t > 0, \tag{8.4}$$

which follow from the semigroup property of the Gaussian kernel.

The coordinate process $(X_t)_{t\geq 0}$ under $\mathbb{P}^x$ is a Gaussian process with mean x and covariance matrix

$$\mathbb{E}^x\big[(X_s^{(i)} - x^{(i)})(X_t^{(j)} - x^{(j)})\big] = \delta_{ij}\,(s \wedge t), \qquad 1 \leq i, j \leq d, \tag{8.5}$$

and by the Kolmogorov–Chentsov criterion the finite-dimensional Gaussian moment estimates imply that the canonical process admits a version with continuous sample paths; since the construction is carried out on $\Omega = C([0, \infty); E)$ the chosen version is continuous $\mathbb{P}^x$-almost surely. In fact one has the stronger Hölder-regularity statement: for every $\gamma \in (0, 1/2)$,

$$\mathbb{P}^x\Big(\sup_{0\leq s<t\leq T} \frac{|X_t - X_s|}{|t - s|^\gamma} < \infty \Big) = 1 \qquad \text{for every } T < \infty, \tag{8.6}$$

so sample paths are almost surely Hölder continuous of any order $< \frac{1}{2}$.

Define the natural (raw) filtration

$$\mathcal{F}_t^0 := \sigma(X_s : 0 \leq s \leq t) \tag{8.7}$$

and its right-continuous, $\mathbb{P}^x$-completed augmentation

$$\mathcal{F}_t := \bigcap_{u>t} (\mathcal{F}_u^0 \vee \mathcal{N}^{\mathbb{P}^x}) \tag{8.8}$$

where $\mathcal{N}^{\mathbb{P}^x}$ denotes the $\mathbb{P}^x$-null sets. The filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}^x)$ satisfies the usual conditions (complete and right-continuous) and the coordinate process is adapted and has continuous trajectories. The family $(\mathbb{P}^x)_{x\in E}$ is Markovian in the strong sense: for each bounded measurable $f : E \to \mathbb{R}$ and $s, t \geq 0$,

$$\mathbb{E}^x\big[f(X_{s+t}) \mid \mathcal{F}_s\big] = (T_t f)(X_s) \quad \text{a.s. under } \mathbb{P}^x, \tag{8.9}$$

where the transition operator T_t acts by convolution with the heat kernel,

$$(T_t f)(x) = \int_E p_t(x, y)\, f(y)\, dy. \tag{8.10}$$

Consequently $(T_t)_{t\geq 0}$ is a strongly continuous contraction semigroup on $L^p(\mathbb{R}^d)$ for each $1 \leq p < \infty$ and a Feller semigroup on $C_0(\mathbb{R}^d)$; in particular for bounded continuous f, $x \mapsto T_t f(x)$ is continuous and $\lim_{t\downarrow 0} T_t f = f$ uniformly on compacts.

From the Gaussian structure one derives explicit moment and exponential estimates used in potential theory and PDE connections: for every multi-index $\alpha \in \mathbb{N}^d$ and $t > 0$, the heat kernel is C^∞ in (x, y) and satisfies the Gaussian bounds

$$|\partial_x^\alpha p_t(x,y)| \leq C_\alpha\, t^{-(|\alpha|+d)/2} \exp\Big(-\frac{|x-y|^2}{8t}\Big), \tag{8.11}$$

which imply smoothing (instantaneous regularization) properties of T_t. In particular T_t maps L^2 into C^∞ for $t > 0$, and the generator of the semigroup is the self-adjoint Laplacian $\frac{1}{2}\Delta$ on its maximal L^2-domain.

Two further structural facts are central and frequently used. First, the Cameron–Martin space H associated with Wiener measure is the space of absolutely continuous paths $h : [0,\infty) \to \mathbb{R}^d$ with $h(0) = 0$ and $\dot{h} \in L^2_{\text{loc}}([0,\infty); \mathbb{R}^d)$; the Cameron–Martin theorem states that for $h \in H$ the shifted measure $\mathbb{P}^x \circ (\omega \mapsto \omega + h)^{-1}$ is equivalent to $\mathbb{P}^x$ and the Radon–Nikodym derivative is given by the stochastic exponential

$$\frac{d(\mathbb{P}^x \circ (\cdot + h)^{-1})}{d\mathbb{P}^x}(\omega) = \exp\Big(\int_0^\infty \dot{h}_s \cdot dX_s(\omega) - \tfrac{1}{2}\int_0^\infty |\dot{h}_s|^2\, ds\Big), \tag{8.12}$$

interpreted via Itô integrals on finite horizons and limiting arguments; this formula is the analytic backbone for Girsanov transformations in the diffusion setting. Second, time-shifts and scaling relationships are explicit: for any $c > 0$ the rescaled process $(c^{-1/2} X_{ct})_{t\geq 0}$ under $\mathbb{P}^{cx}$ has the same law as $(X_t)_{t\geq 0}$ under $\mathbb{P}^x$, reflecting Brownian scaling and enabling precise short-time asymptotics of the semigroup.

Finally, the Wiener measure and the canonical process provide the prototypical example of a Hunt process: (X_t) is strong Markov and has continuous trajectories, possesses transition densities satisfying Chapman–Kolmogorov, admits the Feller property, and its potential theory (Green kernels, capacity, excessive functions) coincides with the classical analytic potential theory for the Laplacian. These properties justify using $\mathbb{P}^x$ as the reference measure in Feynman–Kac constructions, in perturbative path-integral formulas, and in the probabilistic representation of solutions to parabolic and elliptic PDEs.

8.2 Feynman–Kac Functional and Schrödinger Semigroup Representation

We now present a rigorous formulation of the Feynman–Kac representation theorem, which connects the analytic theory of Schrödinger semigroups with the probabilistic theory of Brownian motion and its associated exponential additive functionals.

Let $E = \mathbb{R}^d$ equipped with the Borel σ-algebra $\mathcal{B}(E)$, and denote by $(B_t)_{t\geq 0}$ the standard

d-dimensional Brownian motion on the canonical path space

$$\Omega = C([0,\infty); \mathbb{R}^d), \tag{8.13}$$

with coordinate process $X_t(\omega) = \omega(t)$, natural filtration $\mathcal{F}_t = \sigma(X_s : 0 \le s \le t)$, and Wiener measure $\mathbb{P}^x$ corresponding to Brownian motion starting at $x \in \mathbb{R}^d$. The expectation with respect to $\mathbb{P}^x$ will be denoted by $\mathbb{E}^x[\cdot]$.

Let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable potential function which is *bounded from below*, i.e., there exists a constant $c \in \mathbb{R}$ such that $V(x) \ge c$ for all $x \in \mathbb{R}^d$. We consider the Schrödinger-type differential operator

$$H = -\frac{1}{2}\Delta + V, \tag{8.14}$$

initially defined on the core $C_c^\infty(\mathbb{R}^d)$. Since V is bounded from below, the quadratic form

$$\mathcal{E}^V(f,f) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla f(x)|^2\,dx + \int_{\mathbb{R}^d} V(x)|f(x)|^2\,dx, \qquad f \in C_c^\infty(\mathbb{R}^d), \tag{8.15}$$

is lower bounded and closable in $L^2(\mathbb{R}^d)$. Its closure defines a densely defined, closed, symmetric, semi-bounded quadratic form $(\mathcal{E}^V, \mathcal{D}(\mathcal{E}^V))$, which, by the Friedrichs extension theorem, corresponds to a unique self-adjoint operator H on $L^2(\mathbb{R}^d)$. We continue to denote this operator by $H = -\frac{1}{2}\Delta + V$, with domain

$$\mathcal{D}(H) = \left\{ f \in \mathcal{D}(\mathcal{E}^V) \;\middle|\; \exists g \in L^2(\mathbb{R}^d) \text{ s.t. } \mathcal{E}^V(f,h) = \langle g, h\rangle_{L^2(\mathbb{R}^d)}\ \forall h \in \mathcal{D}(\mathcal{E}^V) \right\}. \tag{8.16}$$

By the general spectral theorem for self-adjoint operators, H generates a strongly continuous, symmetric contraction semigroup $(e^{-tH})_{t\ge 0}$ on $L^2(\mathbb{R}^d)$, which is positivity-preserving and sub-Markovian. Explicitly, for $f \in L^2(\mathbb{R}^d)$,

$$e^{-tH} f = \int_0^\infty e^{-t\lambda}\, dE_\lambda f, \tag{8.17}$$

where E_λ denotes the spectral resolution of H.

On the probabilistic side, define the *Feynman–Kac functional* associated with the potential V by

$$A_t^V(\omega) := \int_0^t V(X_s(\omega))\,ds, \tag{8.18}$$

for each trajectory $\omega \in \Omega$. This is a continuous, adapted, additive functional of the Brownian motion. The exponential weight

$$\exp\Big(-A_t^V(\omega)\Big) = \exp\Big(-\int_0^t V(X_s(\omega))\,ds\Big) \tag{8.19}$$

acts as a multiplicative functional on the Wiener space, modifying the original Brownian measure to encode the influence of the potential energy accumulated along the path. Specifically, the random variable

$$M_t^V(\omega) := \exp\Big(-\int_0^t V(X_s(\omega))\,ds\Big) \tag{8.20}$$

serves as a Radon–Nikodym derivative defining a new weighted measure $\mathbb{Q}_t^x$ on $\mathcal{F}_t$ by

$$\frac{d\mathbb{Q}_t^x}{d\mathbb{P}^x}(\omega) = M_t^V(\omega). \tag{8.21}$$

This measure reweights the paths of Brownian motion according to their accumulated potential energy, yielding a precise probabilistic analogue of the Feynman–Kac path integral measure.

We may now state the Feynman–Kac representation in its rigorous form: for every $f \in L^2(\mathbb{R}^d)$ and almost every $x \in \mathbb{R}^d$,

$$(e^{-tH} f)(x) = \mathbb{E}^x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right]. \tag{8.22}$$

Equivalently, in semigroup notation,

$$P_t^V f(x) := \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right] = (e^{-tH} f)(x), \tag{8.23}$$

where $(P_t^V)_{t\geq 0}$ is referred to as the *Feynman–Kac semigroup*. This semigroup is strongly continuous on $L^2(\mathbb{R}^d)$, self-adjoint, and preserves positivity.

The equality above provides a rigorous probabilistic realization of the quantum-mechanical path integral:

$$\langle x|e^{-tH}|y\rangle = \int_{\substack{X(0)=x \\ X(t)=y}} \exp\left(-\int_0^t \left[\frac{1}{2}|\dot{X}_s|^2 + V(X_s)\right] ds\right) \mathcal{D}X, \tag{8.24}$$

where $\mathcal{D}X$ denotes the formal (nonexistent) Lebesgue measure on path space, which in the rigorous theory is replaced by the Wiener measure $\mathbb{P}^x$. Thus, the Feynman–Kac formula establishes the exact analytic–probabilistic equivalence

$$\boxed{(e^{-tH} f)(x) = \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right], \qquad H = -\tfrac{1}{2}\Delta + V,} \tag{8.25}$$

thereby connecting Schrödinger operators, semigroup theory, and stochastic analysis through the framework of Dirichlet forms and probabilistic path integration.

8.3 Analytic Justification and Domain Correspondence

We now provide a analytic framework connecting the potential-perturbed Dirichlet form with the corresponding Schrödinger operator, thereby establishing a precise equivalence between the analytic and probabilistic formulations of the Feynman–Kac semigroup.

Let us consider the classical Dirichlet form on $L^2(\mathbb{R}^d, dx)$ associated with the Laplacian operator:

$$\mathcal{E}(f,g) = \frac{1}{2}\int_{\mathbb{R}^d} \nabla f(x)\cdot\nabla g(x)\,dx, \qquad f,g \in H^1(\mathbb{R}^d), \tag{8.26}$$

where $H^1(\mathbb{R}^d)$ denotes the standard Sobolev space of square-integrable functions with square-integrable weak derivatives. The form $(\mathcal{E}, H^1(\mathbb{R}^d))$ is symmetric, closed, and Markovian, and hence constitutes a strongly local, regular Dirichlet form whose self-adjoint generator on $L^2(\mathbb{R}^d)$ is the operator

$$\mathcal{L} = \frac{1}{2}\Delta, \tag{8.27}$$

with domain $\mathcal{D}(\mathcal{L}) = H^2(\mathbb{R}^d)$. The associated symmetric Markov process is the standard d-dimensional Brownian motion $(B_t)_{t\geq 0}$ on $\mathbb{R}^d$, defined on its canonical path space.

Let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable potential function, and consider the perturbed bilinear form

$$\mathcal{E}^V(f,g) = \frac{1}{2}\int_{\mathbb{R}^d} \nabla f(x)\cdot\nabla g(x)\,dx + \int_{\mathbb{R}^d} V(x)f(x)g(x)\,dx, \qquad f,g \in C_c^\infty(\mathbb{R}^d), \tag{8.28}$$

which is initially defined on the dense subspace $C_c^\infty(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$.

To ensure that the potential term defines a relatively form-bounded perturbation of $\mathcal{E}$, we impose the *Kato-class condition* on V, given by

$$\lim_{t\downarrow 0}\sup_{x\in\mathbb{R}^d} \mathbb{E}^x\left[\int_0^t |V(B_s)|\,ds\right] = 0. \tag{8.29}$$

This condition guarantees that V is locally integrable with respect to the transition probabilities of Brownian motion and ensures, in particular, that the additive functional

$$A_t^V := \int_0^t V(B_s)\,ds \tag{8.30}$$

is well-defined for almost all trajectories and that its exponential moments satisfy

$$\mathbb{E}^x\left[e^{-\int_0^t V(B_s)\,ds}\right] < \infty, \qquad \forall t > 0. \tag{8.31}$$

Under the Kato-class assumption, the form $\mathcal{E}^V$ admits a unique closed extension on $L^2(\mathbb{R}^d)$, which we again denote by $(\mathcal{E}^V, \mathcal{D}(\mathcal{E}^V))$. This form is densely defined, lower semi-bounded, and symmetric. By the general representation theorem for closed semi-bounded quadratic forms, there exists a unique self-adjoint operator H_V on $L^2(\mathbb{R}^d)$ such that

$$\mathcal{E}^V(f,g) = \langle H_V^{1/2} f, H_V^{1/2} g\rangle_{L^2(\mathbb{R}^d)}, \qquad f,g \in \mathcal{D}(\mathcal{E}^V). \tag{8.32}$$

This operator H_V is precisely the Friedrichs extension of the formal differential operator

$$H_V = -\frac{1}{2}\Delta + V, \tag{8.33}$$

initially defined on $C_c^\infty(\mathbb{R}^d)$. Its domain $\mathcal{D}(H_V)$ satisfies

$$\mathcal{D}(H_V) = \left\{ f \in \mathcal{D}(\mathcal{E}^V) \;\middle|\; \exists\, g \in L^2(\mathbb{R}^d) \text{ such that } \mathcal{E}^V(f,h) = \langle g, h\rangle_{L^2(\mathbb{R}^d)}\ \forall h \in \mathcal{D}(\mathcal{E}^V) \right\}, \tag{8.34}$$

and $H_V f = g$.

The self-adjointness of H_V implies the existence of a strongly continuous contraction semigroup $(e^{-tH_V})_{t\geq 0}$ on $L^2(\mathbb{R}^d)$, which is positivity-preserving and sub-Markovian. Analytically, this semigroup can be represented through the spectral theorem as

$$e^{-tH_V} f = \int_0^\infty e^{-t\lambda}\, dE_\lambda f, \tag{8.35}$$

where E_λ denotes the spectral family associated with H_V.

On the other hand, the Feynman–Kac semigroup is defined probabilistically by

$$P_t^V f(x) := \mathbb{E}^x\Big[\exp\Big(-\int_0^t V(B_s)\, ds\Big) f(B_t)\Big], \qquad f \in L^2(\mathbb{R}^d). \tag{8.36}$$

Standard results in potential theory and Dirichlet form theory (see, e.g., Fukushima–Oshima–Takeda, *Dirichlet Forms and Symmetric Markov Processes*) assert that $(P_t^V)_{t\geq 0}$ forms a strongly continuous symmetric contraction semigroup on $L^2(\mathbb{R}^d)$, and moreover,

$$P_t^V = e^{-tH_V}, \qquad \forall t \geq 0. \tag{8.37}$$

Thus, the analytic semigroup generated by H_V coincides exactly with the probabilistic Feynman–Kac semigroup.

Consequently, the domains, resolvents, and spectral measures of H_V are completely determined by the corresponding probabilistic structure of the weighted Brownian motion, and vice versa. The identification

$$\boxed{\mathcal{E}^V(f,f) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla f|^2\, dx + \int_{\mathbb{R}^d} V f^2\, dx \quad \Longleftrightarrow \quad (e^{-tH_V} f)(x) = \mathbb{E}^x\Big[e^{-\int_0^t V(B_s)\, ds} f(B_t)\Big]} \tag{8.38}$$

rigorously establishes the complete analytic–probabilistic correspondence between Schrödinger operators and Feynman–Kac-type path integral representations.

8.4 Path Integral as Infinite-Dimensional Limit

The construction of the Feynman–Kac path integral representation proceeds through a limiting procedure that replaces the heuristic "integration over paths" by a well-defined limit of finite-dimensional Gaussian integrals.

Let $E = \mathbb{R}^d$ and let $H = -\frac{1}{2}\Delta + V$ be the Schrödinger operator acting on $L^2(\mathbb{R}^d)$ with suitable domain (for instance, $V \in L^1_{\text{loc}}$ and $V \geq -C$ ensuring semi-boundedness). Denote by $(T_t)_{t\geq 0} = (e^{-tH})_{t\geq 0}$ its strongly continuous contraction semigroup. By the Trotter product formula one has, for any $f \in L^2(\mathbb{R}^d)$,

$$e^{-tH} f = \lim_{n\to\infty} \Big(e^{-\frac{t}{n}V} e^{\frac{t}{2n}\Delta}\Big)^n f, \tag{8.39}$$

where the limit is in the strong operator topology of $L^2(\mathbb{R}^d)$. The semigroup $e^{\frac{t}{2n}\Delta}$ acts by convolution with the Gaussian kernel $p_{t/(2n)}(x,y)$, while the multiplication operator $e^{-\frac{t}{n}V}$ acts pointwise by $e^{-(t/n)V(x)}$. Expanding this product iteratively and writing out the kernel explicitly yields the finite-dimensional approximation formula

$$(e^{-tH}f)(x) = \lim_{n\to\infty} \int_{(\mathbb{R}^d)^n} K_n(x, x_1, \dots, x_n)\, f(x_n)\, dx_1 \cdots dx_n, \tag{8.40}$$

where

$$K_n(x_0, \dots, x_n) := \prod_{i=0}^{n-1} \Big[p_{t/n}(x_i, x_{i+1})\, e^{-(t/n)V(x_i)} \Big], \qquad x_0 := x. \tag{8.41}$$

Substituting the explicit Gaussian form

$$p_{t/n}(x_i, x_{i+1}) = \frac{1}{(2\pi t/n)^{d/2}} \exp\left(-\frac{|x_{i+1} - x_i|^2}{2(t/n)} \right), \tag{8.42}$$

one obtains

$$(e^{-tH}f)(x) = \lim_{n\to\infty} c_n \int_{(\mathbb{R}^d)^n} \exp\left[-\sum_{i=0}^{n-1} \left(\frac{|x_{i+1} - x_i|^2}{2(t/n)} + \frac{t}{n} V(x_i) \right) \right] f(x_n)\, dx_1 \cdots dx_n, \tag{8.43}$$

where $c_n = (2\pi t/n)^{-nd/2}$. This expression already exhibits the discrete path-integral structure: each factor describes the contribution of a short Brownian bridge increment $\Delta X_i = x_{i+1} - x_i$ with Gaussian weight $\exp(-|\Delta X_i|^2/(2\Delta t))$, and the potential term contributes the exponential damping $\exp(-V(x_i)\Delta t)$.

To justify the passage $n \to \infty$, we view the family of Gaussian densities

$$\mu_{t,n}(dx_1, \dots, dx_n) := c_n \exp\left(-\sum_{i=0}^{n-1} \frac{|x_{i+1} - x_i|^2}{2(t/n)} \right) dx_1 \cdots dx_n \tag{8.44}$$

as the joint distribution of the discrete-time Gaussian process obtained by sampling a Brownian motion at the grid points $t_i = i(t/n)$. In the limit $n \to \infty$ the family $(\mu_{t,n})$ is projectively consistent, and by Kolmogorov's extension theorem it defines the Wiener measure $\mathbb{P}^x$ on the path space $C([0,t];\mathbb{R}^d)$. Under this identification, for any bounded measurable F depending only on finitely many coordinates,

$$\int F(x_1, \dots, x_n)\, \mu_{t,n}(dx_1, \dots, dx_n) = \mathbb{E}^x[F(X_{t_1}, \dots, X_{t_n})], \tag{8.45}$$

where X_s is Brownian motion started at x. Hence, in the limit, the discrete sum $\sum_{i=0}^{n-1} |x_{i+1} - x_i|^2/(2\Delta t)$ converges in distribution to the quadratic variation $\frac{1}{2}\int_0^t |\dot{X}_s|^2 ds$ that appears in the formal Wiener measure exponent, while the Riemann sum $\sum_i V(x_i)\Delta t$ converges to $\int_0^t V(X_s)\, ds$.

Passing to the limit $n \to \infty$ under these identifications therefore yields the *Feynman–Kac representation*:

$$(e^{-tH}f)(x) = \mathbb{E}^x\left[\exp\left(-\int_0^t V(X_s)\, ds \right) f(X_t) \right], \tag{8.46}$$

which is rigorous for all $f \in L^2(\mathbb{R}^d)$ and measurable potentials V satisfying standard form-boundedness conditions ensuring that

$$\mathcal{E}^V(u,u) = \tfrac{1}{2}\int |\nabla u|^2 + \int V u^2 \tag{8.47}$$

defines a closed, lower-bounded quadratic form. The kernel of the semigroup, $K_t(x,y)$, admits the probabilistic representation

$$K_t(x,y) = p_t(x,y)\,\mathbb{E}_t^{x,y}\Big[\exp\Big(-\int_0^t V(X_s)\,ds\Big)\Big], \tag{8.48}$$

where $\mathbb{E}_t^{x,y}$ denotes expectation under the Brownian bridge measure from x to y in time t. This expression is the rigorous probabilistic counterpart of the heuristic Feynman path integral

$$\langle x|e^{-tH}|y\rangle = \int_{X(0)=x}^{X(t)=y} \exp\Big(-\int_0^t \Big[\tfrac{1}{2}|\dot{X}_s|^2 + V(X_s)\Big]ds\Big)\,\mathcal{D}X, \tag{8.49}$$

where $\mathcal{D}X$ is the formally uniform (nonexistent) "Lebesgue measure" on path space, replaced rigorously by the Wiener measure $\mathbb{P}^x$. Thus the Feynman–Kac formula realizes, within the framework of measure theory and stochastic analysis, the infinite-dimensional limit of Gaussian integrals that physicists denote as the *Euclidean path integral*.

8.5 Connection to Dirichlet Forms and Potential Theory

The Feynman–Kac framework extends with full mathematical generality to the theory of symmetric Dirichlet forms and their associated Markov processes.

Let $(E, \mathcal{B}, \mu)$ be a σ-finite measure space, and let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a symmetric, closed, and Markovian bilinear form on $L^2(E,\mu)$. Suppose further that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is *quasi-regular*, ensuring the existence of an associated Hunt process $(X_t, \mathbb{P}^x)$ with state space E and properly measurable sample paths.

For a measurable potential $V : E \to \mathbb{R}$ that is $\mathcal{E}$-measurable and satisfies the integrability condition

$$\mathbb{E}^x\Big[\exp\Big(-\int_0^t |V(X_s)|\,ds\Big)\Big] < \infty, \quad \forall t > 0, \text{ for quasi-every } x \in E, \tag{8.50}$$

one defines the *Feynman–Kac semigroup* $(P_t^V)_{t\geq 0}$ acting on $L^2(E,\mu)$ by

$$P_t^V f(x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \quad f \in L^2(E,\mu). \tag{8.51}$$

This semigroup is strongly continuous, positivity-preserving, and symmetric in $L^2(E,\mu)$. It corresponds to the analytic perturbation of the original Dirichlet form by the potential V, in the sense that

$$\mathcal{E}^V(f,g) = \mathcal{E}(f,g) + \int_E V(x)f(x)g(x)\,d\mu(x), \qquad f,g \in \mathcal{D}(\mathcal{E}) \cap L^2(E, |V|\,d\mu), \tag{8.52}$$

and its generator is precisely the Schrödinger-type operator

$$H = -\mathcal{L} + V, \tag{8.53}$$

where $\mathcal{L}$ is the self-adjoint generator associated with $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$.

From the probabilistic perspective, the Feynman–Kac semigroup provides an expectation over paths of the Hunt process (X_t), weighted by the exponential of the negative time-integral of the potential. That is, for all $f \in L^2(E, \mu)$,

$$P_t^V f(x) = \mathbb{E}^x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right]. \tag{8.54}$$

From the analytic viewpoint, P_t^V is the strongly continuous symmetric semigroup whose generator H corresponds to the perturbed form $\mathcal{E}^V$. The fundamental equivalence can thus be expressed as the following dual characterization:

$$\boxed{\text{Analytic: } \mathcal{E}^V(f,f) = \mathcal{E}(f,f) + \int_E V f^2\,d\mu \quad \Longleftrightarrow \quad \text{Probabilistic: } P_t^V f(x) = \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right].} \tag{8.55}$$

Hence, the Feynman–Kac formula realizes a rigorous equivalence principle between potential perturbations in the analytic theory of Dirichlet forms and exponential path-weighting in stochastic process theory. Every symmetric Schrödinger operator of the form

$$H = -\mathcal{L} + V, \tag{8.56}$$

where $\mathcal{L}$ is the generator of a quasi-regular symmetric Dirichlet form, possesses a precise probabilistic representation via the expected exponential of the path integral of the potential along the trajectories of X_t.

8.6 Conclusion

The path integral representation thus bridges three distinct mathematical structures: the analytic semigroup e^{-tH}, the probabilistic expectation $\mathbb{E}^x[e^{-\int_0^t V(X_s)\,ds} f(X_t)]$, and the functional integral on infinite-dimensional space. It unifies potential theory, spectral theory, and stochastic calculus within a single conceptual framework, providing a rigorous measure-theoretic foundation for the Feynman–Kac correspondence and, by extension, for the Euclidean formulation of quantum mechanics.

References: Ghosh (2025) [32], Simon (1979) [19], Albeverio et. al. (2012) [30], Simon (2005) [20].

Chapter 9

Functional Analytic Properties

"Analysis is the art of taming infinity by approximation."

— Jean Dieudonné

This chapter presents the fundamental functional-analytic structure underlying the theory of Dirichlet forms, semigroups, and Schrödinger-type operators. We rigorously study domains, generators, spectral properties, and compactness criteria within the framework of self-adjoint operators on L^2-spaces. The interplay between semigroup theory, spectral resolution, and variational principles will serve as the analytic backbone of subsequent probabilistic constructions.

9.1 Strongly Continuous Semigroups and Their Generators

Let $\mathcal{H} = L^2(E, \mu)$ be a real or complex separable Hilbert space and let $(T_t)_{t\geq 0}$ be a family of bounded linear operators on $\mathcal{H}$ satisfying the semigroup law $T_{t+s} = T_t T_s$ for all $s, t \geq 0$ and $T_0 = I$. We say (T_t) is a *strongly continuous semigroup* (or C_0-semigroup) if for every $f \in \mathcal{H}$,

$$\lim_{t\downarrow 0} \|T_t f - f\|_{\mathcal{H}} = 0. \tag{9.1}$$

Strong continuity implies several elementary but important consequences which we record first. There exist constants $M \geq 1$ and $\omega \in \mathbb{R}$ (the *growth bound*) such that for all $t \geq 0$,

$$\|T_t\|_{\mathcal{B}(\mathcal{H})} \leq M e^{\omega t}. \tag{9.2}$$

Equivalently the map $t \mapsto T_t$ is locally uniformly bounded in operator norm and, for every $\alpha > \omega$, the Laplace transform

$$R(\alpha) := \int_0^\infty e^{-\alpha t} T_t \, dt \tag{9.3}$$

converges in operator norm and defines a bounded operator on $\mathcal{H}$. The operator $R(\alpha)$ is the resolvent of the infinitesimal generator L at α, see below.

The *infinitesimal generator* $(L, \mathcal{D}(L))$ of (T_t) is defined by

$$\mathcal{D}(L) := \Big\{ f \in \mathcal{H} : \lim_{t\downarrow 0} \frac{T_t f - f}{t} \text{ exists in } \mathcal{H} \Big\}, \qquad Lf := \lim_{t\downarrow 0} \frac{T_t f - f}{t}, \quad f \in \mathcal{D}(L). \tag{9.4}$$

One checks that L is a closed, densely defined linear operator and that for every $f \in \mathcal{D}(L)$ the map $t \mapsto T_t f$ is continuously differentiable with

$$\frac{d}{dt}T_t f = T_t L f = L T_t f, \qquad t \geq 0. \tag{9.5}$$

Thus (T_t) is the unique (mild) solution operator of the abstract Cauchy problem

$$\begin{cases} \frac{du}{dt}(t) = Lu(t), & t > 0, \\ u(0) = f, \end{cases} \qquad u(t) = T_t f. \tag{9.6}$$

Two fundamental generation results characterize which closed operators arise as generators of C_0-semigroups. We state them in the forms most used in the theory of linear evolution equations on Hilbert spaces.

Theorem 9.1 (Hille–Yosida (Hilbert-space form))**.** *A linear operator $(L, \mathcal{D}(L))$ on $\mathcal{H}$ is the generator of a strongly continuous semigroup $(T_t)_{t\geq 0}$ with $\|T_t\| \leq Me^{\omega t}$ if and only if L is closed, densely defined, and for every $\alpha > \omega$ the resolvent $(\alpha I - L)^{-1}$ exists as a bounded operator on $\mathcal{H}$ and satisfies the uniform resolvent estimate*

$$\|(\alpha I - L)^{-n}\| \leq \frac{M}{(\alpha - \omega)^n}, \qquad n \in \mathbb{N}. \tag{9.7}$$

Proof. We prove the Hilbert-space version of the Hille–Yosida theorem in two directions. All limits and integrals below are taken in the strong (i.e. norm) topology of the Hilbert space $\mathcal{H}$ unless otherwise stated.

Let $(T_t)_{t\geq 0}$ be a strongly continuous semigroup on $\mathcal{H}$ with the growth bound $|T_t| \leq Me^{\omega t}$ for all $t \geq 0$. Denote by L its infinitesimal generator. First we show necessity: L is closed, densely defined, and the resolvent estimates hold.

Because (T_t) is a C_0-semigroup, its generator L is by standard semigroup theory densely defined and closed; moreover $\mathcal{D}(L)$ is dense in $\mathcal{H}$. For $\alpha > \omega$ define the operator

$$R(\alpha) := \int_0^\infty e^{-\alpha t} T_t, dt, \tag{9.8}$$

where the Bochner integral converges in operator norm because

$$|e^{-\alpha t} T_t| \leq M e^{-(\alpha - \omega)t} \tag{9.9}$$

and

$$\int_0^\infty e^{-(\alpha-\omega)t}, dt = (\alpha - \omega)^{-1} < \infty \tag{9.10}$$

Differentiating under the integral sign (justified by the uniform exponential bound) yields

$$(\alpha I - L)R(\alpha) = \int_0^\infty e^{-\alpha t}(\alpha I - L)T_t, dt = \int_0^\infty e^{-\alpha t}\frac{d}{dt}(T_t), dt = [-e^{-\alpha t}T_t]_{t=0}^{t=\infty} + \alpha \int_0^\infty e^{-\alpha t} T_t, dt = I. \tag{9.11}$$

A similar calculation shows $R(\alpha)(\alpha I - L) = I$. Hence $R(\alpha) = (\alpha I - L)^{-1}$ exists and is bounded for every $\alpha > \omega$. Moreover

$$|(\alpha I - L)^{-1}| = |R(\alpha)| \leq \int_0^\infty e^{-\alpha t}|T_t|, dt \leq M \int_0^\infty e^{-(\alpha-\omega)t}, dt = \frac{M}{\alpha - \omega}. \tag{9.12}$$

To obtain the stated power-norm estimates, use the Laplace representation for higher resolvent powers. For every integer $n \geq 1$ and $\alpha > \omega$ we may write (Bochner integral, iterated)

$$(\alpha I - L)^{-n} = \frac{1}{(n-1)!} \int_0^\infty t^{,n-1} e^{-\alpha t} T_t, dt, \tag{9.13}$$

which is obtained by successive differentiation of the scalar Laplace transform or by convolution powers of the kernel. Taking norms and using the growth bound one gets

$$|(\alpha I - L)^{-n}| \leq \frac{1}{(n-1)!} \int_0^\infty t^{,n-1} e^{-\alpha t}|T_t|, dt \leq \frac{M}{(n-1)!} \int_0^\infty t^{,n-1} e^{-(\alpha-\omega)t}, dt = \frac{M}{(n-1)!} \cdot \frac{(n-1)!}{(\alpha-\omega)^n} = \frac{M}{(\alpha-\omega)^n} \tag{9.14}$$

This proves the necessity.

We now prove sufficiency: assume L is a closed densely defined operator and that for some constants $M \geq 1$, $\omega \in \mathbb{R}$ and every $\alpha > \omega$ the resolvent $(\alpha I - L)^{-1}$ exists on $\mathcal{H}$ and satisfies

$$|(\alpha I - L)^{-n}| \leq \frac{M}{(\alpha - \omega)^n}, \qquad n \in \mathbb{N}. \tag{9.15}$$

We will construct a C_0-semigroup $(T_t)_{t \geq 0}$ with $|T_t| \leq M e^{\omega t}$ and generator L.

For each fixed $\alpha > \omega$ define the bounded (Yosida) approximant

$$L_\alpha := \alpha^2 (\alpha I - L)^{-1} - \alpha I = \alpha L (\alpha I - L)^{-1}. \tag{9.16}$$

One checks algebraically, using

$$(\alpha I - L)^{-1} L = \alpha(\alpha I - L)^{-1} - I \tag{9.17}$$

that L_α is a bounded operator on $\mathcal{H}$. The resolvent bounds give the uniform estimate (for $\alpha > \omega$)

$$|L_\alpha| \leq \alpha^2 |(\alpha I - L)^{-1}| + \alpha \leq \frac{\alpha^2 M}{\alpha - \omega} + \alpha < \infty, \tag{9.18}$$

so L_α generates the uniformly continuous semigroup e^{tL_α} on $\mathcal{H}$. We shall show that the family $T_t^\alpha := e^{tL_\alpha}$ converges strongly, as $\alpha \downarrow \infty$, to a C_0-semigroup T_t whose generator is L.

Two standard and equivalent constructions may be used; we present the one via Laplace transforms (resolvents) because it yields clean operator-norm estimates and the growth bound directly.

Fix $\lambda > \omega$. For each $\alpha > \lambda$ compute the resolvent of L_α. Using the algebraic identity between resolvents one obtains (straightforward algebra; details follow)

$$(\lambda I - L_\alpha)^{-1} = (\lambda I - L)^{-1}\big(I + (L_\alpha - L)(\lambda I - L_\alpha)^{-1}\big), \tag{9.19}$$

and, more usefully for our purpose, one derives the explicit relation

$$(\lambda I - L_\alpha)^{-1} = (\lambda I - L)^{-1}\Big(I - (\alpha - \lambda)(\alpha I - L)^{-1}\Big)^{-1}. \tag{9.20}$$

From the hypothesis

$$|(\alpha I - L)^{-1}| \leq M/(\alpha - \omega) \tag{9.21}$$

and the choice $\lambda > \omega$ we may choose α large enough that

$$|(\alpha - \lambda)(\alpha I - L)^{-1}| < 1 \tag{9.22}$$

hence the Neumann series is valid and $(\lambda I - L_\alpha)^{-1}$ is well defined and bounded uniformly in α for α large. A direct computation (using the resolvent bounds and the Neumann series summation) yields the estimate

$$|(\lambda I - L_\alpha)^{-1}| \leq \frac{M}{\lambda - \omega} \cdot \frac{1}{1 - (\alpha - \lambda)\frac{M}{\alpha - \omega}} \leq C(\lambda, M, \omega) \tag{9.23}$$

for α sufficiently large; in particular the resolvents $(\lambda I - L_\alpha)^{-1}$ are uniformly bounded and converge strongly to $(\lambda I - L)^{-1}$ as $\alpha \to \infty$. Indeed, this strong convergence follows from the identity

$$(\lambda I - L_\alpha)^{-1} - (\lambda I - L)^{-1} = (\lambda I - L_\alpha)^{-1}(L_\alpha - L)(\lambda I - L)^{-1}, \tag{9.24}$$

combined with $\lim_{\alpha\to\infty} |(L_\alpha - L)u| = 0$ for each $u \in \mathcal{D}(L)$ (a standard property of Yosida approximants), and density of $\mathcal{D}(L)$.

Having established that the bounded-resolvent family $(\lambda I - L_\alpha)^{-1}$ converges strongly to $(\lambda I - L)^{-1}$ for some (hence all) $\lambda > \omega$, we apply the Post–Widder inversion (or the Laplace inversion formula for resolvents) to pass to the semigroup level. For each α the semigroup $T_t^\alpha = e^{tL_\alpha}$ satisfies the Laplace representation

$$(\lambda I - L_\alpha)^{-1} = \int_0^\infty e^{-\lambda t} T_t^\alpha, dt, \qquad \lambda > |L_\alpha|. \tag{9.25}$$

By dominated convergence in Bochner integrals (using uniform bounds on $|T_t^\alpha|$ to be established below) and the strong resolvent convergence we deduce that the family $(T_t^\alpha)_{t\geq 0}$ converges strongly, for each fixed $t \geq 0$, to an operator T_t characterized by

$$(\lambda I - L)^{-1} = \int_0^\infty e^{-\lambda t} T_t, dt, \qquad \lambda > \omega. \tag{9.26}$$

The operator-valued Laplace transform inversion then ensures that $(T_t)_{t\geq 0}$ is a strongly continuous semigroup whose generator has resolvent $(\lambda I - L)^{-1}$ and therefore must coincide with L. We must verify the uniform bound $|T_t| \leq Me^{\omega t}$ and the strong continuity.

To obtain the growth bound, we use the uniform resolvent estimates. For each $\alpha > \omega$, the bounded generator L_α produces a uniformly continuous semigroup T_t^α with

$$T_t^\alpha = \frac{1}{2\pi i} \int_\gamma e^{t\zeta} (\zeta I - L_\alpha)^{-1}, d\zeta, \tag{9.27}$$

where γ is a vertical contour in the complex plane to the right of the spectrum of L_α. Using the resolvent estimates inherited from the hypothesis one obtains, after contour deformation and standard resolvent–semigroup estimates, that for all $t \geq 0$ and all $\alpha > \omega$,

$$|T_t^\alpha| \leq M e^{\omega t}. \tag{9.28}$$

One concrete way to see this is to use the inequality

$$|(\beta I - L_\alpha)^{-1}| \leq M/(\beta - \omega) \tag{9.29}$$

for $\beta > \omega$, which follows from the resolvent identities and the assumed bounds; then integrate the scalar bound

$$\int_0^\infty e^{-\beta t} |T_t^\alpha|, dt \leq M/(\beta - \omega) \tag{9.30}$$

and apply the Post–Widder inversion to obtain the multiplicative factor $Me^{\omega t}$.

Since $|T_t^\alpha| \leq Me^{\omega t}$ uniformly in α and since T_t^α converges strongly to T_t for each fixed (t) (as argued via resolvent convergence and Laplace inversion), it follows that $|T_t| \leq Me^{\omega t}$ as well. Strong continuity of $t \mapsto T_t f$ comes from the dominated convergence in the representation by Yosida approximants and the fact that each T_t^α is strongly continuous; explicitly, for any $f \in \mathcal{H}$,

$$|T_t f - f| \leq |T_t f - T_t^\alpha f| + |T_t^\alpha f - f| \tag{9.31}$$

and the first term tends to 0 as $\alpha \to \infty$ for fixed t (strong convergence), while the second tends to 0 as $t \downarrow 0$ (continuity of the uniformly continuous semigroup T_t^α), which yields the desired strong continuity of T_t.

Finally, by standard uniqueness theorems for generators of C_0-semigroups, the generator of (T_t) is the closure of L (which we assumed closed) and thus equals L. This completes the proof of sufficiency.

Combining the two directions yields the equivalence: L is the generator of a strongly continuous semigroup with bound $|T_t| \leq Me^{\omega t}$ if and only if L is closed, densely defined, and its resolvent satisfies the uniform bounds

$$|(\alpha I - L)^{-n}| \leq M/(\alpha - \omega)^n \tag{9.32}$$

for all integers $n \geq 1$ and all $\alpha > \omega$. □

In practice one uses the equivalent resolvent estimate for $n = 1$ together with density and closedness. The Hille–Yosida theorem also yields the representation of the resolvent via the

Laplace transform:

$$(\alpha I - L)^{-1} f = \int_0^\infty e^{-\alpha t} T_t f \, dt, \qquad \alpha > \omega, \; f \in \mathcal{H}, \tag{9.33}$$

which is the basic bridge between the semigroup and generator viewpoints.

Another characterization, tailor-made for contractive semigroups, is given by the Lumer–Phillips theorem.

Theorem 9.2 (Lumer–Phillips)**.** *A densely defined linear operator L on a Hilbert space $\mathcal{H}$ generates a contraction semigroup $(T_t)_{t\geq 0}$ (i.e. $\|T_t\| \leq 1$ for all $t \geq 0$) if and only if L is dissipative*

$$\operatorname{Re}\langle Lf, f\rangle_{\mathcal{H}} \leq 0, \qquad \forall f \in \mathcal{D}(L), \tag{9.34}$$

and the range condition $\operatorname{Range}(\lambda I - L) = \mathcal{H}$ *holds for some (equivalently all) $\lambda > 0$.*

Proof. We prove both implications.

(Necessity.) Assume L generates a contraction C_0-semigroup $(T_t)_{t\geq 0}$ on $\mathcal{H}$, so $\|T_t\| \leq 1$ for all $t \geq 0$. The generator L is densely defined and closed. For any $f \in \mathcal{D}(L)$ consider the function $\phi(t) := \|T_t f\|^2$. Since $t \mapsto T_t f$ is differentiable in norm for $f \in \mathcal{D}(L)$ with derivative $T_t L f$, we compute

$$\phi'(t) = 2 \operatorname{Re}\langle T_t L f, T_t f\rangle. \tag{9.35}$$

Contractivity implies $\phi(t) \leq \phi(0) = \|f\|^2$ for all $t \geq 0$, hence $\phi'(0) \leq 0$. Evaluating at $t = 0$ gives

$$\phi'(0) = 2 \operatorname{Re}\langle Lf, f\rangle \leq 0, \tag{9.36}$$

which yields the dissipativity condition

$$\operatorname{Re}\langle Lf, f\rangle \leq 0, \qquad \forall f \in \mathcal{D}(L). \tag{9.37}$$

Next, fix $\lambda > 0$. The Laplace transform representation of the resolvent for generators gives

$$(\lambda I - L)^{-1} f = \int_0^\infty e^{-\lambda t} T_t f \, dt, \tag{9.38}$$

valid for all $f \in \mathcal{H}$. In particular the range of $\lambda I - L$ is all of $\mathcal{H}$, since the resolvent operator is everywhere defined and bounded. Thus the range condition holds for every $\lambda > 0$.

(Sufficiency.) Assume L is densely defined, dissipative, and that $\operatorname{Range}(\lambda I - L) = \mathcal{H}$ for some $\lambda > 0$. We will show that L generates a contraction semigroup.

Step 1: resolvent estimates. Fix $\lambda > 0$ with $\operatorname{Range}(\lambda I - L) = \mathcal{H}$. For arbitrary $f \in \mathcal{H}$ let $u \in \mathcal{D}(L)$ be the (unique) solution of $(\lambda I - L)u = f$. Taking the inner product of the equation with u and taking real parts yields

$$\lambda \|u\|^2 - \operatorname{Re}\langle Lu, u\rangle = \operatorname{Re}\langle f, u\rangle. \tag{9.39}$$

Using dissipativity $\mathrm{Re}\langle Lu, u\rangle \le 0$ we obtain

$$\lambda \|u\|^2 \le \|f\| \, \|u\|, \tag{9.40}$$

hence $\|u\| \le \|f\|/\lambda$. Therefore the resolvent satisfies

$$\|(\lambda I - L)^{-1}\| \le \frac{1}{\lambda}. \tag{9.41}$$

In particular the resolvent is a bounded operator on $\mathcal{H}$.

Step 2: Yosida approximants and approximating semigroups. For each $n \in \mathbb{N}$ set $\lambda_n := \lambda + n$ and define the bounded operator

$$L_n := \lambda_n L(\lambda_n I - L)^{-1} = \lambda_n\big(I - \lambda_n(\lambda_n I - L)^{-1}\big). \tag{9.42}$$

Each L_n is everywhere defined and bounded on $\mathcal{H}$. Moreover for $u \in \mathcal{D}(L)$ one verifies

$$L_n u = \lambda_n(\lambda_n I - L)^{-1} L u \to Lu \quad (n \to \infty), \tag{9.43}$$

so L_n converges strongly to L on $\mathcal{D}(L)$. Each L_n generates a uniformly continuous semigroup e^{tL_n} and, using the resolvent estimate from Step 1, one checks the contraction bound

$$\|e^{tL_n}\| \le 1 \qquad \forall t \ge 0, \ \forall n \in \mathbb{N}. \tag{9.44}$$

Indeed, for bounded generators one may use the exponential formula and the resolvent bound to obtain contractivity; alternatively one can show dissipativity of L_n and apply the Lumer–Phillips criterion for bounded operators.

Step 3: strong convergence to a contraction semigroup. Fix $f \in \mathcal{H}$. For each $t \ge 0$ the family $(e^{tL_n} f)_{n\in\mathbb{N}}$ is bounded in $\mathcal{H}$. We claim that $e^{tL_n} f$ converges as $n \to \infty$ and that the limit defines a strongly continuous contraction semigroup. To see this, observe that for $\lambda' > \lambda$ the resolvents satisfy

$$(\lambda' I - L_n)^{-1} = (\lambda' I - L)(\lambda_n I - L)^{-1}(\lambda_n I - L_n)^{-1}, \tag{9.45}$$

and using the resolvent bounds one establishes that $(\lambda' I - L_n)^{-1}$ converges strongly to $(\lambda' I - L)^{-1}$ as $n \to \infty$. The Laplace inversion formula (Post–Widder) then implies that $e^{tL_n} f$ converges strongly for each fixed t to some vector $T_t f$, and the limit operators T_t satisfy

$$(\lambda' I - L)^{-1} = \int_0^\infty e^{-\lambda' t} T_t \, dt, \qquad \lambda' > \lambda. \tag{9.46}$$

Using dominated convergence and the uniform contractive bound of e^{tL_n} we deduce $\|T_t\| \le 1$ for every $t \ge 0$. The semigroup property $T_{t+s} = T_t T_s$ follows from taking limits of the semigroup property for e^{tL_n}, and strong continuity at $t = 0$ follows from the same limit argument combined with uniform boundedness and the fact that each approximating semigroup is uniformly continuous.

Step 4: identification of the generator. It remains to show that the generator of (T_t) is the closure of L, which by assumption equals L since L is closed. From the Laplace representation and standard resolvent–generator inversion (compare the Hille–Yosida correspondence) we obtain that the resolvent of the generator of (T_t) equals $(\lambda I - L)^{-1}$ for $\lambda > \lambda_0$, hence the generator coincides with L. Therefore L generates the contraction semigroup $(T_t)_{t\geq 0}$.

Combining necessity and sufficiency completes the proof of the theorem. □

In the semigroup-to-generator direction one obtains the dissipativity automatically: if $\|T_t\| \leq 1$ then for $f \in \mathcal{D}(L)$,

$$\mathrm{Re}\langle Lf, f\rangle = \lim_{t\downarrow 0} \frac{\mathrm{Re}\langle T_t f - f, f\rangle}{t} \leq 0, \tag{9.47}$$

since $\|T_t f\| \leq \|f\|$. Conversely, dissipativity plus the range condition yields generation of a contraction semigroup.

Important operational tools associated to generators are the Yosida approximants and the resolvent identities. For $\alpha > \omega$ define the bounded Yosida approximation

$$L_\alpha := \alpha L(\alpha I - L)^{-1} = \alpha\big(I - \alpha(\alpha I - L)^{-1}\big), \tag{9.48}$$

which satisfies $L_\alpha \in \mathcal{B}(\mathcal{H})$, $\|L_\alpha\| \leq C(\alpha)$, and for every $f \in \mathcal{D}(L)$,

$$\lim_{\alpha\to\infty} L_\alpha f = Lf. \tag{9.49}$$

The resolvent identity

$$(\alpha I - L)^{-1} - (\beta I - L)^{-1} = (\beta - \alpha)(\alpha I - L)^{-1}(\beta I - L)^{-1} \tag{9.50}$$

is frequently used to pass between different resolvent parameters and to prove strong continuity properties.

We now record several finer properties of the semigroup and generator which are used repeatedly in analysis of PDEs and stochastic processes.

Domain invariance and regularity. If $f \in \mathcal{D}(L)$ then $T_t f \in \mathcal{D}(L)$ for all $t \geq 0$ and the map $t \mapsto T_t f$ is C^1 with derivative $LT_t f$. Moreover, if T_t is analytic (which occurs, for instance, when L is sectorial), then T_t maps $\mathcal{H}$ into $\mathcal{D}(L^k)$ for every $k \in \mathbb{N}$ and each fixed $t > 0$.

Approximation by bounded operators (Chernoff product formula). If A is a bounded operator approximating $I + tL$ at small times in a suitable sense, then

$$T_t = \lim_{n\to\infty} \Big(A(\tfrac{t}{n})\Big)^n, \tag{9.51}$$

with strong convergence on $\mathcal{H}$. The Trotter–Kato product formula is a central instance used to combine semigroups generated by simpler operators to produce the semigroup of a sum.

Spectral mapping and growth/decay rates. The spectral mapping theorem for C_0-semigroups states that for the point spectrum,

$$e^{t\sigma_p(L)} \subset \sigma_p(T_t), \tag{9.52}$$

and for the full spectrum one has the inclusion $e^{t\sigma(L)} \setminus \{0\} \subset \sigma(T_t)$; equality may fail in full generality but holds under additional hypotheses (e.g. boundedness or normality). The growth bound ω_0 defined by

$$\omega_0 := \inf\{\omega \in \mathbb{R} : \exists M \geq 1, \ \|T_t\| \leq M e^{\omega t} \ \forall t \geq 0\} \tag{9.53}$$

satisfies $\omega_0 \geq s(L) := \sup\{\operatorname{Re}\lambda : \ \lambda \in \sigma(L)\}$ and in many self-adjoint or sectorial settings equality holds.

Cores and perturbation stability. A subspace $\mathcal{C} \subset \mathcal{D}(L)$ is called a *core* for L if the closure of $L|_{\mathcal{C}}$ equals L. In practice one often verifies generation or properties of L on a convenient core (e.g. C_c^∞ for differential generators) and extends by closure. Under relatively bounded perturbations (Kato–Rellich type), generation and essential properties of the semigroup persist; specifically, if B is L-bounded with relative bound $a < 1$ then $L + B$ is closed and generates a C_0-semigroup.

Concrete examples. The prototypical examples on $L^2(\mathbb{R}^d)$ are the (scaled) Laplacian $(\frac{1}{2}\Delta, \ H^2(\mathbb{R}^d))$ generating the heat semigroup $T_t f = p_t * f$, and second-order uniformly elliptic operators with lower-order drift terms, which under suitable form or sectorial conditions generate analytic semigroups. In probability these are the transition semigroups of diffusion processes and form the bridge to Dirichlet-form constructions.

In summary, strongly continuous semigroups on Hilbert spaces are completely characterized by their infinitesimal generators via Hille–Yosida and Lumer–Phillips theorems; resolvent and Yosida approximations provide constructive tools; cores and perturbation results enable verification in concrete settings; and spectral mapping plus analyticity furnish fine regularity and long-time asymptotics essential for both PDE and stochastic analyses.

9.2 Self-Adjoint Operators and Spectral Theorem

Let $(T_t)_{t\geq 0}$ be a strongly continuous semigroup on the Hilbert space $L^2(E,\mu)$ such that for all $f, g \in L^2(E,\mu)$ and all $t \geq 0$,

$$\langle T_t f, g\rangle_{L^2(E,\mu)} = \langle f, T_t g\rangle_{L^2(E,\mu)}. \tag{9.54}$$

That is, each T_t is self-adjoint. We will rigorously show that its infinitesimal generator L is self-adjoint, nonpositive, and that the spectral theorem provides a complete functional calculus for L and (T_t).

9.2.1 Self-adjointness of the generator

Lemma 9.3. *If $(T_t)_{t\geq 0}$ is a symmetric, strongly continuous semigroup of contractions on $L^2(E,\mu)$, then its infinitesimal generator L is self-adjoint and satisfies $\langle Lf, f\rangle_{L^2(E,\mu)} \leq 0$ for all $f \in \mathcal{D}(L)$.*

Proof. Let $f, g \in \mathcal{D}(L)$ and recall that by definition of the generator,

$$Lf = \lim_{t\downarrow 0} \frac{T_t f - f}{t}, \qquad \text{in } L^2(E,\mu). \tag{9.55}$$

Since T_t is symmetric, we have for all $t > 0$,

$$\langle T_t f, g\rangle = \langle f, T_t g\rangle. \tag{9.56}$$

Dividing both sides by t and subtracting $\langle f, g\rangle / t$ yields

$$\left\langle \frac{T_t f - f}{t}, g \right\rangle = \left\langle f, \frac{T_t g - g}{t} \right\rangle. \tag{9.57}$$

Passing to the limit $t \downarrow 0$ gives

$$\langle Lf, g\rangle = \langle f, Lg\rangle, \qquad \forall f, g \in \mathcal{D}(L), \tag{9.58}$$

showing that L is symmetric. Furthermore, contractivity implies $\|T_t f\| \leq \|f\|$, and differentiating $\|T_t f\|^2$ at $t = 0$ yields

$$\frac{d}{dt}\|T_t f\|^2\Big|_{t=0} = 2\,\mathrm{Re}\langle Lf, f\rangle \leq 0, \tag{9.59}$$

so $\mathrm{Re}\langle Lf, f\rangle \leq 0$. Thus L is dissipative and symmetric. By a standard result in semigroup theory (see, e.g., Reed–Simon, Functional Analysis, Thm. X.49), a densely defined symmetric dissipative operator generating a contraction semigroup is self-adjoint and nonpositive. □

9.2.2 Spectral representation

Since L is self-adjoint and nonpositive, the spectral theorem ensures the existence of a unique projection-valued measure E_λ on $\mathbb{R}_-$ such that

$$L = \int_{\mathbb{R}_-} \lambda\, dE_\lambda. \tag{9.60}$$

For any Borel measurable function $\varphi : \mathbb{R}_- \to \mathbb{C}$, we define

$$\varphi(L) = \int_{\mathbb{R}_-} \varphi(\lambda)\, dE_\lambda, \tag{9.61}$$

which gives the full functional calculus for L. In particular, setting $\varphi_t(\lambda) = e^{t\lambda}$, we obtain the semigroup representation

$$T_t = e^{tL} = \int_{\mathbb{R}_-} e^{t\lambda}\, dE_\lambda, \qquad t \geq 0. \tag{9.62}$$

This representation is strongly continuous in t, and since $|e^{t\lambda}| \le 1$ for $\lambda \le 0$, we have $\|T_t\| \le 1$, confirming that (T_t) is a contraction semigroup. Moreover, differentiability in t follows from dominated convergence under the spectral integral, hence (T_t) is analytic.

9.2.3 Quadratic form and fractional powers

Define the quadratic form $\mathcal{E}$ associated with L by

$$\mathcal{E}(f,g) := -\langle Lf, g\rangle, \qquad f,g \in \mathcal{D}(L). \tag{9.63}$$

Since L is self-adjoint and nonpositive, $\mathcal{E}$ is symmetric, closed, and positive definite. The spectral calculus allows us to define fractional powers of $(-L)$ by

$$(-L)^\alpha = \int_{\mathbb{R}_-} (-\lambda)^\alpha \, dE_\lambda, \qquad \alpha > 0. \tag{9.64}$$

In particular, the domain of the square root is given by

$$\mathcal{D}((-L)^{1/2}) = \left\{ f \in L^2(E,\mu) : \int_{\mathbb{R}_-} (-\lambda)\, d\langle E_\lambda f, f\rangle < \infty \right\}, \tag{9.65}$$

and satisfies

$$\|(-L)^{1/2} f\|_2^2 = \int_{\mathbb{R}_-} (-\lambda)\, d\langle E_\lambda f, f\rangle = \mathcal{E}(f,f). \tag{9.66}$$

Hence the quadratic form $\mathcal{E}$ coincides with the Dirichlet form associated to L, and the correspondence

$$\mathcal{E}(f,f) = \|(-L)^{1/2} f\|_2^2, \qquad \mathcal{D}(\mathcal{E}) = \mathcal{D}((-L)^{1/2}), \tag{9.67}$$

establishes the equivalence between the analytic structure defined by the form and the spectral properties of the self-adjoint generator.

9.3 Compactness and Spectral Discreteness

A central analytic property in the study of self-adjoint generators of Dirichlet forms and Schrödinger-type operators concerns the *compactness* of the semigroup $(T_t)_{t\ge 0}$. Compactness has deep implications for the spectral structure of the generator L and the asymptotic behaviour of the semigroup.

9.3.1 Compactness criterion and its consequences

Let $(T_t)_{t\ge 0}$ be a strongly continuous, symmetric semigroup on the Hilbert space $L^2(E,\mu)$ with self-adjoint, nonpositive generator L. We say that T_t is *compact* if, for some (hence all) $t > 0$, the operator $T_t : L^2(E,\mu) \to L^2(E,\mu)$ is compact, i.e. it maps bounded sets into relatively compact sets.

Lemma 9.4. *If T_t is compact for one $t > 0$, then T_s is compact for all $s > 0$. Moreover, L has purely discrete spectrum.*

Proof. Suppose that T_{t_0} is compact for some $t_0 > 0$. For any $s > 0$, write

$$T_s = \begin{cases} T_{t_0}^{\lfloor s/t_0 \rfloor} T_r, & \text{if } s > t_0, \text{ with } r = s - \lfloor s/t_0 \rfloor t_0, \\ T_{t_0}^{1/2} T_{s-t_0/2} T_{t_0}^{1/2}, & \text{if } s < t_0. \end{cases} \tag{9.68}$$

Since products of compact and bounded operators are compact, it follows that T_s is compact for every $s > 0$. Hence, T_t forms a compact semigroup.

By the spectral theorem, the compactness of T_t implies that its spectrum $\sigma(T_t)$ consists of a countable sequence of real eigenvalues $\{e^{-t\lambda_n}\}_{n=1}^{\infty}$ accumulating only at 0, each with finite multiplicity. Correspondingly, the spectrum of the generator L is given by

$$\sigma(L) = \{-\lambda_n\}_{n=1}^{\infty}, \qquad \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots, \quad \lambda_n \to +\infty. \tag{9.69}$$

Each $\lambda_n > 0$ is an eigenvalue of $-L$ with eigenfunction ϕ_n, satisfying

$$L\phi_n = -\lambda_n \phi_n. \tag{9.70}$$

Since $T_t = e^{tL}$, we have

$$T_t \phi_n = e^{-t\lambda_n} \phi_n. \tag{9.71}$$

The family $\{\phi_n\}_{n \geq 1}$ forms an orthonormal basis of $L^2(E, \mu)$, by the completeness property of spectral projections associated with compact self-adjoint operators. □

9.3.2 Examples and sufficient conditions for compactness

Compactness of (T_t) (and hence discreteness of the spectrum of L) arises in several important contexts:

(i) **Finite measure space:** If $\mu(E) < \infty$ and (T_t) is a symmetric, positivity-preserving Markov semigroup (e.g. the heat semigroup on a bounded domain with Dirichlet boundary conditions), then compactness of T_t follows from the Rellich–Kondrachov compact embedding theorem:

$$H_0^1(E) \hookrightarrow L^2(E, \mu) \text{ compactly}, \tag{9.72}$$

which ensures that the corresponding generator (typically $-\frac{1}{2}\Delta$) has purely discrete spectrum.

(ii) **Confining potential:** For the Schrödinger operator

$$H = -\frac{1}{2}\Delta + V(x) \tag{9.73}$$

on $\mathbb{R}^d$, if the potential $V(x)$ satisfies $V(x) \to +\infty$ as $|x| \to \infty$, then the embedding of the associated form domain

$$\mathcal{D}(\mathcal{E}^V) = \{f \in H^1(\mathbb{R}^d) : V^{1/2} f \in L^2(\mathbb{R}^d)\} \tag{9.74}$$

into $L^2(\mathbb{R}^d)$ is compact. Consequently, e^{-tH} is a compact operator for all $t > 0$, and the spectrum of H consists of discrete eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ with finite multiplicities and $\lambda_n \to +\infty$.

(iii) **Elliptic operators on compact manifolds:** If E is a compact Riemannian manifold and L is the Laplace–Beltrami operator, then $T_t = e^{tL}$ is compact on $L^2(E)$ because the embedding $H^1(E) \hookrightarrow L^2(E)$ is compact. In this case, the eigenfunctions $\{\phi_n\}$ form an orthonormal basis of smooth functions on E, and the eigenvalues $\{\lambda_n\}$ satisfy Weyl's asymptotic law:

$$\lambda_n \sim C_d n^{2/d}, \qquad n \to \infty. \tag{9.75}$$

9.3.3 Summary of implications

Hence, the following equivalence holds for a symmetric semigroup (T_t) with self-adjoint generator L:

$$T_t \text{ compact for some } t > 0 \quad \Longleftrightarrow \quad L \text{ has purely discrete spectrum.} \tag{9.76}$$

In that case, we have the spectral decompositions

$$T_t f = \sum_{n=1}^{\infty} e^{-t\lambda_n} \langle f, \phi_n \rangle \phi_n, \tag{9.77}$$

$$Lf = -\sum_{n=1}^{\infty} \lambda_n \langle f, \phi_n \rangle \phi_n, \tag{9.78}$$

for all $f \in L^2(E, \mu)$, with $\{\phi_n\}$ forming a complete orthonormal basis and $\lambda_n \to +\infty$.

9.4 Form Methods and the Friedrichs Extension

Let $\mathcal{H} = L^2(E, \mu)$ be a separable Hilbert space, and let $\mathcal{E} : \mathcal{D}(\mathcal{E}) \times \mathcal{D}(\mathcal{E}) \to \mathbb{R}$ be a densely defined, symmetric bilinear form satisfying

$$\mathcal{E}(f, g) = \mathcal{E}(g, f), \qquad \forall f, g \in \mathcal{D}(\mathcal{E}). \tag{9.79}$$

Assume moreover that $\mathcal{E}$ is *nonnegative*, i.e.

$$\mathcal{E}(f, f) \geq 0, \qquad \forall f \in \mathcal{D}(\mathcal{E}). \tag{9.80}$$

Closability of Forms: A form $\mathcal{E}$ is said to be *closable* if for every sequence $(f_n) \subset \mathcal{D}(\mathcal{E})$ satisfying

$$f_n \to 0 \text{ in } L^2(E, \mu), \qquad \text{and} \qquad \mathcal{E}(f_n - f_m, f_n - f_m) \to 0,$$

we have

$$\mathcal{E}(f_n, f_n) \to 0.$$

In this case, there exists a smallest closed extension $(\overline{\mathcal{E}}, \mathcal{D}(\overline{\mathcal{E}}))$ such that $\mathcal{E} \subset \overline{\mathcal{E}}$, obtained by completing $\mathcal{D}(\mathcal{E})$ under the norm

$$\|f\|_{\mathcal{E}} := \left(\mathcal{E}(f, f) + \|f\|_2^2\right)^{1/2}. \tag{9.81}$$

The pair $(\overline{\mathcal{E}}, \mathcal{D}(\overline{\mathcal{E}}))$ is called a *closed symmetric form* on $L^2(E, \mu)$.

9.4.1 First Representation Theorem (Kato–Friedrichs Theorem)

A central result connecting closed forms and self-adjoint operators is the *first representation theorem.* Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a densely defined, symmetric, closed, nonnegative form on $\mathcal{H}$. Then there exists a unique nonnegative self-adjoint operator L on $\mathcal{H}$ such that

$$\mathcal{D}(\mathcal{E}) = \mathcal{D}((-L)^{1/2}), \qquad \mathcal{E}(f,g) = \langle (-L)^{1/2} f, (-L)^{1/2} g\rangle_{L^2}, \quad \forall f,g \in \mathcal{D}(\mathcal{E}). \tag{9.82}$$

Conversely, given a nonnegative self-adjoint operator L, the bilinear form defined by the right-hand side above is closed, symmetric, and nonnegative.

Proof. Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a densely defined, symmetric, closed, nonnegative form on the Hilbert space $\mathcal{H}$. Define the energy norm $\|\cdot\|_{\mathcal{E}}$ on $\mathcal{D}(\mathcal{E})$ by

$$\|u\|_{\mathcal{E}} := \left(\mathcal{E}(u,u) + \|u\|_{\mathcal{H}}^2\right)^{1/2}. \tag{9.83}$$

Closedness of $\mathcal{E}$ means that $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$ is a Hilbert space. The canonical embedding $j : (\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}}) \hookrightarrow \mathcal{H}$ is continuous and has dense range because $\mathcal{D}(\mathcal{E})$ is dense in $\mathcal{H}$.

For each fixed $u \in \mathcal{D}(\mathcal{E})$ the map $\mathcal{D}(\mathcal{E}) \ni v \mapsto \mathcal{E}(u,v)$ is a bounded linear functional on $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$. Indeed, by the Cauchy–Schwarz inequality for the form,

$$|\mathcal{E}(u,v)| \le \mathcal{E}(u,u)^{1/2}\mathcal{E}(v,v)^{1/2} \le \mathcal{E}(u,u)^{1/2}\|v\|_{\mathcal{E}}. \tag{9.84}$$

Hence, by the Riesz representation theorem in the Hilbert space $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$, there exists a unique element $\widetilde{L}u \in \mathcal{D}(\mathcal{E})$ (viewed as the Riesz representative with respect to $\langle\cdot,\cdot\rangle_{\mathcal{E}}$) such that for all $v \in \mathcal{D}(\mathcal{E})$,

$$\mathcal{E}(u,v) + \langle u,v\rangle_{\mathcal{H}} = \langle \widetilde{L}u, v\rangle_{\mathcal{E}}. \tag{9.85}$$

Unwinding the right-hand side in terms of the energy inner product yields, by definition of $\langle\cdot,\cdot\rangle_{\mathcal{E}}$,

$$\mathcal{E}(u,v) + \langle u,v\rangle_{\mathcal{H}} = \mathcal{E}(\widetilde{L}u, v) + \langle \widetilde{L}u, v\rangle_{\mathcal{H}}. \tag{9.86}$$

Rearranging gives the symmetric identity

$$\mathcal{E}(u - \widetilde{L}u, v) + \langle u - \widetilde{L}u, v\rangle_{\mathcal{H}} = 0, \qquad \forall v \in \mathcal{D}(\mathcal{E}). \tag{9.87}$$

Since the left-hand side is the $\mathcal{E}$-inner product with $u - \widetilde{L}u$, this implies $u = \widetilde{L}u$ as an identity in the Hilbert space $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$. Thus the Riesz representative is precisely u itself and the previous representation can be rewritten to exhibit a linear operator relating the form to the Hilbert-space inner product.

Define the linear operator $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ by declaring $u \in \mathcal{D}(A)$ if and only if the linear functional $\mathcal{D}(\mathcal{E}) \ni v \mapsto \mathcal{E}(u,v)$ is bounded in the $\mathcal{H}$-norm, i.e. there exists $h \in \mathcal{H}$ such that

$$\mathcal{E}(u,v) = \langle h, v\rangle_{\mathcal{H}}, \qquad \forall v \in \mathcal{D}(\mathcal{E}). \tag{9.88}$$

For such u we set $Au := h$. Equivalently,

$$\mathcal{D}(A) = \{u \in \mathcal{D}(\mathcal{E}) :\ \exists h \in \mathcal{H} \text{ s.t. } \mathcal{E}(u,v) = \langle h, v\rangle_{\mathcal{H}}\ \forall v \in \mathcal{D}(\mathcal{E})\}, \tag{9.89}$$

and $A : \mathcal{D}(A) \to \mathcal{H}$ is linear by construction. Observe that $\mathcal{D}(A)$ is dense in $\mathcal{H}$ because $\mathcal{D}(\mathcal{E})$ is dense and the form is closed.

We now show that A is self-adjoint and nonnegative. First, for $u \in \mathcal{D}(A)$ and $v \in \mathcal{D}(\mathcal{E})$ we have

$$\langle Au, v\rangle_{\mathcal{H}} = \mathcal{E}(u,v) = \mathcal{E}(v,u) = \overline{\langle Av, u\rangle_{\mathcal{H}}}, \tag{9.90}$$

whenever $v \in \mathcal{D}(A)$ as well; hence A is symmetric on $\mathcal{D}(A)$. Next, for $u \in \mathcal{D}(A)$ taking $v = u$ in the defining relation gives

$$\langle Au, u\rangle_{\mathcal{H}} = \mathcal{E}(u,u) \geq 0, \tag{9.91}$$

so A is nonnegative.

To see that A is self-adjoint, consider the operator $B := I + A$ with domain $\mathcal{D}(B) = \mathcal{D}(A)$. For $u \in \mathcal{D}(B)$ and $v \in \mathcal{D}(\mathcal{E})$ we have

$$\langle Bu, v\rangle_{\mathcal{H}} = \langle u, v\rangle_{\mathcal{H}} + \mathcal{E}(u,v) =: \ell_u(v), \tag{9.92}$$

and the map $v \mapsto \ell_u(v)$ is a bounded linear functional on $(\mathcal{D}(\mathcal{E}), \|\cdot\|_{\mathcal{E}})$. By the Riesz representation theorem there exists a unique $w \in \mathcal{D}(\mathcal{E})$ such that $\ell_u(v) = \langle w, v\rangle_{\mathcal{E}}$ for all v. Unwinding definitions shows B is surjective onto $\mathcal{H}$ and has bounded inverse $B^{-1} : \mathcal{H} \to \mathcal{D}(\mathcal{E}) \subset \mathcal{H}$. In particular B is self-adjoint and strictly positive, hence $A = B - I$ is self-adjoint and nonnegative. We therefore set $L := A$, obtaining a self-adjoint nonnegative operator associated with the form.

It remains to identify the form domain with the domain of the square root of $-L$. Since L is nonnegative and self-adjoint, functional calculus yields the square root $(-L)^{1/2}$ and its domain $\mathcal{D}((-L)^{1/2})$. For $u \in \mathcal{D}(A)$ we have

$$\|(-L)^{1/2}u\|_{\mathcal{H}}^2 = \langle -Lu, u\rangle_{\mathcal{H}} = \mathcal{E}(u,u), \tag{9.93}$$

so $\mathcal{D}(A) \subset \mathcal{D}((-L)^{1/2})$ and the equality of norms shows that the form $\mathcal{E}$ coincides with the quadratic form induced by $(-L)^{1/2}$ on $\mathcal{D}(A)$. By closedness and density the identity extends to

$$\mathcal{D}(\mathcal{E}) = \mathcal{D}((-L)^{1/2}), \qquad \mathcal{E}(u,v) = \langle (-L)^{1/2}u, (-L)^{1/2}v\rangle_{\mathcal{H}}, \quad \forall u, v \in \mathcal{D}(\mathcal{E}). \tag{9.94}$$

Uniqueness of L follows because any self-adjoint nonnegative operator L' satisfying the displayed identity must have the same quadratic form and hence the same spectral measure; by functional calculus the operators L and L' coincide.

Conversely, if L is a given nonnegative self-adjoint operator on $\mathcal{H}$, define the quadratic form $\mathcal{E}_L$ by

$$\mathcal{D}(\mathcal{E}_L) := \mathcal{D}((-L)^{1/2}), \qquad \mathcal{E}_L(u,v) := \langle (-L)^{1/2}u, (-L)^{1/2}v\rangle_{\mathcal{H}}. \tag{9.95}$$

By the spectral theorem this form is densely defined, symmetric, closed and nonnegative, and the correspondence constructed above is inverted. This completes the proof of the first representation theorem. □

9.4.2 Lower Bounded Forms and Friedrichs Extension

If $\mathcal{E}$ is not nonnegative but *lower bounded*, i.e.

$$\mathcal{E}(f,f) \geq -C\|f\|_2^2, \quad \text{for some } C \in \mathbb{R}, \tag{9.96}$$

then we can modify the form to make it nonnegative by setting

$$\mathcal{E}_C(f,g) := \mathcal{E}(f,g) + C\langle f,g\rangle. \tag{9.97}$$

The closedness of $\mathcal{E}_C$ implies that $\mathcal{E}$ is also closed in the sense of lower semibounded forms. In this case, the first representation theorem still applies and produces a unique self-adjoint operator L such that

$$\mathcal{E}(f,g) = \langle(-L)^{1/2}f, (-L)^{1/2}g\rangle - C\langle f,g\rangle. \tag{9.98}$$

If $\mathcal{E}$ is the restriction of a symmetric differential operator A defined on a dense subspace $\mathcal{D}_0 \subset L^2(E,\mu)$ (e.g. $C_c^\infty(E)$), then the self-adjoint operator associated with the closure of $\mathcal{E}$ is called the *Friedrichs extension* of A. Formally, if A is symmetric and lower bounded on $\mathcal{D}_0$, the Friedrichs extension A_F is the unique self-adjoint operator satisfying

$$\mathcal{D}(A_F^{1/2}) = \overline{\mathcal{D}_0}^{\|\cdot\|_{\mathcal{E}}}, \qquad \mathcal{E}(f,g) = \langle A_F^{1/2}f, A_F^{1/2}g\rangle - C\langle f,g\rangle. \tag{9.99}$$

9.4.3 Connection with Dirichlet Forms

In the context of diffusion processes, the symmetric closed form $\mathcal{E}$ represents the *Dirichlet form* associated to the infinitesimal generator of the process. The Friedrichs extension ensures the existence of a self-adjoint realization of the (possibly unbounded) operator describing the energy dissipation, even when the original domain is only dense and not closed. This framework provides a rigorous bridge between partial differential operators, spectral theory, and stochastic analysis.

9.5 Analyticity, Sectorial Forms, and Holomorphic Semigroups

Let $\mathcal{H} = L^2(E,\mu)$ be a complex Hilbert space. We consider a sesquilinear form $\mathcal{E} : \mathcal{D}(\mathcal{E}) \times \mathcal{D}(\mathcal{E}) \to \mathbb{C}$ that is densely defined, closed, and *sectorial*. That is, there exist constants $\theta \in [0, \frac{\pi}{2})$ and $C > 0$ such that for all $f,g \in \mathcal{D}(\mathcal{E})$:

$$|\operatorname{Im}\mathcal{E}(f,f)| \leq \tan(\theta)\ \operatorname{Re}\mathcal{E}(f,f), \qquad \text{and} \qquad |\mathcal{E}(f,g)| \leq C\,[\operatorname{Re}\mathcal{E}(f,f)]^{1/2}[\operatorname{Re}\mathcal{E}(g,g)]^{1/2}. \tag{9.100}$$

In this case, the numerical range of $\mathcal{E}$ lies in the closed sector

$$\Sigma_\theta := \{z \in \mathbb{C} : |\arg z| \leq \theta\}, \tag{9.101}$$

and the form $\mathcal{E}$ is said to be *sectorial with semi-angle θ*.

Closability and the Associated Operator: If $\mathcal{E}$ satisfies the sector condition (9.100) and is closable, its closure $(\overline{\mathcal{E}}, \mathcal{D}(\overline{\mathcal{E}}))$ defines a unique *m-sectorial* operator L on $\mathcal{H}$ via the *second representation theorem* (Kato):

$$\mathcal{E}(f,g) = \langle Lf, g\rangle, \qquad \forall f \in \mathcal{D}(L),\ \forall g \in \mathcal{D}(\mathcal{E}). \tag{9.102}$$

Here, L is the maximal accretive (m-accretive) operator associated to $\mathcal{E}$, characterized by

$$\operatorname{Re}\langle Lf, f\rangle \geq 0, \qquad \forall f \in \mathcal{D}(L), \tag{9.103}$$

and $\lambda I + L$ is surjective for some (hence all) $\lambda > 0$.

M-Sectorial Operators and Holomorphic Semigroups: An operator L on $\mathcal{H}$ is said to be *m-sectorial* if it is closed, densely defined, and satisfies:

1. Its numerical range satisfies

$$W(L) := \{\langle Lf, f\rangle : f \in \mathcal{D}(L),\ \|f\| = 1\} \subseteq \Sigma_\theta, \tag{9.104}$$

   for some $\theta \in [0, \pi/2)$;

2. The range condition holds:

$$\operatorname{Range}(\lambda I + L) = \mathcal{H}, \qquad \forall \lambda > 0. \tag{9.105}$$

A fundamental theorem of semigroup theory (Kato, 1953) states that if L is m-sectorial, then $-L$ generates a strongly continuous *holomorphic semigroup* $(T_t)_{t\geq 0}$ satisfying

$$T_t = e^{-tL}, \qquad \text{and} \qquad T_z = e^{-zL} \text{ is holomorphic in } z \in \Sigma_{\frac{\pi}{2}-\theta}. \tag{9.106}$$

Thus, the analyticity of (T_t) in a sector of the complex plane is a direct consequence of the sectoriality of $\mathcal{E}$ (or equivalently, L).

Angle of Analyticity and Norm Bounds: For L m-sectorial with semi-angle $\theta < \frac{\pi}{2}$, the generated semigroup (T_t) extends analytically to the sector

$$S_{\frac{\pi}{2}-\theta} := \{z \in \mathbb{C} : |\arg z| < \tfrac{\pi}{2} - \theta\}, \tag{9.107}$$

and satisfies the bound

$$\|T_z\| \leq e^{-\operatorname{Re}(z)\,\omega}, \qquad z \in S_{\frac{\pi}{2}-\theta}, \tag{9.108}$$

for some $\omega \in \mathbb{R}$ depending on the lower bound of $\operatorname{Re}\mathcal{E}(f,f)$. In particular, (T_t) is analytic on the real half-axis and exponentially bounded.

Examples (Non-Symmetric Diffusion Operators): Sectorial forms naturally arise in the study of elliptic operators with drift. For instance, consider on $\mathcal{H} = L^2(\mathbb{R}^d, dx)$ the bilinear form

$$\mathcal{E}(f,g) = \int_{\mathbb{R}^d} \nabla f(x) \cdot \nabla \overline{g(x)}\, dx + \int_{\mathbb{R}^d} b(x) \cdot \nabla f(x)\, \overline{g(x)}\, dx, \tag{9.109}$$

with $\operatorname{div} b = 0$ and $b \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$. The second term introduces a non-symmetric component but satisfies the sector condition provided $\|b\|_\infty$ is sufficiently small relative to the diffusion coefficient. The associated operator

$$Lf = \Delta f - b \cdot \nabla f \tag{9.110}$$

is m-sectorial and thus generates a holomorphic semigroup $(T_t)_{t\geq 0}$ on $L^2(\mathbb{R}^d)$, corresponding to the non-symmetric diffusion process with drift b.

Summary: The sector condition ensures that the imaginary part of $\mathcal{E}(f,f)$ is controlled by its real part, guaranteeing accretivity and analyticity of the generated semigroup. Hence, the passage from symmetric to sectorial forms generalizes the Dirichlet form framework to non-symmetric diffusion operators, allowing one to rigorously treat elliptic operators with drift or complex coefficients within the theory of holomorphic semigroups.

9.6 Spectral Measures, Functional Calculus, and Resolvent Identities

Let L be a self-adjoint operator on the Hilbert space $\mathcal{H} = L^2(E, \mu)$. By the spectral theorem there exists a unique projection-valued measure $E(\cdot)$ defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ such that for every bounded Borel measurable function $\varphi : \mathbb{R} \to \mathbb{C}$ the operator

$$\varphi(L) := \int_{\mathbb{R}} \varphi(\lambda)\, dE(\lambda) \tag{9.111}$$

is a bounded linear operator on $\mathcal{H}$ and the map $\varphi \mapsto \varphi(L)$ is a *-homomorphism from $B_b(\mathbb{R})$ (the bounded Borel functions) into $\mathcal{B}(\mathcal{H})$. Moreover, for each $f \in \mathcal{H}$ the complex measure $\mu_f(A) := \langle E(A)f, f\rangle$ is finite and

$$\|\varphi(L)f\|^2 = \int_{\mathbb{R}} |\varphi(\lambda)|^2\, d\mu_f(\lambda). \tag{9.112}$$

The spectral representation yields the following important special cases (we record them in the form convenient for semigroup and resolvent theory). If $\sigma(L)$ denotes the spectrum of L, then for $\alpha \in \mathbb{C} \setminus \sigma(L)$ the resolvent $(\alpha I - L)^{-1}$ exists and admits the spectral integral representation

$$R(\alpha) := (\alpha I - L)^{-1} = \int_{\sigma(L)} \frac{1}{\alpha - \lambda}\, dE(\lambda). \tag{9.113}$$

Since the integrand $\lambda \mapsto (\alpha - \lambda)^{-1}$ is bounded on $\sigma(L)$, the operator on the right-hand side is bounded and equal to the resolvent. In particular, for α in the resolvent set one has the norm bound

$$\|R(\alpha)\| \leq \sup_{\lambda \in \sigma(L)} \frac{1}{|\alpha - \lambda|}. \tag{9.114}$$

Resolvent identity. For any $\alpha, \beta \in \mathbb{C} \setminus \sigma(L)$ the following algebraic identity holds:

$$R(\alpha) - R(\beta) = (\beta - \alpha) R(\alpha) R(\beta). \tag{9.115}$$

Proof: using the spectral representation,

$$\begin{aligned} R(\alpha) - R(\beta) &= \int_{\sigma(L)} \Big(\frac{1}{\alpha-\lambda} - \frac{1}{\beta-\lambda}\Big)\, dE(\lambda) = \int_{\sigma(L)} \frac{\beta-\alpha}{(\alpha-\lambda)(\beta-\lambda)}\, dE(\lambda) \\ &= (\beta-\alpha)\int_{\sigma(L)} \frac{1}{\alpha-\lambda}\frac{1}{\beta-\lambda}\, dE(\lambda) = (\beta-\alpha)R(\alpha)R(\beta), \end{aligned} \tag{9.116}$$

where the last equality follows from the multiplicative property of the functional calculus for bounded Borel functions. This identity is the cornerstone of resolvent calculus and yields many classical consequences (e.g. resolvent expansions, perturbation formulae).

Functional calculus and algebraic properties. The spectral calculus preserves algebraic operations and adjoints: if $\varphi, \psi \in B_b(\mathbb{R})$ then

$$\varphi(L)\psi(L) = (\varphi\psi)(L), \qquad \varphi(L)^* = \overline{\varphi}(L). \tag{9.117}$$

In particular, exponential and fractional powers are defined via the spectral integral:

$$e^{tL} = \int_{\sigma(L)} e^{t\lambda}\, dE(\lambda), \qquad (-L)^\alpha = \int_{\sigma(L)} (-\lambda)^\alpha\, dE(\lambda), \tag{9.118}$$

whenever the right-hand side is well-defined (for the exponential the integrand is bounded on any spectral subset bounded above).

Semigroup generation via spectral calculus. If L is self-adjoint and bounded above by some constant ω (for instance, when $-L$ is nonnegative), then the family $(T_t)_{t\geq 0}$ defined by

$$T_t := e^{tL} = \int_{\sigma(L)} e^{t\lambda}\, dE(\lambda) \tag{9.119}$$

is a strongly continuous one-parameter semigroup with generator L. Strong continuity follows from dominated convergence applied to the spectral integral: for any fixed $f \in \mathcal{H}$,

$$\|T_t f - f\|^2 = \int_{\sigma(L)} |e^{t\lambda} - 1|^2\, d\mu_f(\lambda) \to 0 \quad (t \downarrow 0), \tag{9.120}$$

because $|e^{t\lambda} - 1|^2 \leq C|t|^2(1+\lambda^2)$ locally and μ_f has finite second moment on bounded spectral sets. In the common probabilistic convention where L is the (negative) generator of a contraction semigroup, one usually writes $T_t = e^{tL}$ with $\sigma(L) \subset (-\infty, 0]$, so T_t is contractive.

Laplace transform and inversion. The resolvent and semigroup are related by the (operator-valued) Laplace transform. Suppose L is such that e^{tL} is exponentially bounded, i.e. $\|e^{tL}\| \leq Me^{\omega t}$. Then for $\alpha > \omega$ the Laplace transform identity holds in the strong operator sense:

$$(\alpha I - L)^{-1} = \int_0^\infty e^{-\alpha t} e^{tL}\, dt, \tag{9.121}$$

where the Bochner integral converges in operator norm on $\mathcal{H}$ due to the exponential bound. The proof is immediate from the spectral representations: for α in the right half-plane,

$$\int_0^\infty e^{-\alpha t} e^{tL}\, dt = \int_0^\infty e^{-\alpha t}\Big(\int_{\sigma(L)} e^{t\lambda}\, dE(\lambda)\Big)\, dt = \int_{\sigma(L)} \Big(\int_0^\infty e^{-(\alpha-\lambda)t}\, dt\Big)\, dE(\lambda) = \int_{\sigma(L)} \frac{1}{\alpha-\lambda}\, dE(\lambda), \tag{9.122}$$

justifying the resolvent representation. Conversely, given the resolvent one recovers the semigroup by the Bromwich (inverse Laplace) transform in the strong operator topology:

$$e^{tL} = \frac{1}{2\pi i} \lim_{R\to\infty} \int_{\gamma_R} e^{t\alpha}(\alpha I - L)^{-1}\, d\alpha, \tag{9.123}$$

where γ_R is a suitable vertical contour in the resolvent set; on the level of spectral integrals this inversion reduces to the elementary scalar inversion of Laplace transforms and is thus justified by dominated convergence.

Trotter product and resolvent approximations. The resolvent identity and functional calculus allow one to derive approximation formulas. For instance the exponential can be obtained as the strong limit of resolvent powers (Yosida approximants):

$$e^{tL} = \lim_{n\to\infty} \Big(I - \frac{t}{n}L\Big)^{-n}, \tag{9.124}$$

which again follows by applying the scalar identity $(1 - \frac{t}{n}\lambda)^{-n} \to e^{t\lambda}$ for each $\lambda \in \sigma(L)$ together with dominated convergence under the spectral measure.

Consequences and applications. These spectral and resolvent relations are the analytic backbone of the Feynman–Kac correspondence: the pathwise representation of e^{tL} as an expectation of multiplicative functionals is consistent with the spectral representation, and spectral properties (such as spectral gaps, eigenfunction expansions, trace-class criteria) may be translated into probabilistic statements (large-time decay, ground-state dominance, heat-kernel asymptotics). The functional calculus also enables perturbation analysis, spectral projections, and precise control of operator norms used throughout semigroup and Dirichlet-form theory.

9.7 Spectral Decomposition of Schrödinger Operators

Let H denote the Schrödinger operator

$$H := -\tfrac{1}{2}\Delta + V(x) \tag{9.125}$$

acting in the Hilbert space $L^2(\mathbb{R}^d)$, where the real-valued potential $V : \mathbb{R}^d \to \mathbb{R}$ satisfies $V(x) \geq V_0$ for some $V_0 \in \mathbb{R}$ and is chosen so that the quadratic form

$$\mathcal{E}^V(f,f) := \tfrac{1}{2}\int_{\mathbb{R}^d} |\nabla f(x)|^2\, dx + \int_{\mathbb{R}^d} V(x)|f(x)|^2\, dx, \qquad f \in C_c^\infty(\mathbb{R}^d), \tag{9.126}$$

is closable and lower semibounded. Passing to the closure $(\mathcal{E}^V, \mathcal{D}(\mathcal{E}^V))$ and applying the first representation theorem yields a unique self-adjoint operator H (the Friedrichs extension of the formal differential expression) with

$$\mathcal{D}(\mathcal{E}^V) = \mathcal{D}((-H)^{1/2}), \qquad \mathcal{E}^V(f,f) = \|(-H)^{1/2}f\|_2^2, \quad f \in \mathcal{D}(\mathcal{E}^V). \tag{9.127}$$

In particular H is bounded from below and $\sigma(H) \subset [\lambda_{\inf}, \infty)$ for some $\lambda_{\inf} \in \mathbb{R}$.

By the spectral theorem for self-adjoint operators there exists a unique projection-valued measure $E^H(\cdot)$ on the Borel σ-algebra of $\mathbb{R}$ such that for any bounded Borel function $\varphi : \mathbb{R} \to \mathbb{C}$

the operator $\varphi(H)$ is defined by the spectral integral

$$\varphi(H) = \int_{\mathbb{R}} \varphi(\lambda)\, dE^H(\lambda). \tag{9.128}$$

Hence one obtains the spectral decomposition

$$H = \int_{\mathbb{R}} \lambda\, dE^H(\lambda) \tag{9.129}$$

and, for the heat (Schrödinger) semigroup,

$$e^{-tH} = \int_{\mathbb{R}} e^{-t\lambda}\, dE^H(\lambda), \qquad t \geq 0. \tag{9.130}$$

The spectral measure E^H decomposes the spectrum $\sigma(H)$ into mutually singular parts

$$\sigma(H) = \sigma_{\mathrm{pp}}(H) \;\cup\; \sigma_{\mathrm{sc}}(H) \;\cup\; \sigma_{\mathrm{ac}}(H), \tag{9.131}$$

corresponding respectively to the pure point (eigenvalues), singular continuous, and absolutely continuous spectral components.

The semigroup representation yields the Laplace transform identity for matrix elements: for any $f, g \in L^2(\mathbb{R}^d)$ and α larger than the lower spectral bound,

$$\langle (\alpha I + H)^{-1} f, g\rangle = \int_0^\infty e^{-\alpha t} \langle e^{-tH} f, g\rangle\, dt = \int_{\mathbb{R}} \frac{1}{\alpha + \lambda}\, d\langle E^H(\lambda) f, g\rangle, \tag{9.132}$$

so that resolvent, semigroup and spectral measure are equivalent encodings of the operator H.

The Feynman–Kac formula furnishes a concrete probabilistic realization of the spectral measure via the semigroup kernel. Under hypotheses ensuring the Feynman–Kac representation (for example V Kato-class or form-bounded perturbation), the semigroup admits an integral kernel $K_t(x, y)$ satisfying

$$(e^{-tH} f)(x) = \int_{\mathbb{R}^d} K_t(x, y)\, f(y)\, dy, \qquad K_t(x, y) = p_t(x, y)\, \mathbb{E}_t^{x,y}\Big[e^{-\int_0^t V(X_s)\, ds}\Big], \tag{9.133}$$

where $p_t(x, y)$ is the Gaussian heat kernel and $\mathbb{E}_t^{x,y}$ denotes expectation with respect to the Brownian bridge from x to y over time t. The kernel $K_t(x, y)$ is related to the spectral measure by the inverse Laplace (or spectral) transform:

$$K_t(x, y) = \int_{\mathbb{R}} e^{-t\lambda}\, d\mu_{x,y}^H(\lambda), \qquad \mu_{x,y}^H(A) := \langle E^H(A)\delta_y, \delta_x\rangle, \tag{9.134}$$

with δ_x denoting the Dirac distribution at x (understood in the sense of distributional kernels when legitimate). Consequently, spectral properties (e.g. existence of a ground state, spectral gap, density of states) have immediate probabilistic manifestations in the long-time behaviour, trace asymptotics, and large deviations of the path measures.

When the semigroup e^{-tH} is compact (for instance, due to confining growth of V at infinity or working on a bounded domain with Dirichlet boundary), the spectrum is purely discrete and one has a complete eigen-expansion. Concretely, if $\sigma(H) = \{\lambda_n\}_{n=1}^{\infty}$ with $\lambda_1 \le \lambda_2 \le \cdots$ and $\lambda_n \to \infty$, and $\{\phi_n\}$ is an orthonormal basis of eigenfunctions, then

$$K_t(x,y) = \sum_{n=1}^{\infty} e^{-t\lambda_n} \phi_n(x)\overline{\phi_n(y)}, \tag{9.135}$$

with the series converging uniformly for $t \ge t_0 > 0$ and in L^2-sense for each fixed $t > 0$. In particular,

$$(e^{-tH} f)(x) = \sum_{n=1}^{\infty} e^{-t\lambda_n} \langle f, \phi_n\rangle \phi_n(x). \tag{9.136}$$

The ground state ϕ_1 corresponding to the lowest eigenvalue λ_1 can be chosen strictly positive (Perron–Frobenius principle / positivity-improving property of e^{-tH} when applicable), and spectral gap estimates $\lambda_2 - \lambda_1 > 0$ govern the exponential rate of convergence to equilibrium in L^2 and pointwise senses:

$$\lim_{t\to\infty} e^{t\lambda_1} K_t(x,y) = \phi_1(x)\overline{\phi_1(y)}. \tag{9.137}$$

In the presence of absolutely continuous spectrum, one typically has no eigenfunction expansion but rather a spectral resolution in terms of generalized eigenfunctions or scattering states; the semigroup still admits the spectral integral representation

$$K_t(x,y) = \int_{\sigma(H)} e^{-t\lambda}\, \rho_{x,y}(\lambda)\, d\lambda, \tag{9.138}$$

where $\rho_{x,y}$ denotes the density of the spectral measure with respect to Lebesgue measure on the absolutely continuous component. This representation underpins dispersive estimates and the short-time/large-energy asymptotics of the kernel via stationary phase and semiclassical analysis.

Trace-class and Schatten-class properties of e^{-tH} are read off from the spectral side: e^{-tH} is trace-class if and only if

$$\mathrm{Tr}(e^{-tH}) = \sum_n e^{-t\lambda_n} < \infty, \tag{9.139}$$

which holds e.g. when V is confining or when one restricts to compact domains. The trace yields the partition function in statistical mechanics and is connected to heat-kernel asymptotics and spectral invariants via the Minakshisundaram–Pleijel expansion or the Weyl law:

$$\mathrm{Tr}(e^{-tH}) \sim (4\pi t)^{-d/2} \int_{\mathbb{R}^d} e^{-tV(x)}\, dx + \cdots, \qquad t \downarrow 0, \tag{9.140}$$

under suitable regularity/growth assumptions on V.

The functional calculus for H permits one to translate analytic operations directly into probabilistic statements. For example, for $f \in L^2$ and measurable $\phi \ge 0$,

$$\langle \phi(H) f, f\rangle = \int_{\sigma(H)} \phi(\lambda)\, d\langle E^H(\lambda) f, f\rangle = \int_0^{\infty} \phi(\lambda)\, d\mu_f^H(\lambda), \tag{9.141}$$

and when $\phi(\lambda) = e^{-t\lambda}$ the right-hand side equals $\langle e^{-tH} f, f\rangle = \mathbb{E}[\cdots]$ via Feynman–Kac. This equivalence is the cornerstone of the analytic–probabilistic duality: spectral information (eigenvalues, spectral measures, resolvent poles) corresponds to pathwise phenomena (long-time decay, ground-state dominance, resonance behaviour).

Finally, perturbation theory (Kato–Rellich, Birman–Schwinger) and semiclassical analysis connect fine spectral features of H to the small-noise / large-parameter asymptotics of the path integrals. For instance, the Birman–Schwinger principle relates eigenvalues below the essential spectrum to compactness properties of a suitable integral operator built from the free resolvent and the potential V, while semiclassical (WKB) methods approximate high-energy spectral densities by phase-space integrals, which also admit probabilistic interpretations via large-deviation principles for the Brownian motion path measures.

In summary, the spectral decomposition of H and the Feynman–Kac representation are two equivalent and complementary lenses: the former is the algebraic/spectral description furnished by functional calculus, and the latter is the analytic/probabilistic description furnished by expectations over path space. Each perspective supplies tools and estimates that translate directly into the other, yielding a rich and exact correspondence between spectral theory, PDE analysis, potential theory, and stochastic processes.

9.8 Conclusion

In this chapter, we examined the functional analytic structure underlying Feynman–Kac semigroups and their generators. The discussion unified semigroup theory, spectral theory, and Sobolev space analysis to characterize the mapping and regularity properties of the associated operators.

We began by investigating the domain of the generator $H = -\frac{1}{2}\Delta_g + V$ as a self-adjoint operator on $L^2(M, d\mathrm{vol}_g)$, ensuring that H admits a well-defined spectral decomposition. Using this, we connected the Feynman–Kac semigroup e^{-tH} to the spectral resolution of H via

$$e^{-tH} = \int_0^\infty e^{-t\lambda}\, dE_\lambda, \tag{9.142}$$

where E_λ is the spectral measure. This relation provided a precise quantitative framework for understanding diffusion smoothing, spectral gaps, and the asymptotic decay of heat kernels. We then analyzed compactness, ultracontractivity, and trace-class properties of e^{-tH}, relating them to geometric quantities such as Ricci curvature and volume growth. The equivalence between functional inequalities (Poincaré, log-Sobolev) and analytic semigroup properties was established, showing that curvature bounds yield precise L^p-regularization and entropy dissipation estimates.

In conclusion, this chapter provided the deepest analytic insight into how geometry, probability, and spectral theory intertwine in the study of diffusion semigroups. Through the functional analytic lens, the Feynman–Kac representation emerges as a unifying theorem connecting path expectations, operator theory, and geometric analysis on manifolds.

References: Reed and Simon (1980) [21], Davies (1989) [13], DaPrato and Zabczyk (1992) [10].

Part IV

Generalizations and Modern Extensions

Chapter 10

Extensions to Lévy and Jump Processes

"Every discontinuity hides a structure; every jump conceals a law."
— Paul Lévy

10.1 Motivation and General Framework

The analysis of diffusion semigroups and their generators naturally extends to stochastic processes with jumps. While Brownian motion yields continuous trajectories governed by local operators such as the Laplacian, jump processes correspond to nonlocal generators — operators involving integro-differential terms. This framework unifies diffusion and jump phenomena under the general theory of Lévy processes, which serve as the building blocks of all stochastic processes with stationary independent increments.

A Lévy process $(X_t)_{t\geq 0}$ on $\mathbb{R}^d$ is characterized by three components: a drift vector $b \in \mathbb{R}^d$, a covariance matrix $Q \in \mathbb{R}^{d\times d}$, and a Lévy measure ν on $\mathbb{R}^d \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}^d\setminus\{0\}} \min(1, |x|^2)\, \nu(dx) < \infty. \tag{10.1}$$

These data determine both the infinitesimal generator of the process and the structure of its sample paths.

10.2 Lévy–Khintchine Representation

The Lévy–Khintchine representation theorem provides the most general analytic characterization of infinitely divisible probability distributions and, equivalently, of Lévy processes. A *Lévy process* $(X_t)_{t\geq 0}$ on $\mathbb{R}^d$ is a stochastic process satisfying:

$$X_0 = 0 \text{ a.s.}, \quad X_t - X_s \text{ is independent of } \mathcal{F}_s, \quad X_t - X_s \sim X_{t-s}, \quad \text{and } X_t \text{ is stochastically continuous.} \tag{10.2}$$

That is, X_t has stationary and independent increments with càdlàg sample paths.

For such a process, each random variable X_t is *infinitely divisible*, meaning that for any $n \in \mathbb{N}$, there exist i.i.d. random variables $Y_1, \ldots, Y_n$ such that

$$X_t \overset{d}{=} Y_1 + \cdots + Y_n. \tag{10.3}$$

Hence, the characteristic function $\phi_t(\xi) = \mathbb{E}\big[e^{i\xi \cdot X_t}\big]$ satisfies

$$\phi_t(\xi) = e^{-t\psi(\xi)}, \qquad \xi \in \mathbb{R}^d, \tag{10.4}$$

where $\psi : \mathbb{R}^d \to \mathbb{C}$ is called the *characteristic exponent* or *Lévy symbol*. The mapping $t \mapsto \phi_t(\xi)$ satisfies the semigroup property

$$\phi_{s+t}(\xi) = \phi_s(\xi)\phi_t(\xi), \tag{10.5}$$

which implies that $\psi(\xi)$ uniquely determines the process.

The *Lévy–Khintchine formula* asserts that any function $\psi(\xi)$ of the form

$$\psi(\xi) = ib \cdot \xi + \frac{1}{2}\xi^T Q\xi + \int_{\mathbb{R}^d \setminus \{0\}} \Big(1 - e^{i\xi \cdot x} + i\xi \cdot x\, \mathbf{1}_{\{|x|<1\}}\Big)\, \nu(dx) \tag{10.6}$$

corresponds to the characteristic exponent of a Lévy process, and conversely, every Lévy process admits such a representation. Here:

- $b \in \mathbb{R}^d$ is the *drift vector*;
- $Q \in \mathbb{R}^{d \times d}$ is a symmetric nonnegative-definite matrix representing the Gaussian covariance;
- ν is a *Lévy measure* on $\mathbb{R}^d \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}^d \setminus \{0\}} \min(1, |x|^2)\, \nu(dx) < \infty. \tag{10.7}$$

The three terms in $\psi(\xi)$ have distinct probabilistic interpretations:

1. $ib \cdot \xi$ corresponds to the deterministic *drift component*;
2. $\frac{1}{2}\xi^T Q\xi$ corresponds to the *Gaussian diffusion component*;
3. the integral term represents the *jump component*, describing random discontinuities with intensity governed by ν.

The Lévy measure ν encodes the frequency and magnitude of jumps, with small jumps regularized by the term $i\xi \cdot x\, \mathbf{1}_{\{|x|<1\}}$ to ensure convergence of the integral. If $\nu = 0$, the process reduces to a Gaussian process with drift and covariance Q.

The infinitesimal generator L of the Lévy process, acting on test functions $f \in C_c^\infty(\mathbb{R}^d)$, is then given by the Lévy–Itô operator:

$$(Lf)(x) = b \cdot \nabla f(x) + \frac{1}{2}\,\mathrm{Tr}(Q\nabla^2 f(x)) + \int_{\mathbb{R}^d \setminus \{0\}} (f(x+y) - f(x) - y \cdot \nabla f(x)\, \mathbf{1}_{\{|y|<1\}})\nu(dy). \tag{10.8}$$

This operator is the infinitesimal generator of the strongly continuous contraction semigroup $(P_t)_{t\geq 0}$ defined by

$$(P_t f)(x) = \mathbb{E}^x[f(X_t)] = \int_{\mathbb{R}^d} f(x+y)\, p_t(dy), \tag{10.9}$$

where p_t is the transition probability measure of X_t.

Finally, the spectral symbol of the generator in Fourier space satisfies

$$\widehat{Lf}(\xi) = -\psi(\xi)\, \widehat{f}(\xi), \tag{10.10}$$

showing that L acts as a pseudo-differential operator with symbol $-\psi(\xi)$. Hence, the Lévy–Khintchine formula not only characterizes the law of Lévy processes but also provides the analytic bridge between probability theory and harmonic analysis via the spectral representation of nonlocal generators.

10.3 Nonlocal Dirichlet Forms

In the context of symmetric Lévy processes, the infinitesimal generator is inherently *nonlocal*, reflecting the possibility of instantaneous jumps from one spatial location to another. The associated bilinear Dirichlet form captures this nonlocality through its dependence on spatial increments rather than derivatives.

Let X_t be a symmetric Lévy process on $\mathbb{R}^d$ with Lévy triplet $(0, 0, \nu)$, where ν is a symmetric Lévy measure satisfying

$$\nu(A) = \nu(-A), \qquad \forall A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}), \tag{10.11}$$

and

$$\int_{\mathbb{R}^d\setminus\{0\}} \min(1, |y|^2)\, \nu(dy) < \infty. \tag{10.12}$$

The symmetry condition implies that the process X_t has no drift or Gaussian component and that its transition probabilities are symmetric, i.e.,

$$p_t(x, y) = p_t(y, x), \qquad \forall x, y \in \mathbb{R}^d. \tag{10.13}$$

The corresponding Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ acting on $L^2(\mathbb{R}^d, dx)$ is defined by

$$\mathcal{E}(f, g) = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \big(f(x+y) - f(x)\big)\big(g(x+y) - g(x)\big)\, \nu(dy)\, dx, \tag{10.14}$$

with the form domain

$$\mathcal{D}(\mathcal{E}) = \left\{ f \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x+y) - f(x)|^2\, \nu(dy)\, dx < \infty \right\}. \tag{10.15}$$

This form arises from the L^2-limit of the energy associated with the process, and it satisfies the following fundamental analytic properties:

Symmetry: For all $f, g \in \mathcal{D}(\mathcal{E})$,

$$\mathcal{E}(f,g) = \mathcal{E}(g,f), \tag{10.16}$$

which follows directly from the symmetry of the Lévy measure $\nu(A) = \nu(-A)$ and the change of variables $y \mapsto -y$ in the integral.

Positivity: For all $f \in \mathcal{D}(\mathcal{E})$,

$$\mathcal{E}(f,f) = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x+y) - f(x)|^2 \, \nu(dy) \, dx \geq 0, \tag{10.17}$$

so that the form is nonnegative definite.

Closedness: The form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is *closed*, meaning that if $(f_n)_{n\in\mathbb{N}} \subset \mathcal{D}(\mathcal{E})$ satisfies

$$f_n \to f \text{ in } L^2(\mathbb{R}^d) \quad \text{and} \quad \mathcal{E}(f_n - f_m, f_n - f_m) \to 0 \text{ as } n, m \to \infty, \tag{10.18}$$

then $f \in \mathcal{D}(\mathcal{E})$ and $\mathcal{E}(f_n - f, f_n - f) \to 0$. This follows from Fatou's lemma and the completeness of $L^2(\mathbb{R}^d)$.

Markov Property: For any $f \in \mathcal{D}(\mathcal{E})$ and the normal contraction $\tilde{f} = (0 \vee f) \wedge 1$, one has $\tilde{f} \in \mathcal{D}(\mathcal{E})$ and

$$\mathcal{E}(\tilde{f}, \tilde{f}) \leq \mathcal{E}(f,f). \tag{10.19}$$

Thus, $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a *Dirichlet form* in the sense of Fukushima, Oshima, and Takeda.

Associated Nonlocal Generator: The self-adjoint operator L associated with $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ via the first representation theorem satisfies

$$\mathcal{E}(f,g) = \langle -Lf, g \rangle_{L^2(\mathbb{R}^d)}, \qquad f \in \mathcal{D}(L),\ g \in \mathcal{D}(\mathcal{E}), \tag{10.20}$$

and for $f \in C_c^\infty(\mathbb{R}^d)$,

$$(Lf)(x) = \int_{\mathbb{R}^d} \left(f(x+y) - f(x)\right) \nu(dy), \tag{10.21}$$

or, after compensation for small jumps (ensuring convergence),

$$(Lf)(x) = \int_{\mathbb{R}^d \setminus \{0\}} \left(f(x+y) - f(x) - y \cdot \nabla f(x) \, \mathbf{1}_{\{|y|<1\}}\right) \nu(dy). \tag{10.22}$$

Fourier Representation: By Parseval's identity, one obtains the equivalent spectral characterization:

$$\mathcal{E}(f,f) = \int_{\mathbb{R}^d} \psi(\xi) \, |\widehat{f}(\xi)|^2 \, d\xi, \tag{10.23}$$

where $\psi(\xi)$ is the characteristic exponent associated with ν,

$$\psi(\xi) = \int_{\mathbb{R}^d \setminus \{0\}} \left(1 - \cos(\xi \cdot y)\right) \nu(dy). \tag{10.24}$$

Hence, the Dirichlet form $\mathcal{E}$ fully captures the nonlocal jump structure of the process in both spatial and Fourier domains. By the general theory of Dirichlet forms, $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ generates a symmetric Hunt process with càdlàg trajectories and jump kernel ν, thereby providing a complete analytic characterization of symmetric pure-jump Lévy processes.

10.4 Fractional Laplacian as a Paradigm

A central and paradigmatic example of a nonlocal generator arising from a symmetric Lévy process is the *fractional Laplacian* $(-\Delta)^{\alpha/2}$, where the order of differentiation $\alpha \in (0,2)$ determines the jump intensity and tail behavior of the corresponding Lévy process. This operator generalizes the classical Laplacian to the nonlocal setting and plays a fundamental role in probability theory, harmonic analysis, and the theory of partial differential equations.

Definition via Fourier Symbol: For $f \in \mathcal{S}(\mathbb{R}^d)$, the Schwartz space of rapidly decaying smooth functions, the fractional Laplacian is defined in the Fourier domain by

$$\mathcal{F}\big[(-\Delta)^{\alpha/2} f\big](\xi) = |\xi|^{\alpha}\, \widehat{f}(\xi), \tag{10.25}$$

where

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} e^{-i\xi\cdot x} f(x)\, dx \tag{10.26}$$

is the Fourier transform of f. This definition exhibits the operator as a pseudodifferential operator with symbol $|\xi|^{\alpha}$, showing that $(-\Delta)^{\alpha/2}$ is positive, self-adjoint, and homogeneous of order α.

Singular Integral Representation: The same operator admits a real-space representation as a singular integral:

$$(-\Delta)^{\alpha/2} f(x) = C_{d,\alpha}\, \text{P.V.} \int_{\mathbb{R}^d} \frac{f(x) - f(y)}{|x-y|^{d+\alpha}}\, dy, \tag{10.27}$$

where P.V. denotes the Cauchy principal value and the normalization constant $C_{d,\alpha} > 0$ is explicitly given by

$$C_{d,\alpha} = \frac{2^{\alpha-1}\alpha\, \Gamma\!\left(\frac{d+\alpha}{2}\right)}{\pi^{d/2}\Gamma\!\left(1-\frac{\alpha}{2}\right)}. \tag{10.28}$$

This normalization ensures the identity

$$\mathcal{F}\big[(-\Delta)^{\alpha/2} f\big](\xi) = |\xi|^{\alpha} \widehat{f}(\xi) \tag{10.29}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$. The integral form reveals the *nonlocality* of $(-\Delta)^{\alpha/2}$: the value of the operator at a point x depends on all values of f throughout $\mathbb{R}^d$, with the kernel $|x-y|^{-d-\alpha}$ governing the strength of long-range interactions.

Dirichlet Form Representation: The corresponding bilinear Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ associated with the fractional Laplacian is defined by

$$\mathcal{E}(f,g) = \frac{C_{d,\alpha}}{2} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(f(x)-f(y))(g(x)-g(y))}{|x-y|^{d+\alpha}}\, dx\, dy, \tag{10.30}$$

with form domain

$$\mathcal{D}(\mathcal{E}) = \left\{ f \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|f(x)-f(y)|^2}{|x-y|^{d+\alpha}}\,dx\,dy < \infty \right\}. \tag{10.31}$$

The form $\mathcal{E}$ is symmetric, nonnegative, and closed, and hence by the first representation theorem there exists a unique self-adjoint operator $L = -(-\Delta)^{\alpha/2}$ such that

$$\mathcal{E}(f,g) = \langle (-L)^{1/2} f, (-L)^{1/2} g\rangle_{L^2(\mathbb{R}^d)}. \tag{10.32}$$

Energy Identity and Spectral Representation: In the Fourier domain, the energy functional can be equivalently expressed as

$$\mathcal{E}(f,f) = \int_{\mathbb{R}^d} |\xi|^\alpha\, |\widehat{f}(\xi)|^2\, d\xi. \tag{10.33}$$

This equality follows from Plancherel's theorem and provides the spectral characterization of the fractional Laplacian as the multiplication operator with symbol $|\xi|^\alpha$. It also implies that $(-\Delta)^{\alpha/2}$ is the infinitesimal generator of an α-stable Lévy process, whose characteristic exponent is precisely $\psi(\xi) = |\xi|^\alpha$.

Probabilistic Interpretation: Let X_t be a rotationally symmetric α-stable Lévy process on $\mathbb{R}^d$ with characteristic function

$$\mathbb{E}\Big[e^{i\xi\cdot X_t}\Big] = e^{-t|\xi|^\alpha}, \qquad \xi \in \mathbb{R}^d. \tag{10.34}$$

Then the infinitesimal generator of the associated semigroup $(T_t)_{t\geq 0}$, given by

$$(T_t f)(x) = \mathbb{E}_x[f(X_t)] = \int_{\mathbb{R}^d} p_t(x-y) f(y)\, dy, \tag{10.35}$$

is exactly $L = -(-\Delta)^{\alpha/2}$. Thus, the fractional Laplacian serves as the analytic generator of the isotropic α-stable process, linking nonlocal analysis and jump-type stochastic dynamics.

Variational Characterization: The space $\mathcal{D}(\mathcal{E})$, endowed with the inner product

$$\langle f,g\rangle_{\mathcal{E}} = \langle f,g\rangle_{L^2(\mathbb{R}^d)} + \mathcal{E}(f,g), \tag{10.36}$$

is a Hilbert space coinciding with the fractional Sobolev space $H^{\alpha/2}(\mathbb{R}^d)$. The energy form $\mathcal{E}$ thus represents the quadratic form associated with $(-\Delta)^{\alpha/2}$, providing both an analytic and probabilistic characterization of the fractional Laplacian.

Hence, $(-\Delta)^{\alpha/2}$ stands as the canonical prototype of a nonlocal, self-adjoint, and Markovian operator, embodying the structure of symmetric α-stable Lévy generators and their associated Dirichlet forms.

10.5 Feynman–Kac Representation with Jumps

In the nonlocal and jump-diffusion setting, the classical Feynman–Kac formula extends naturally to operators whose generators include both differential and integro-differential components. Let X_t be a symmetric Lévy process on $\mathbb{R}^d$ with generator L acting on $C_c^\infty(\mathbb{R}^d)$ by

$$(Lf)(x) = b \cdot \nabla f(x) + \frac{1}{2}\operatorname{Tr}(Q\nabla^2 f(x)) + \int_{\mathbb{R}^d\setminus\{0\}} \Big(f(x+y) - f(x) - y\cdot\nabla f(x)\mathbf{1}_{\{|y|<1\}}\Big)\,\nu(dy), \tag{10.37}$$

where $b \in \mathbb{R}^d$ is the drift vector, Q is a positive semidefinite covariance matrix, and ν is the Lévy measure satisfying

$$\int_{\mathbb{R}^d\setminus\{0\}} \min(1,|y|^2)\,\nu(dy) < \infty. \tag{10.38}$$

The process X_t has càdlàg (right-continuous with left limits) paths in the Skorokhod space $D([0,\infty),\mathbb{R}^d)$, and it can be decomposed as

$$X_t = bt + W_t + \int_0^t \int_{|y|<1} y\,(\tilde{N}(ds,dy)) + \int_0^t \int_{|y|\geq 1} y\,N(ds,dy), \tag{10.39}$$

where W_t is a Brownian motion, $N(ds,dy)$ is the Poisson random measure of jumps, and $\tilde{N}(ds,dy) = N(ds,dy) - \nu(dy)ds$ is the compensated Poisson measure.

Schrödinger-type Nonlocal Operators: Let $V : \mathbb{R}^d \to \mathbb{R}$ be a measurable potential function, bounded from below or belonging to the Kato class associated with X_t. We consider the nonlocal Schrödinger operator

$$H = -L + V, \tag{10.40}$$

initially defined on $C_c^\infty(\mathbb{R}^d)$, and then extended to a self-adjoint operator on $L^2(\mathbb{R}^d)$ via the Friedrichs extension of the associated quadratic form

$$\mathcal{E}^V(f,f) = \mathcal{E}(f,f) + \int_{\mathbb{R}^d} V(x)|f(x)|^2\,dx, \tag{10.41}$$

where $\mathcal{E}$ is the symmetric Dirichlet form corresponding to the Lévy process.

Generalized Feynman–Kac Formula: The semigroup $(e^{-tH})_{t\geq 0}$ generated by H admits a probabilistic representation in terms of expectations over the Lévy process:

$$(e^{-tH}f)(x) = \mathbb{E}^x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right], \qquad f \in L^2(\mathbb{R}^d). \tag{10.42}$$

This equality defines the generalized *Feynman–Kac formula with jumps*. The expectation is taken with respect to the law of the process X_t starting at x, and the exponential weight incorporates the pathwise accumulated potential energy $\int_0^t V(X_s)\,ds$.

Analytic Foundation: From the analytic perspective, the mapping

$$P_t^V f(x) := \mathbb{E}^x\left[e^{-\int_0^t V(X_s)\,ds} f(X_t)\right] \tag{10.43}$$

defines a strongly continuous contraction semigroup on $L^2(\mathbb{R}^d)$ under suitable assumptions on V, such as the Kato smallness condition

$$\lim_{t\downarrow 0} \sup_{x\in\mathbb{R}^d} \mathbb{E}^x\left[\int_0^t |V(X_s)|\, ds\right] = 0. \tag{10.44}$$

By the general semigroup theory, its generator coincides with the self-adjoint realization of $H = -L + V$. Hence,

$$P_t^V = e^{-tH}, \qquad \forall\, t \geq 0. \tag{10.45}$$

Path Integral over Skorokhod Space: The probabilistic representation can be rigorously viewed as an expectation over the Skorokhod path space $D([0,t],\mathbb{R}^d)$ endowed with the canonical measure induced by the Lévy process:

$$(e^{-tH} f)(x) = \int_{D([0,t],\mathbb{R}^d)} e^{-\int_0^t V(\omega(s))\, ds} f(\omega(t))\, \mathbb{P}^x(d\omega). \tag{10.46}$$

This integral defines a nonlocal analogue of the classical Feynman path integral, where the trajectories ω include both continuous Brownian paths and jump discontinuities governed by ν.

Analytic–Probabilistic Duality: The above representation preserves the fundamental duality between analytic and probabilistic descriptions:

- The analytic side is represented by the semigroup e^{-tH} acting on $L^2(\mathbb{R}^d)$, with infinitesimal generator $H = -L + V$.
- The probabilistic side is given by the expectation over Lévy trajectories, with the exponential weight incorporating the potential energy functional.

Even in the presence of jumps, this duality remains exact, thus extending the classical Feynman–Kac correspondence to the realm of nonlocal operators and discontinuous stochastic processes.

10.6 Generators with Drift and Jump–Diffusion Operators

In the most general setting of Markov processes with both continuous and discontinuous components, one encounters *jump–diffusion generators* that combine drift, diffusion, and jump contributions in a single integro-differential operator. These generators characterize the infinitesimal behavior of a large class of Itô semimartingales, encompassing both local (diffusion) and nonlocal (jump) phenomena.

General Form of the Generator: Let L be an operator acting on sufficiently smooth functions $f \in C_c^2(\mathbb{R}^d)$ by

$$(Lf)(x) = b(x)\cdot\nabla f(x) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \int_{\mathbb{R}^d\setminus\{0\}} \Big(f(x+y) - f(x) - y\cdot\nabla f(x)\mathbf{1}_{\{|y|<1\}}\Big)\, \nu(x,dy), \tag{10.47}$$

where:

- $b : \mathbb{R}^d \to \mathbb{R}^d$ is a measurable *drift field* describing the deterministic first-order motion;

- $a(x) = (a_{ij}(x))_{1\le i,j\le d}$ is a measurable, symmetric, positive semidefinite matrix field representing the local diffusion coefficients;
- $\nu(x, dy)$ is a *position-dependent Lévy kernel* (or *jump measure*) satisfying the integrability condition

$$\int_{\mathbb{R}^d\setminus\{0\}} (1 \wedge |y|^2)\, \nu(x, dy) < \infty, \qquad \forall\, x \in \mathbb{R}^d. \tag{10.48}$$

The truncation term $y \cdot \nabla f(x)\mathbf{1}_{\{|y|<1\}}$ guarantees the convergence of the integral near $y = 0$ by compensating the linear part of small jumps, thus ensuring that $Lf(x)$ is well-defined for twice continuously differentiable f.

Probabilistic Representation via Stochastic Differential Equations: The operator L in (10.47) arises as the infinitesimal generator of a *jump–diffusion process* $(X_t)_{t\ge 0}$ satisfying the stochastic differential equation

$$dX_t = b(X_{t-})\, dt + \sigma(X_{t-})\, dW_t + \int_{|y|<1} y\, \tilde{N}(dt, dy) + \int_{|y|\ge 1} y\, N(dt, dy), \tag{10.49}$$

where:

- W_t is a standard d-dimensional Brownian motion;
- $\sigma(x)$ is a measurable diffusion coefficient matrix satisfying $a(x) = \sigma(x)\sigma(x)^\top$;
- $N(dt, dy)$ is a Poisson random measure on $(0,\infty)\times(\mathbb{R}^d\setminus\{0\})$ with compensator $\nu(X_{t-}, dy)\, dt$;
- $\tilde{N}(dt, dy) = N(dt, dy) - \nu(X_{t-}, dy)\, dt$ denotes the compensated Poisson measure, ensuring martingale properties for small jumps.

The process X_t is a càdlàg semimartingale taking values in $\mathbb{R}^d$, and its infinitesimal generator is precisely L defined in (10.47). Formally, for all $f \in C_c^2(\mathbb{R}^d)$,

$$f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)\, ds \tag{10.50}$$

is a martingale with respect to the natural filtration of X_t.

Analytic Properties: The operator L is generally *non-symmetric* because of the drift term $b(x)\cdot\nabla f(x)$ and the possible asymmetry of $\nu(x, dy)$. However, it retains several essential analytic properties:

- L is *local* in the diffusion and drift parts but *nonlocal* in the jump component.
- When $b(x) \equiv 0$, $a(x) \equiv 0$, and $\nu(x, dy) = \nu(dy)$ is symmetric, L reduces to the generator of a pure Lévy process, often associated with a Dirichlet form.
- When $\nu(x, dy) \equiv 0$, L reduces to a second-order differential operator with drift, corresponding to an Itô diffusion.

The combination of these features makes L a unifying operator encompassing both continuous and jump-driven dynamics.

Domain and Regularity Considerations: To ensure that $Lf(x)$ is well-defined and belongs to $L^2_{\text{loc}}(\mathbb{R}^d)$, one typically requires that

$$b(\cdot) \in L^p_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^d), \quad a_{ij}(\cdot) \in L^\infty_{\text{loc}}(\mathbb{R}^d), \quad \text{and} \quad \nu(x, dy) \text{ measurable in } x. \tag{10.51}$$

If these conditions hold, the associated operator L generates a (possibly non-symmetric) strongly continuous semigroup $(P_t)_{t\geq 0}$ on $L^2(\mathbb{R}^d)$, which acts via

$$(P_t f)(x) = \mathbb{E}^x[f(X_t)]. \tag{10.52}$$

Interpretation and Applications: The generator (10.47) serves as the mathematical foundation for a wide variety of models:

- In mathematical finance, it appears in jump–diffusion models of asset prices (e.g., Merton's model, Kou's double-exponential jump diffusion).
- In physics, it describes transport phenomena with both diffusive and ballistic (jump) components.
- In biology and ecology, it models spatial population dynamics with long-range dispersal.

Hence, the operator L provides a *unifying analytical framework* for stochastic dynamics that exhibit both continuous diffusion and discontinuous jumps, bridging the theory of partial differential equations and the probabilistic theory of semimartingales.

10.7 Analytic and Probabilistic Duality

The profound connection between nonlocal analytic operators and stochastic processes is established through the framework of *Dirichlet forms* and their generalizations. When symmetry fails or nonlocal interactions dominate, one must extend the classical Dirichlet form theory to the broader setting of *semi-Dirichlet forms* and *sectorial operators*, which retain enough analytic structure to ensure the existence of an associated Markov process.

Semi-Dirichlet and Sectorial Forms: Let E be a locally compact separable metric space with a positive Radon measure m having full support. A bilinear form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(E; m)$ is called a *semi-Dirichlet form* if it satisfies the following conditions:

1. **Densely defined:** The domain $\mathcal{D}(\mathcal{E})$ is dense in $L^2(E; m)$.
2. **Closedness:** For any sequence $(f_n) \subset \mathcal{D}(\mathcal{E})$ with $f_n \to f$ in $L^2(E; m)$ and $\mathcal{E}(f_n - f_m, f_n - f_m) \to 0$, we have $f \in \mathcal{D}(\mathcal{E})$ and $\mathcal{E}(f_n - f, f_n - f) \to 0$.
3. **Sector condition:** There exists a constant $K \geq 0$ such that for all $f, g \in \mathcal{D}(\mathcal{E})$,

$$|\mathcal{E}(f, g) - \mathcal{E}(g, f)| \leq K \left[\mathcal{E}(f, f)^{1/2} \mathcal{E}(g, g)^{1/2} \right]. \tag{10.53}$$

   This ensures that the form is *sectorial*, i.e., its numerical range lies in a sector of the complex plane.

4. **Markov property:** For every $f \in \mathcal{D}(\mathcal{E})$, the truncated function $f^\sharp := (0 \vee f) \wedge 1$ also belongs to $\mathcal{D}(\mathcal{E})$, and
$$\mathcal{E}(f^\sharp, f^\sharp) \leq \mathcal{E}(f, f). \tag{10.54}$$

The pair $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is then associated with a (not necessarily symmetric) strongly continuous semigroup $(T_t)_{t\geq 0}$ on $L^2(E; m)$ satisfying the sectorial bound

$$\|T_t\|_{L^2\to L^2} \leq e^{Kt}, \qquad t \geq 0. \tag{10.55}$$

Beurling–Deny Representation: The central analytic structure underlying nonlocal and jump processes is provided by the *Beurling–Deny decomposition*, which expresses any regular semi-Dirichlet form as a sum of three mutually orthogonal components:

$$\mathcal{E}(f, g) = \mathcal{E}^{(c)}(f, g) + \int_{E\times E\setminus \mathrm{diag}} (f(x) - f(y))(g(x) - g(y))\, J(dx, dy) + \int_E f(x)g(x)\, k(dx), \tag{10.56}$$

for all $f, g \in \mathcal{D}(\mathcal{E})$, where:

- $\mathcal{E}^{(c)}$ is the *local (continuous)* part corresponding to diffusion; it is strongly local in the sense that

$$\mathcal{E}^{(c)}(f, g) = 0 \quad \text{whenever } f \text{ is constant on a neighborhood of } \mathrm{supp}(g); \tag{10.57}$$

- $J(dx, dy)$ is a symmetric, positive Radon measure on $E \times E \setminus \mathrm{diag}$, referred to as the *jump measure*, encoding the intensity and distribution of discontinuous jumps;
- $k(dx)$ is a positive Radon measure on E known as the *killing measure*, representing the rate at which the process is terminated (or killed).

The form (10.56) thus decomposes any semi-Dirichlet form into diffusion, jump, and killing contributions—an exact analytic counterpart of the drift–diffusion–jump structure of general stochastic generators.

Correspondence with Markov Processes: For a regular semi-Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(E; m)$, there exists a unique (up to m-equivalence) *Hunt process* $(X_t)_{t\geq 0}$ such that:

$$T_t f(x) = \mathbb{E}^x[f(X_t)], \qquad \forall f \in L^2(E; m), \tag{10.58}$$

where $(T_t)_{t\geq 0}$ is the semigroup associated with $\mathcal{E}$. This establishes the *analytic–probabilistic duality*: every semi-Dirichlet form corresponds to a (possibly non-symmetric) Markov process, and conversely, every such process admits an energy form of the Beurling–Deny type.

Relation to the Lévy–Khintchine Structure: The decomposition (10.56) generalizes the Lévy–Khintchine representation of infinitely divisible processes. Indeed:

- the local part $\mathcal{E}^{(c)}$ corresponds to the Gaussian (Brownian) component;
- the jump measure $J(dx, dy)$ corresponds to the Lévy measure $\nu(dy)$;
- the killing term $k(dx)$ corresponds to an exponential killing rate or absorption potential.

Hence, the Beurling–Deny representation unifies both local (differential) and nonlocal (integral) aspects of stochastic dynamics within a single variational framework.

Fundamental Duality Principle: The analytic form and its associated stochastic process are two realizations of the same underlying object:

$$\text{Energy form } \mathcal{E}(f,g) \quad \longleftrightarrow \quad \text{Infinitesimal generator } L, \quad \longleftrightarrow \quad \text{Hunt process } (X_t)_{t\geq 0}. \tag{10.59}$$

This triad expresses the deep equivalence between the variational, analytic, and probabilistic formulations of nonlocal stochastic dynamics, extending classical diffusion theory to the general setting of Lévy-type and jump processes.

10.8 Spectral and Regularity Properties

The spectral and regularity theory of nonlocal operators such as the fractional Laplacian and, more generally, Lévy-type generators, forms one of the central pillars of modern nonlocal analysis. Although these operators share deep analogies with their local elliptic counterparts, their spectral structure and the nature of regularity of their eigenfunctions differ in subtle but fundamental ways, owing to the nonlocal interactions encoded in their jump kernels.

Spectral Structure and Discreteness: Let $L = (-\Delta)^{\alpha/2} + V(x)$ on $L^2(\mathbb{R}^d)$, with $0 < \alpha < 2$, where $V : \mathbb{R}^d \to [0, \infty)$ is a measurable confining potential satisfying

$$V(x) \to +\infty \quad \text{as } |x| \to \infty. \tag{10.60}$$

Under these conditions, L defines a self-adjoint, nonnegative operator on $L^2(\mathbb{R}^d)$ with form domain

$$\mathcal{D}(\mathcal{E}_V) = \left\{ f \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|f(x)-f(y)|^2}{|x-y|^{d+\alpha}}\,dx\,dy + \int_{\mathbb{R}^d} V(x)|f(x)|^2\,dx < \infty \right\}. \tag{10.61}$$

The associated quadratic form

$$\mathcal{E}_V(f,f) = C_{d,\alpha}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|f(x)-f(y)|^2}{|x-y|^{d+\alpha}}\,dx\,dy + \int_{\mathbb{R}^d} V(x)|f(x)|^2\,dx \tag{10.62}$$

is closed, symmetric, and coercive. By standard compactness arguments (Rellich-type embedding theorems for fractional Sobolev spaces), one deduces that the resolvent $(L + \lambda I)^{-1}$ is compact on $L^2(\mathbb{R}^d)$ for all $\lambda > 0$. Consequently, L admits a purely discrete spectrum:

$$0 \leq \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots, \qquad \lambda_n \to +\infty, \tag{10.63}$$

and the corresponding eigenfunctions $\{\phi_n\}_{n\geq 1}$ form a complete orthonormal basis of $L^2(\mathbb{R}^d)$.

Regularity of Eigenfunctions: Unlike classical elliptic operators, eigenfunctions of nonlocal operators are not smooth in general, due to the singular and integral nature of the generator. However, they enjoy a form of fractional regularity.

Precisely, if $L = (-\Delta)^{\alpha/2} + V(x)$ with $V \in L^\infty_{\text{loc}}(\mathbb{R}^d)$, and $L\phi = \lambda\phi$, then by fractional elliptic regularity theory, one has

$$\phi \in C^{\alpha+\beta}_{\text{loc}}(\mathbb{R}^d) \quad \text{for some } \beta \in (0, 1-\alpha), \tag{10.64}$$

and moreover, if V is Hölder continuous, one obtains global Hölder continuity

$$\phi \in C^{\alpha+\gamma}(\mathbb{R}^d) \quad \text{for some } \gamma > 0. \tag{10.65}$$

This fractional regularity reflects the nonlocal smoothing properties of the operator: jumps connect distant regions of the space, propagating regularity in a weaker, integral sense rather than through differential smoothing.

Heat Semigroup and Analyticity: The semigroup $(T_t)_{t\geq 0}$ generated by $-(-\Delta)^{\alpha/2}$ is strongly continuous, symmetric, and contractive on $L^2(\mathbb{R}^d)$. Moreover, it extends analytically to a sector in the complex plane:

$$T_z f = \mathcal{F}^{-1}\Big[e^{-z|\xi|^\alpha}\widehat{f}(\xi)\Big], \qquad |\arg z| < \frac{\pi\alpha}{2}. \tag{10.66}$$

The analyticity follows from the fact that the generator $-(-\Delta)^{\alpha/2}$ is a sectorial operator on L^2, i.e.

$$\|(\lambda I + (-\Delta)^{\alpha/2})^{-1}\|_{L^2\to L^2} \leq \frac{C}{|\lambda|}, \qquad |\arg\lambda| < \pi - \theta_\alpha, \tag{10.67}$$

for some sector angle $\theta_\alpha < \pi/2$. Hence, (T_t) is analytic and contractive, satisfying

$$\|T_t f\|_{L^2} \leq \|f\|_{L^2}, \qquad \forall t \geq 0. \tag{10.68}$$

Heat Kernel and Scaling Behavior: The transition probability density (heat kernel) $p_t(x, y)$ of the symmetric α-stable process satisfies the Fourier representation

$$p_t(x, y) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)}e^{-t|\xi|^\alpha}\,d\xi. \tag{10.69}$$

From this, one derives the fundamental two-sided heat kernel estimate:

$$p_t(x, y) \asymp \frac{t}{\big(t^{1/\alpha} + |x-y|\big)^{d+\alpha}}, \tag{10.70}$$

uniformly for all $t > 0$ and $x, y \in \mathbb{R}^d$. The notation $f(t, x, y) \asymp g(t, x, y)$ indicates the existence of constants $C_1, C_2 > 0$ such that

$$C_1\, g(t, x, y) \leq f(t, x, y) \leq C_2\, g(t, x, y), \tag{10.71}$$

for all arguments.

This estimate illustrates the heavy-tailed nature of jump propagation: whereas the classical Gaussian heat kernel decays exponentially in $|x-y|^2/t$, the fractional kernel decays polynomially, allowing for long-range interactions. Consequently, the process generated by $-(-\Delta)^{\alpha/2}$ exhibits nonlocal diffusion and infinite propagation speed, reflecting the essential difference between diffusive and jump-driven stochastic dynamics.

Summary of Analytic and Spectral Properties: In summary, for the fractional Laplacian and its general Lévy-type analogues:

- the operator is nonlocal, self-adjoint, and nonnegative on $L^2(\mathbb{R}^d)$;
- the spectrum is discrete under confinement potentials $V(x) \to +\infty$;
- eigenfunctions are Hölder continuous but not smooth;
- the semigroup is analytic and contractive on L^2;
- the heat kernel obeys heavy-tailed decay consistent with jump behavior.

These properties together reveal the deep interplay between spectral theory, regularity, and probabilistic structure in nonlocal analysis.

10.9 Conclusion

The extension from diffusive to jump processes represents a profound generalization of the classical probabilistic–analytic correspondence. While Brownian motion leads to local differential operators and continuous trajectories, Lévy processes and their Dirichlet forms give rise to nonlocal operators and càdlàg paths. This unified framework provides the foundation for modern analysis of stochastic systems with discontinuities, including anomalous diffusion, stable processes, and kinetic models, and forms a natural bridge between partial differential equations, harmonic analysis, and probability theory on infinite-dimensional spaces.

References: Applebaum (2009) [31], Sato (1999) [33], Bass et al. (2010) [34].

Chapter 11

Nonlocal PDEs and Fractional Diffusions

"To understand diffusion beyond locality is to touch the geometry of space itself."
— I. M. Gel'fand

11.1 Motivation and Overview

Fractional and nonlocal partial differential equations arise naturally as the infinitesimal descriptions of stochastic processes with jumps. While local diffusions correspond to second-order differential operators, nonlocal diffusions are governed by integro-differential operators whose behavior reflects the long-range interactions inherent in Lévy-type stochastic dynamics. Such equations have found central applications in anomalous transport, turbulence, quantitative finance, and nonlocal continuum mechanics.

The prototype of nonlocal diffusion is the *fractional heat equation*

$$\partial_t u(t,x) + (-\Delta)^{\alpha/2} u(t,x) = 0, \qquad (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{11.1}$$

whose fundamental solution describes the transition density of an isotropic α-stable Lévy process. The fractional Laplacian here represents the generator of nonlocal diffusion, producing heavy-tailed propagation and infinite propagation speed in contrast to the Gaussian case.

11.2 Fractional Laplacian: Analytic Definitions

The fractional Laplacian $(-\Delta)^{\alpha/2}$ for $0 < \alpha < 2$ is a paradigmatic example of a nonlocal operator arising in analysis, probability, and geometry. It can be rigorously defined in several equivalent ways, each revealing different structural aspects of the operator — spectral, integral, and variational. In what follows, we present a mathematically rigorous derivation and justification of the three most fundamental formulations.

Fourier Representation: Let $f \in \mathcal{S}(\mathbb{R}^d)$, the Schwartz space of rapidly decreasing smooth functions. The Fourier transform and its inverse are defined, respectively, as

$$\widehat{f}(\xi) = \mathcal{F}[f](\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx, \qquad f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} \widehat{f}(\xi)\, d\xi. \tag{11.2}$$

The fractional Laplacian is then defined spectrally by

$$\mathcal{F}\Big[(-\Delta)^{\alpha/2} f\Big](\xi) = |\xi|^{\alpha}\, \widehat{f}(\xi), \qquad 0 < \alpha < 2. \tag{11.3}$$

This identity is a natural extension of the classical Laplacian since for $\alpha = 2$,

$$\mathcal{F}[(-\Delta) f](\xi) = |\xi|^2 \widehat{f}(\xi).$$

Thus $(-\Delta)^{\alpha/2}$ can be viewed as a fractional power of the positive self-adjoint operator $-\Delta$ in the sense of the spectral theorem. Formally, if $-\Delta = \int_0^\infty \lambda\, dE(\lambda)$ denotes its spectral resolution, then

$$(-\Delta)^{\alpha/2} = \int_0^\infty \lambda^{\alpha/2}\, dE(\lambda), \tag{11.4}$$

which provides a rigorous spectral definition valid in $L^2(\mathbb{R}^d)$.

Singular Integral Representation: For functions $f \in C_c^\infty(\mathbb{R}^d)$, one can equivalently define the fractional Laplacian as a singular integral operator:

$$(-\Delta)^{\alpha/2} f(x) = C_{d,\alpha}\, \text{P.V.} \int_{\mathbb{R}^d} \frac{f(x) - f(y)}{|x-y|^{d+\alpha}}\, dy, \qquad 0 < \alpha < 2, \tag{11.5}$$

where P.V. denotes the Cauchy principal value defined as

$$\text{P.V.} \int_{\mathbb{R}^d} g(y)\, dy := \lim_{\varepsilon \to 0} \int_{|y-x| > \varepsilon} g(y)\, dy, \tag{11.6}$$

and $C_{d,\alpha}$ is a dimension- and order-dependent normalization constant given explicitly by

$$C_{d,\alpha} = \frac{2^\alpha\, \Gamma\!\left(\frac{d+\alpha}{2}\right)}{\pi^{d/2}\, |\Gamma(-\alpha/2)|}. \tag{11.7}$$

This normalization ensures that the Fourier symbol of the operator in (11.5) equals $|\xi|^\alpha$, thereby making the integral and Fourier definitions equivalent. For each $f \in C_c^\infty(\mathbb{R}^d)$, the function

$$K_\alpha(x-y) = \frac{C_{d,\alpha}}{|x-y|^{d+\alpha}}$$

defines a singular, non-integrable kernel at the origin. However, the antisymmetrization in the numerator $f(x) - f(y)$ renders the integral finite, since by Taylor expansion,

$$f(x) - f(y) = \nabla f(x) \cdot (x-y) + O(|x-y|^2),$$

and hence the local singularity behaves like $|x-y|^{-d-\alpha+1}$, which is integrable near $y = x$ provided $\alpha < 2$. Thus the principal value integral exists pointwise for all smooth f.

This representation explicitly manifests the *nonlocality* of $(-\Delta)^{\alpha/2}$: the value at a point x depends on all values of f across $\mathbb{R}^d$, weighted by a kernel decaying like $|x-y|^{-d-\alpha}$. This property underlies the jump-diffusion character of the associated stochastic processes.

Dirichlet Form Representation: Let $H^{\alpha/2}(\mathbb{R}^d)$ denote the fractional Sobolev space, defined as

$$H^{\alpha/2}(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|f(x)-f(y)|^2}{|x-y|^{d+\alpha}}\,dx\,dy < \infty \right\}. \tag{11.8}$$

On this space, we define the bilinear form

$$\mathcal{E}(f,g) = \frac{C_{d,\alpha}}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(f(x)-f(y))(\overline{g(x)}-\overline{g(y)})}{|x-y|^{d+\alpha}}\,dx\,dy, \tag{11.9}$$

and the corresponding energy

$$\mathcal{E}(f,f) = \frac{C_{d,\alpha}}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|f(x)-f(y)|^2}{|x-y|^{d+\alpha}}\,dx\,dy. \tag{11.10}$$

The form $\mathcal{E}$ is:

- **Symmetric:** $\mathcal{E}(f,g) = \overline{\mathcal{E}(g,f)}$.
- **Closed:** The domain $H^{\alpha/2}(\mathbb{R}^d)$ is complete under the norm

$$\|f\|^2_{H^{\alpha/2}} = \|f\|^2_{L^2} + \mathcal{E}(f,f).$$

- **Markovian:** For any normal contraction $\eta : \mathbb{R} \to \mathbb{R}$ with $|\eta(a)-\eta(b)| \le |a-b|$, we have

$$\mathcal{E}(\eta \circ f, \eta \circ f) \le \mathcal{E}(f,f),$$

  ensuring that the corresponding semigroup preserves positivity and contractivity.

By the theory of symmetric Dirichlet forms (Fukushima–Oshima–Takeda), the form $\mathcal{E}$ uniquely determines a self-adjoint, nonnegative operator A on $L^2(\mathbb{R}^d)$ satisfying

$$\mathcal{E}(f,g) = \langle Af, g\rangle_{L^2}, \qquad f \in \mathcal{D}(A),\ g \in H^{\alpha/2}(\mathbb{R}^d), \tag{11.11}$$

and this operator coincides with the fractional Laplacian $A = (-\Delta)^{\alpha/2}$. Hence, the three formulations — Fourier, singular integral, and Dirichlet form — are rigorously equivalent on their common domains.

Summary:

Definition Type	**Formula**	**Key Feature**
Fourier	$\mathcal{F}[(-\Delta)^{\alpha/2}f](\xi) = \|\xi\|^\alpha \widehat{f}(\xi)$	Spectral structure
Singular Integral	$(-\Delta)^{\alpha/2}f(x) = C_{d,\alpha}\,\text{P.V.}\int \frac{f(x)-f(y)}{\|x-y\|^{d+\alpha}}dy$	Nonlocality
Dirichlet Form	$\mathcal{E}(f,f) = \frac{C_{d,\alpha}}{2}\iint \frac{\|f(x)-f(y)\|^2}{\|x-y\|^{d+\alpha}}dx\,dy$	Variational energy

Each perspective exposes a different analytic facet of the fractional Laplacian — its spectral nature, its nonlocal integral structure, and its variational interpretation as the generator of a symmetric α-stable Lévy process.

11.3 Fractional Heat Equation and Its Kernel

The *fractional heat equation* is the evolution equation governed by the nonlocal operator $(-\Delta)^{\alpha/2}$ for $0 < \alpha < 2$. Formally, it reads

$$\partial_t u(t,x) + (-\Delta)^{\alpha/2} u(t,x) = 0, \qquad t > 0,\ x \in \mathbb{R}^d, \tag{11.12}$$

with the initial condition $u(0,x) = f(x) \in L^2(\mathbb{R}^d)$. The solution defines a strongly continuous semigroup $(P_t)_{t\geq 0}$ on $L^2(\mathbb{R}^d)$, given by

$$P_t f = e^{-t(-\Delta)^{\alpha/2}} f. \tag{11.13}$$

Fourier Representation of the Semigroup: By the spectral characterization of the fractional Laplacian, we have

$$\mathcal{F}[(-\Delta)^{\alpha/2} f](\xi) = |\xi|^\alpha \widehat{f}(\xi),$$

which immediately implies that the semigroup acts diagonally in Fourier space:

$$\mathcal{F}[P_t f](\xi) = e^{-t|\xi|^\alpha} \widehat{f}(\xi), \qquad t \geq 0,\ \xi \in \mathbb{R}^d. \tag{11.14}$$

Taking the inverse Fourier transform yields the convolution representation

$$P_t f(x) = (p_t * f)(x), \qquad \text{where } p_t(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} e^{-t|\xi|^\alpha}\, d\xi. \tag{11.15}$$

The function $p_t(x)$ serves as the *fractional heat kernel*, or equivalently, the transition density of the isotropic symmetric α-stable Lévy process.

Self-Similar Scaling Property: A fundamental structural property of $p_t(x)$ is its *self-similarity* (or scaling invariance). By performing a change of variables $\xi' = t^{1/\alpha}\xi$ in (11.15), one obtains

$$\begin{aligned} p_t(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} e^{-t|\xi|^\alpha}\, d\xi = \frac{t^{-d/\alpha}}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i(t^{-1/\alpha}x)\cdot\xi'} e^{-|\xi'|^\alpha}\, d\xi' \\ &= t^{-d/\alpha} p_1(t^{-1/\alpha}x). \end{aligned} \tag{11.16}$$

This homogeneity relation reflects the scaling behavior of the generator:

$$(-\Delta)^{\alpha/2}(f(\lambda\cdot))(x) = \lambda^\alpha((-\Delta)^{\alpha/2} f)(\lambda x),$$

showing that the process is self-similar with index $1/\alpha$.

Probabilistic Interpretation: The fractional heat semigroup admits the probabilistic representation

$$(P_t f)(x) = \mathbb{E}^x[f(X_t)], \tag{11.17}$$

where $(X_t)_{t\geq 0}$ is an isotropic symmetric α-stable Lévy process with characteristic function

$$\mathbb{E}^0[e^{i\xi\cdot X_t}] = e^{-t|\xi|^\alpha}.$$

Thus, $p_t(x)$ is the probability density function of X_t, i.e.,

$$p_t(x)\,dx = \mathbb{P}(X_t \in dx).$$

In particular, the process exhibits jumps with heavy-tailed increments, in contrast to the Gaussian increments of Brownian motion.

Heat Kernel Estimates: The kernel $p_t(x)$ is positive, radially symmetric, and satisfies the following precise asymptotic estimates: there exist constants $c_1, c_2 > 0$, depending only on d and α, such that for all $t > 0$ and $x \in \mathbb{R}^d$,

$$c_1 \frac{t}{(t^{1/\alpha} + |x|)^{d+\alpha}} \le p_t(x) \le c_2 \frac{t}{(t^{1/\alpha} + |x|)^{d+\alpha}}. \tag{11.18}$$

The proof of these inequalities rests on detailed asymptotic analysis of the Fourier integral and Tauberian theorems relating the decay of $e^{-t|\xi|^\alpha}$ to the behavior of $p_t(x)$ at infinity.

Asymptotic Behavior: From (11.18), we infer:

- For small displacements ($|x| \ll t^{1/\alpha}$):

  $$p_t(x) \approx c\, t^{-d/\alpha},$$

  which matches the density scaling of a stable process.

- For large displacements ($|x| \gg t^{1/\alpha}$):

  $$p_t(x) \approx c' \frac{t}{|x|^{d+\alpha}},$$

  showing a power-law, heavy-tailed decay unlike the Gaussian exponential decay of the classical heat kernel.

Comparison with Gaussian Heat Kernel: For the classical Laplacian ($\alpha = 2$), the heat kernel is

$$p_t^{(2)}(x) = (4\pi t)^{-d/2} e^{-\frac{|x|^2}{4t}},$$

which exhibits exponential decay in $|x|$. In contrast, for $0 < \alpha < 2$, the fractional heat kernel $p_t(x)$ decays polynomially, encoding the presence of long-range jumps and Lévy flight behavior. This difference marks the transition from local diffusion to nonlocal anomalous diffusion.

Summary: The fractional heat semigroup $(P_t)_{t\ge 0}$ generated by $-(-\Delta)^{\alpha/2}$ satisfies:

$$\mathcal{F}[p_t](\xi) = e^{-t|\xi|^\alpha}, \tag{11.19}$$

$$p_t(x) = t^{-d/\alpha} p_1(t^{-1/\alpha} x), \tag{11.20}$$

$$c_1 \frac{t}{(t^{1/\alpha} + |x|)^{d+\alpha}} \le p_t(x) \le c_2 \frac{t}{(t^{1/\alpha} + |x|)^{d+\alpha}}. \tag{11.21}$$

These results rigorously characterize the analytic and probabilistic structure of the fractional heat flow, revealing its self-similar scaling, nonlocal propagation, and heavy-tailed behavior intrinsic to stable Lévy processes.

11.4 Cauchy Problems and Well-Posedness

We now provide a mathematically rigorous treatment of the Cauchy problem associated with the fractional Laplacian. Let $0 < \alpha < 2$ and consider the equation

$$\begin{cases} \partial_t u(t,x) + (-\Delta)^{\alpha/2} u(t,x) = 0, & t > 0,\ x \in \mathbb{R}^d, \\ u(0,x) = u_0(x), & x \in \mathbb{R}^d. \end{cases} \tag{11.22}$$

We seek to establish existence, uniqueness, and continuous dependence of solutions in an appropriate functional setting.

Functional Analytic Framework: Let $L^2(\mathbb{R}^d)$ denote the Hilbert space with the standard inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx,$$

and norm $\|f\|_{L^2} = \langle f, f \rangle^{1/2}$. Define the operator $A := (-\Delta)^{\alpha/2}$ on the domain

$$\mathcal{D}(A) = \left\{ f \in L^2(\mathbb{R}^d) : |\xi|^\alpha \widehat{f}(\xi) \in L^2(\mathbb{R}^d) \right\},$$

that is, $\mathcal{D}(A) = H^\alpha(\mathbb{R}^d)$, the fractional Sobolev space.

For $f \in \mathcal{D}(A)$, we have

$$\|Af\|_{L^2}^2 = \int_{\mathbb{R}^d} |\xi|^{2\alpha} |\widehat{f}(\xi)|^2\, d\xi < \infty.$$

The operator A is self-adjoint, positive definite, and densely defined. Therefore, by the spectral theorem, it generates a strongly continuous contraction semigroup $\{P_t\}_{t\geq 0}$ on $L^2(\mathbb{R}^d)$ given by

$$P_t = e^{-tA}.$$

Mild and Weak Formulations: The mild solution to (11.22) is defined by

$$u(t) = P_t u_0 = e^{-tA} u_0. \tag{11.23}$$

In the Fourier domain, this is explicitly given by

$$\widehat{u}(t,\xi) = e^{-t|\xi|^\alpha} \widehat{u_0}(\xi),$$

which, by Plancherel's theorem, ensures that $u(t) \in L^2(\mathbb{R}^d)$ for all $t \geq 0$. Moreover, the strong continuity of (P_t) implies

$$\lim_{t \to 0^+} \|P_t u_0 - u_0\|_{L^2} = 0,$$

so that $u \in C([0,\infty); L^2(\mathbb{R}^d))$. A function $u : [0,\infty) \to L^2(\mathbb{R}^d)$ is called a weak solution of (11.22) if, for all $\varphi \in \mathcal{D}(A)$,

$$\frac{d}{dt}\langle u(t), \varphi \rangle + \langle A^{1/2} u(t), A^{1/2} \varphi \rangle = 0, \qquad \text{for almost every } t > 0,$$

and $u(0) = u_0$. By the semigroup representation (11.23), the mild and weak solutions coincide.

Contractivity and Energy Dissipation: The semigroup $(P_t)_{t\geq 0}$ satisfies the following fundamental properties:

- **(Strong continuity)**: $\lim_{t\to 0^+} \|P_t f - f\|_{L^2} = 0$ for all $f \in L^2(\mathbb{R}^d)$.
- **(Self-adjointness)**: $\langle P_t f, g\rangle = \langle f, P_t g\rangle$ for all $f, g \in L^2(\mathbb{R}^d)$.
- **(Contractivity)**: $\|P_t f\|_{L^2} \leq \|f\|_{L^2}$ for all $t \geq 0$.

Differentiating the L^2-norm along the solution $u(t)$ yields the energy dissipation identity

$$\frac{d}{dt}\|u(t)\|_{L^2}^2 = -2\langle A^{1/2}u(t), A^{1/2}u(t)\rangle = -2\|A^{1/2}u(t)\|_{L^2}^2 \leq 0. \tag{11.24}$$

Thus, the L^2-norm of the solution is non-increasing in time:

$$\|u(t)\|_{L^2} \leq \|u_0\|_{L^2}, \qquad \forall t \geq 0.$$

Uniqueness and Well-Posedness: To show uniqueness, suppose u_1 and u_2 are two mild solutions with the same initial data. Then $w = u_1 - u_2$ satisfies

$$\partial_t w + Aw = 0, \qquad w(0) = 0.$$

Taking the inner product with $w(t)$ and integrating over $\mathbb{R}^d$, we find

$$\frac{1}{2}\frac{d}{dt}\|w(t)\|_{L^2}^2 + \|A^{1/2}w(t)\|_{L^2}^2 = 0.$$

Since $\|w(0)\|_{L^2} = 0$, Grönwall's inequality gives $\|w(t)\|_{L^2} = 0$ for all t, hence $u_1 = u_2$. Therefore, the solution is unique in $C([0,\infty); L^2(\mathbb{R}^d))$.

Conclusion: The Cauchy problem (11.22) admits a unique mild (equivalently, weak) solution

$$u(t,x) = (P_t u_0)(x) = (p_t * u_0)(x),$$

with p_t the fractional heat kernel. The mapping $u_0 \mapsto u(t)$ defines a strongly continuous, symmetric, and contractive semigroup on $L^2(\mathbb{R}^d)$. Hence, the fractional heat equation is *well-posed* in the L^2 sense, with continuous dependence on the initial data and conservation (or decay) of the L^2 energy.

11.5 Fractional Schrödinger Operators

We now construct and analyze the fractional Schrödinger operator, which generalizes the classical Schrödinger operator to incorporate nonlocal effects governed by the fractional Laplacian.

Definition and Functional Setting: Let $0 < \alpha < 2$. The fractional Schrödinger operator is defined formally as

$$H = (-\Delta)^{\alpha/2} + V(x), \tag{11.25}$$

where $V : \mathbb{R}^d \to \mathbb{R}$ is a measurable potential function. The operator acts on the Hilbert space $L^2(\mathbb{R}^d)$. The natural domain of H is determined by the intersection of the domains of its two components:

$$\mathcal{D}(H) = \left\{ f \in H^\alpha(\mathbb{R}^d) : V(x)f(x) \in L^2(\mathbb{R}^d) \right\}.$$

Thus, for $f \in \mathcal{D}(H)$, we define

$$Hf = (-\Delta)^{\alpha/2} f + V(x)f(x),$$

where $(-\Delta)^{\alpha/2}$ is the fractional Laplacian understood in the Fourier sense,

$$\widehat{(-\Delta)^{\alpha/2}} f(\xi) = |\xi|^\alpha \widehat{f}(\xi).$$

Quadratic form representation: Define the quadratic form associated with H by

$$\mathcal{E}_V(f,g) = \frac{C_{d,\alpha}}{2} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(f(x)-f(y))\overline{(g(x)-g(y))}}{|x-y|^{d+\alpha}}\, dx\, dy + \int_{\mathbb{R}^d} V(x) f(x)\overline{g(x)}\, dx, \tag{11.26}$$

for $f, g \in C_c^\infty(\mathbb{R}^d)$, where $C_{d,\alpha}$ is the normalization constant of the fractional Laplacian. The form $\mathcal{E}_V$ is closable on $L^2(\mathbb{R}^d)$ whenever V belongs to the Kato class $\mathcal{K}_{d,\alpha}$, which ensures suitable integrability and smallness properties of the potential.

Self-Adjointness and Semiboundedness: The quadratic form (11.26) is densely defined, symmetric, and semibounded from below if

$$\inf_{x \in \mathbb{R}^d} V(x) > -\infty.$$

By the Kato–Friedrichs representation theorem, there exists a unique self-adjoint operator H on $L^2(\mathbb{R}^d)$ such that

$$\mathcal{E}_V(f,g) = \langle H^{1/2} f, H^{1/2} g\rangle, \qquad f, g \in \mathcal{D}(\mathcal{E}_V),$$

and this operator coincides with (11.25) in the distributional sense. Moreover, the spectrum $\sigma(H) \subseteq [\inf V, \infty)$ is real and bounded from below.

Feynman–Kac Representation for the Semigroup: Let $(X_t)_{t\geq 0}$ denote a symmetric α-stable Lévy process on $\mathbb{R}^d$, characterized by the Fourier transform

$$\mathbb{E}^0\left[e^{i\langle \xi, X_t\rangle}\right] = e^{-t|\xi|^\alpha}.$$

The generator of (X_t) is $-(-\Delta)^{\alpha/2}$. For V in the Kato class, the semigroup $\{e^{-tH}\}_{t\geq 0}$ associated with H admits a probabilistic representation known as the *nonlocal Feynman–Kac formula*:

$$(e^{-tH} f)(x) = \mathbb{E}^x\left[\exp\left(-\int_0^t V(X_s)\, ds\right) f(X_t)\right], \tag{11.27}$$

where $\mathbb{E}^x$ denotes the expectation for the process started at $X_0 = x$.

For the justification of (11.27), the identity (11.27) follows from the following steps:

1. The semigroup $(P_t)_{t\geq 0}$ generated by $-(-\Delta)^{\alpha/2}$ acts as

$$(P_t f)(x) = \mathbb{E}^x[f(X_t)] = (p_t * f)(x),$$

with Fourier symbol $e^{-t|\xi|^\alpha}$.

2. Define the multiplicative functional

$$M_t = \exp\left(-\int_0^t V(X_s)\,ds\right),$$

which is positive and satisfies $M_t \leq e^{t\|V_-\|_\infty}$ if V_- denotes the negative part of V.

3. Then, the family $\{T_t\}_{t\geq 0}$ defined by

$$(T_t f)(x) = \mathbb{E}^x[M_t f(X_t)]$$

forms a strongly continuous, symmetric, positivity-preserving semigroup on $L^2(\mathbb{R}^d)$.

4. The generator of (T_t) is the self-adjoint operator H defined in (11.25).

Analytic and Probabilistic Equivalence: The representation (11.27) establishes an exact analytic–probabilistic equivalence:

$$e^{-tH} = \mathbb{E}\left[\exp\left(-\int_0^t V(X_s)\,ds\right)\cdot f(X_t)\right],$$

linking the semigroup generated by the nonlocal operator H with expectations over the paths of the Lévy process (X_t).

This duality provides a powerful bridge between spectral theory and stochastic processes:

- The analytic properties of H (self-adjointness, spectral bounds, domain structure) correspond to path properties of (X_t) and integrability of the exponential functional of V.
- The lower spectral bound of H satisfies

$$\inf \sigma(H) \geq \inf_{x\in\mathbb{R}^d} V(x),$$

reflecting the fact that the potential term shifts the energy spectrum upward by its infimum.

Summary: Under the assumption that V belongs to the Kato class $\mathcal{K}_{d,\alpha}$ and is bounded from below, the fractional Schrödinger operator

$$H = (-\Delta)^{\alpha/2} + V(x)$$

is self-adjoint, semibounded, and generates a strongly continuous symmetric semigroup $(e^{-tH})_{t\geq 0}$ on $L^2(\mathbb{R}^d)$ satisfying the nonlocal Feynman–Kac representation (11.27). The spectrum of H is purely real and satisfies

$$\sigma(H) \subseteq [\inf V, \infty),$$

demonstrating that the analytic and probabilistic formulations of the nonlocal quantum evolution are fully consistent.

11.6 Boundary Value Problems and Extension Technique

A remarkable and deep connection between nonlocal and local partial differential equations was discovered by Caffarelli and Silvestre, showing that the fractional Laplacian can be realized as a Dirichlet-to-Neumann map for a degenerate elliptic problem posed in one higher dimension. This correspondence allows one to translate nonlocal problems into local PDEs with weighted ellipticity, thereby enabling the use of classical analytic tools in nonlocal analysis.

The Extension Problem: Let $\alpha \in (0, 2)$ and define the extended domain as the upper half-space

$$\mathbb{R}^{d+1}_+ = \{(x, y) \in \mathbb{R}^d \times (0, \infty)\}.$$

Given boundary data $f : \mathbb{R}^d \to \mathbb{R}$, consider the weighted elliptic boundary value problem

$$\begin{cases} \nabla \cdot (y^{1-\alpha} \nabla U(x, y)) = 0, & (x, y) \in \mathbb{R}^{d+1}_+, \\ U(x, 0) = f(x), & x \in \mathbb{R}^d. \end{cases} \tag{11.28}$$

The function U is called the *α-harmonic extension* of f. The equation in (11.28) is a *degenerate elliptic* PDE, since the coefficient $y^{1-\alpha}$ vanishes or blows up near $y = 0$ depending on whether $\alpha < 1$ or $\alpha > 1$.

The operator

$$L_\alpha U := \nabla \cdot (y^{1-\alpha} \nabla U)$$

is self-adjoint on the weighted space $L^2(\mathbb{R}^{d+1}_+, y^{1-\alpha}\, dx\, dy)$, and the corresponding weak formulation reads:

$$\int_{\mathbb{R}^{d+1}_+} y^{1-\alpha} \nabla U \cdot \nabla \varphi \, dx\, dy = 0, \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^{d+1}_+).$$

Poisson Kernel Representation: The solution U to (11.28) can be represented explicitly by convolution with a suitable Poisson kernel:

$$U(x, y) = (P_y^{(\alpha)} * f)(x), \tag{11.29}$$

where the α-Poisson kernel is given by

$$P_y^{(\alpha)}(x) = c_{d,\alpha} \frac{y^\alpha}{(|x|^2 + y^2)^{\frac{d+\alpha}{2}}}, \qquad c_{d,\alpha} = \frac{\Gamma\left(\frac{d+\alpha}{2}\right)}{\pi^{d/2}\Gamma\left(\frac{\alpha}{2}\right)}. \tag{11.30}$$

This kernel satisfies $\int_{\mathbb{R}^d} P_y^{(\alpha)}(x)\, dx = 1$ for all $y > 0$, and one easily checks that $L_\alpha U = 0$ in $\mathbb{R}^{d+1}_+$ and $\lim_{y\to 0^+} U(x, y) = f(x)$.

Dirichlet-to-Neumann Map and the Fractional Laplacian: The key identity discovered by Caffarelli and Silvestre is the correspondence

$$(-\Delta)^{\alpha/2} f(x) = -\,C_\alpha \lim_{y \to 0^+} y^{1-\alpha}\, \partial_y U(x,y), \tag{11.31}$$

where U solves (11.28) and

$$C_\alpha = \frac{2^{\alpha-1}\Gamma(\alpha/2)}{\Gamma(1-\alpha/2)}.$$

This identity expresses the fractional Laplacian as a boundary flux associated with the weighted harmonic extension of f. The right-hand side is well-defined for $f \in H^{\alpha/2}(\mathbb{R}^d)$, and the limit exists in the sense of distributions.

Variational Formulation and Energy Identity: The extension problem admits a natural variational structure. Define the weighted energy functional

$$\mathcal{E}_\alpha(U) = \frac{1}{2} \int_{\mathbb{R}^{d+1}_+} y^{1-\alpha} |\nabla U(x,y)|^2 \, dx \, dy. \tag{11.32}$$

Among all functions U with prescribed boundary trace $U(\cdot, 0) = f$, the minimizer of $\mathcal{E}_\alpha$ satisfies the Euler–Lagrange equation $L_\alpha U = 0$, i.e., it solves the extension problem (11.28).

Moreover, there is an exact equivalence of energies:

$$\mathcal{E}_\alpha(U) = C_{d,\alpha} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|f(x) - f(y)|^2}{|x-y|^{d+\alpha}} \, dx \, dy = 2C_{d,\alpha}\, \mathcal{E}(f,f), \tag{11.33}$$

where $\mathcal{E}(f,f)$ is the Dirichlet form of the fractional Laplacian. This shows that minimizing the weighted local energy $\mathcal{E}_\alpha(U)$ in $\mathbb{R}^{d+1}_+$ is equivalent to minimizing the nonlocal energy on $\mathbb{R}^d$.

Analytic Consequences: The representation (11.31) has several fundamental consequences:

- **Boundary regularity:** Solutions to fractional elliptic equations $(-\Delta)^{\alpha/2} f = g$ inherit regularity from the corresponding degenerate elliptic equation $L_\alpha U = 0$.
- **Spectral interpretation:** On bounded domains $\Omega \subset \mathbb{R}^d$, the spectral fractional Laplacian defined via eigenfunction expansions coincides with the Dirichlet-to-Neumann map of the extension problem with vanishing boundary condition $U|_{\partial\Omega \times (0,\infty)} = 0$.
- **Variational analysis:** The correspondence (11.33) permits the application of tools from weighted Sobolev spaces and calculus of variations to fractional problems.
- **Harnack and Liouville principles:** The weighted harmonic nature of U yields nonlocal analogues of classical results such as Harnack inequalities and Liouville theorems for α-harmonic functions.

Summary: The Caffarelli–Silvestre extension provides a rigorous geometric and analytic realization of the fractional Laplacian as a boundary flux:

$$(-\Delta)^{\alpha/2} f(x) = -C_\alpha \lim_{y \to 0^+} y^{1-\alpha} \partial_y U(x,y),$$

where U solves the degenerate elliptic problem $\nabla \cdot (y^{1-\alpha}\nabla U) = 0$. This construction converts nonlocal problems on $\mathbb{R}^d$ into local PDEs on $\mathbb{R}^{d+1}_+$, thereby bridging fractional analysis, variational methods, and geometric PDE theory in a unified framework.

11.7 Nonlocal Variational Problems and Energy Minimization

In the mathematical theory of nonlocal partial differential equations, a central principle is that weak or distributional solutions can often be interpreted as critical points or minimizers of appropriately defined nonlocal energy functionals. This variational viewpoint extends the classical Dirichlet principle to the nonlocal framework governed by fractional Laplacians.

Consider the fractional Poisson equation

$$(-\Delta)^{\alpha/2}u = f \quad \text{in } \mathbb{R}^d, \tag{11.34}$$

where $0 < \alpha < 2$ and $f \in H^{-\alpha/2}(\mathbb{R}^d)$ is a given distribution. The operator $(-\Delta)^{\alpha/2}$ is nonlocal and self-adjoint on $L^2(\mathbb{R}^d)$, with domain $H^{\alpha/2}(\mathbb{R}^d)$. Its action may be represented weakly as

$$\langle (-\Delta)^{\alpha/2}u, v\rangle = C_{d,\alpha}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{(u(x)-u(y))(v(x)-v(y))}{|x-y|^{d+\alpha}}\,dx\,dy, \tag{11.35}$$

for all $u, v \in H^{\alpha/2}(\mathbb{R}^d)$, where $C_{d,\alpha}$ is a normalization constant depending only on d and α.

The weak formulation of the fractional Poisson equation seeks $u \in H^{\alpha/2}(\mathbb{R}^d)$ such that

$$\mathcal{E}(u, v) = \langle f, v\rangle, \qquad \forall v \in H^{\alpha/2}(\mathbb{R}^d), \tag{11.36}$$

where the bilinear form $\mathcal{E}(\cdot,\cdot)$ is defined by

$$\mathcal{E}(u, v) := \frac{C_{d,\alpha}}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{(u(x)-u(y))(v(x)-v(y))}{|x-y|^{d+\alpha}}\,dx\,dy. \tag{11.37}$$

This bilinear form is symmetric, coercive, and continuous on $H^{\alpha/2}(\mathbb{R}^d)$, that is,

$$\mathcal{E}(u, u) \geq c_{\alpha,d}\|u\|^2_{H^{\alpha/2}(\mathbb{R}^d)}, \qquad |\mathcal{E}(u, v)| \leq C_{\alpha,d}\|u\|_{H^{\alpha/2}}\|v\|_{H^{\alpha/2}},$$

for positive constants $c_{\alpha,d}, C_{\alpha,d}$. By the Lax–Milgram theorem, there exists a unique weak solution u satisfying this equation.

Equivalently, u minimizes the nonlocal energy functional

$$\mathcal{J}(u) = \frac{C_{d,\alpha}}{4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{|u(x)-u(y)|^2}{|x-y|^{d+\alpha}}\,dx\,dy - \int_{\mathbb{R}^d} f(x)u(x)\,dx. \tag{11.38}$$

To see this, one observes that for any perturbation $u+\varepsilon v$ with $v \in H^{\alpha/2}(\mathbb{R}^d)$, the first variation of $\mathcal{J}$ satisfies

$$\begin{aligned}\frac{d}{d\varepsilon}\mathcal{J}(u+\varepsilon v)\bigg|_{\varepsilon=0} &= \frac{C_{d,\alpha}}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{(u(x)-u(y))(v(x)-v(y))}{|x-y|^{d+\alpha}}\,dx\,dy - \int_{\mathbb{R}^d} f(x)v(x)\,dx \\ &= \mathcal{E}(u, v) - \langle f, v\rangle. \end{aligned} \tag{11.39}$$

Setting this to zero for all admissible v yields precisely the weak form of the fractional Poisson problem. Thus, the variational minimizer of $\mathcal{J}$ is the weak solution to $(-\Delta)^{\alpha/2}u = f$.

The energy $\mathcal{E}(u,u)$ admits a Fourier-space representation:

$$\mathcal{E}(u,u) = c_{d,\alpha} \int_{\mathbb{R}^d} |\xi|^\alpha |\widehat{u}(\xi)|^2 \, d\xi, \tag{11.40}$$

establishing the equivalence between the Gagliardo seminorm and the spectral norm of $H^{\alpha/2}$. This characterization emphasizes that the fractional Laplacian penalizes high-frequency oscillations according to a power-law weight, capturing nonlocal smoothness properties intrinsic to fractional-order Sobolev spaces.

Therefore, the fractional Poisson equation is well-posed in $H^{\alpha/2}(\mathbb{R}^d)$: there exists a unique weak solution depending continuously on f, and the solution minimizes the nonlocal Dirichlet energy subject to the forcing term. This framework generalizes the classical elliptic variational principle to encompass long-range interactions governed by power-law kernels.

11.8 Spectral Properties and Regularity

The spectral structure of the fractional Laplacian on bounded domains provides a rigorous nonlocal analogue of the classical elliptic spectral theory. Let $\Omega \subset \mathbb{R}^d$ be a bounded open domain with sufficiently smooth boundary, and let $0 < \alpha < 2$. We define the Dirichlet fractional Laplacian $(-\Delta)^{\alpha/2}$ as the self-adjoint operator associated with the quadratic form

$$\mathcal{E}(u,v) = \frac{C_{d,\alpha}}{2} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(u(x)-u(y))(v(x)-v(y))}{|x-y|^{d+\alpha}} \, dx \, dy, \tag{11.41}$$

with domain

$$\mathcal{D}(\mathcal{E}) = \{u \in H^{\alpha/2}(\mathbb{R}^d) : u = 0 \text{ a.e. in } \mathbb{R}^d \setminus \Omega\}.$$

This definition enforces a nonlocal Dirichlet condition by penalizing jumps across the boundary. The operator $(-\Delta)^{\alpha/2}$ acts as the generator of a symmetric Markov semigroup $(P_t)_{t\geq 0}$ on $L^2(\Omega)$, defined by $P_t = e^{-t(-\Delta)^{\alpha/2}}$.

Spectral decomposition: Since $\mathcal{E}$ is a symmetric, densely defined, closed, and coercive bilinear form on $L^2(\Omega)$, the associated operator $(-\Delta)^{\alpha/2}$ is self-adjoint, positive, and possesses a compact resolvent. Consequently, by the spectral theorem for compact self-adjoint operators, there exists an orthonormal basis of eigenfunctions

$$\{\phi_k\}_{k=1}^\infty \subset L^2(\Omega)$$

and an increasing sequence of eigenvalues

$$0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots, \quad \lambda_k \to +\infty,$$

such that

$$(-\Delta)^{\alpha/2}\phi_k = \lambda_k \phi_k, \qquad \text{with } \phi_k|_{\mathbb{R}^d\setminus\Omega} = 0. \tag{11.42}$$

The eigenfunctions form a complete orthonormal basis of $L^2(\Omega)$, that is,

$$\langle \phi_j, \phi_k \rangle_{L^2(\Omega)} = \delta_{jk}, \qquad f = \sum_{k=1}^{\infty} \langle f, \phi_k \rangle \phi_k, \quad \forall f \in L^2(\Omega).$$

Semigroup representation: The semigroup (P_t) generated by $-(-\Delta)^{\alpha/2}$ admits the spectral expansion

$$(P_t f)(x) = \sum_{k=1}^{\infty} e^{-\lambda_k t} \langle f, \phi_k \rangle \phi_k(x), \tag{11.43}$$

which converges in $L^2(\Omega)$ for each $t > 0$. This representation provides an explicit solution formula for the fractional heat equation

$$\partial_t u + (-\Delta)^{\alpha/2} u = 0, \quad u(0, \cdot) = f,$$

given by $u(t) = P_t f$.

Spectral asymptotics: The eigenvalues satisfy a Weyl-type asymptotic law analogous to that for the classical Laplacian:

$$\lambda_k \sim C_{d,\alpha} \left(\frac{k}{|\Omega|} \right)^{\alpha/d}, \qquad \text{as } k \to \infty, \tag{11.44}$$

where $C_{d,\alpha}$ depends only on the dimension and the fractional order. This relation quantifies the scaling of high-frequency eigenmodes and encodes the nonlocal dispersion inherent in the operator.

Regularity of eigenfunctions: Unlike the classical Laplacian, the fractional Laplacian lacks elliptic locality; consequently, the eigenfunctions ϕ_k are generally not smooth up to the boundary. For bounded Ω with Lipschitz boundary, one has

$$\phi_k \in C^{\beta}(\overline{\Omega}), \qquad \beta < \alpha,$$

and this Hölder regularity is optimal in general. Interior regularity is stronger: for $\Omega' \subset\subset \Omega$, $\phi_k \in C^{\infty}(\Omega')$, reflecting analytic smoothing in the interior but nonlocal singularity propagation near $\partial\Omega$. These results stem from the nonlocal elliptic estimates of Silvestre and Ros-Oton–Serra, which show that $(-\Delta)^{\alpha/2}$ gains only fractional derivatives in the boundary setting.

Functional analytic interpretation: The fractional Laplacian defines an unbounded self-adjoint operator on $L^2(\Omega)$ with domain

$$\mathcal{D}((-\Delta)^{\alpha/2}) = \left\{ u \in H_0^{\alpha/2}(\Omega) : (-\Delta)^{\alpha/2} u \in L^2(\Omega) \right\}.$$

The semigroup (P_t) is analytic and contractive on $L^2(\Omega)$, and satisfies

$$\|P_t f\|_{L^2} \leq \|f\|_{L^2}, \quad \frac{d}{dt} \|P_t f\|_{L^2}^2 = -2\mathcal{E}(P_t f, P_t f) \leq 0.$$

Thus, the fractional heat flow preserves L^2 norm and dissipates the nonlocal Dirichlet energy, reflecting both spectral decay and probabilistic contractivity of the corresponding α-stable process.

In summary, the spectral theory of the fractional Laplacian retains the orthogonal completeness and semigroup representation of classical elliptic operators, but with markedly different regularity characteristics: smoothness is lost at the boundary, replaced by fractional Hölder continuity that reflects the intrinsic nonlocality of the underlying operator.

11.9 Conclusion

Fractional and nonlocal partial differential equations extend the classical theory of diffusion to encompass long-range interactions and anomalous transport. Their deep connection with Lévy processes, Dirichlet forms, and spectral theory reveals a unified structure binding probability, functional analysis, and geometry. Through the Feynman–Kac correspondence, every nonlocal evolution acquires both analytic and stochastic meaning — establishing a bridge between PDEs, random jumps, and the geometry of nonlocality.

References: Di Nezza et al. 2012 [36], Servadei and Valdinoci 2015 [37].

Chapter 12

Infinite-Dimensional Diffusions and Stochastic Partial Differential Equations

"Probability in infinite dimensions is not mere chance—it is geometry itself."
— S. R. S. Varadhan

12.1 Motivation and Overview

Infinite-dimensional diffusions arise naturally when the state of a stochastic system is described by a random field rather than a finite-dimensional vector. Such systems include stochastic evolution of temperature fields, fluid velocities, concentration profiles, or quantum states, each governed by partial differential equations perturbed by random noise. The rigorous analytic framework for their study combines infinite-dimensional analysis, functional analysis on Hilbert and Banach spaces, and stochastic calculus in infinite dimensions.

A stochastic partial differential equation (SPDE) in abstract form is typically written as

$$dX_t = (AX_t + F(X_t))\,dt + G(X_t)\,dW_t, \qquad X_0 = x \in \mathcal{H}, \tag{12.1}$$

where:

- $\mathcal{H}$ is a separable Hilbert space,
- A is the generator of a strongly continuous semigroup $(S_t)_{t\geq 0}$,
- W_t is a cylindrical Wiener process on $\mathcal{H}$,
- F and G are nonlinear drift and diffusion maps.

The challenge lies in defining, constructing, and analyzing the process (X_t) as an $\mathcal{H}$-valued stochastic process with well-defined regularity and Markov properties.

12.2 Stochastic Integration in Hilbert Spaces

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space satisfying the usual conditions of right-continuity and completeness. Let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and norm $\|\cdot\|_{\mathcal{H}}$.

Cylindrical Wiener process. A *cylindrical Wiener process* $(W_t)_{t\geq 0}$ on $\mathcal{H}$ is a family of linear maps

$$W_t : \mathcal{H} \to L^2(\Omega, \mathcal{F}, \mathbb{P}), \qquad t \geq 0,$$

such that for every $h \in \mathcal{H}$, the real-valued process $\{W_t(h)\}_{t\geq 0}$ is a standard Brownian motion satisfying

$$\mathbb{E}[W_t(h)] = 0, \qquad \mathrm{Cov}(W_s(h_1), W_t(h_2)) = (s \wedge t)\langle h_1, h_2\rangle_{\mathcal{H}}.$$

Equivalently, if $\{e_k\}_{k\geq 1}$ is a fixed orthonormal basis of $\mathcal{H}$ and $\{\beta_k(t)\}_{k\geq 1}$ is a family of independent standard Brownian motions, one can write formally

$$W_t = \sum_{k=1}^{\infty} \beta_k(t) e_k, \qquad t \geq 0, \tag{12.2}$$

where the sum is not convergent in $\mathcal{H}$ but defines a Gaussian process in a larger space of generalized random variables.

Hilbert–Schmidt operators. Let $\mathcal{L}_2(\mathcal{H})$ denote the space of Hilbert–Schmidt operators on $\mathcal{H}$, i.e.

$$\mathcal{L}_2(\mathcal{H}) = \{T : \mathcal{H} \to \mathcal{H} \text{ linear and bounded} \mid \|T\|^2_{\mathcal{L}_2(\mathcal{H})} = \sum_{k=1}^{\infty} \|Te_k\|^2_{\mathcal{H}} < \infty\}.$$

This space is a separable Hilbert space endowed with the inner product

$$\langle S, T\rangle_{\mathcal{L}_2(\mathcal{H})} = \sum_{k=1}^{\infty} \langle Se_k, Te_k\rangle_{\mathcal{H}},$$

independent of the chosen orthonormal basis $\{e_k\}$.

Predictable operator-valued processes. A process $\Phi = (\Phi_t)_{t\in[0,T]}$ is said to be *predictable* with values in $\mathcal{L}_2(\mathcal{H})$ if

$$\Phi : \Omega \times [0, T] \to \mathcal{L}_2(\mathcal{H})$$

is measurable with respect to the predictable σ-algebra $\mathcal{P}$ on $\Omega \times [0, T]$ and satisfies

$$\mathbb{E} \int_0^T \|\Phi_t\|^2_{\mathcal{L}_2(\mathcal{H})}\, dt < \infty.$$

The set of all such processes is denoted by $L^2_{\mathcal{P}}([0, T]; \mathcal{L}_2(\mathcal{H}))$.

Definition of the stochastic integral. For any $\Phi \in L^2_{\mathcal{P}}([0, T]; \mathcal{L}_2(\mathcal{H}))$, the stochastic integral of Φ with respect to the cylindrical Wiener process W_t is defined as

$$\int_0^t \Phi_s\, dW_s := \sum_{k=1}^{\infty} \int_0^t \Phi_s e_k\, d\beta_k(s), \tag{12.3}$$

where the stochastic integrals on the right-hand side are Itô integrals in the sense of real-valued stochastic calculus. The convergence of the series is in $L^2(\Omega; \mathcal{H})$, since

$$\mathbb{E} \left\| \sum_{k=1}^{n} \int_0^t \Phi_s e_k\, d\beta_k(s) - \sum_{k=1}^{m} \int_0^t \Phi_s e_k\, d\beta_k(s) \right\|^2_{\mathcal{H}} \to 0 \quad \text{as } n, m \to \infty.$$

Hence, $\int_0^t \Phi_s \, dW_s$ is well-defined as an element of $L^2(\Omega; \mathcal{H})$.

Itô isometry. The stochastic integral satisfies the fundamental *Itô isometry*:

$$\mathbb{E}\left[\left\|\int_0^t \Phi_s \, dW_s\right\|_{\mathcal{H}}^2\right] = \mathbb{E}\left[\int_0^t \|\Phi_s\|_{\mathcal{L}_2(\mathcal{H})}^2 \, ds\right]. \tag{12.4}$$

Proof. By the definition of the stochastic integral and the orthogonality of Brownian motions β_k,

$$\mathbb{E}\left\|\int_0^t \Phi_s \, dW_s\right\|_{\mathcal{H}}^2 = \mathbb{E}\left[\sum_{k=1}^{\infty}\left\|\int_0^t \Phi_s e_k \, d\beta_k(s)\right\|_{\mathcal{H}}^2\right] = \mathbb{E}\left[\sum_{k=1}^{\infty}\int_0^t \|\Phi_s e_k\|_{\mathcal{H}}^2 \, ds\right],$$

which, by Fubini's theorem, yields

$$\mathbb{E}\left[\int_0^t \|\Phi_s\|_{\mathcal{L}_2(\mathcal{H})}^2 \, ds\right].$$

This proves the identity.

Interpretation and applications. The process $\int_0^t \Phi_s \, dW_s$ is an $\mathcal{H}$-valued square-integrable martingale with quadratic variation

$$[\![M]\!]_t = \int_0^t \|\Phi_s\|_{\mathcal{L}_2(\mathcal{H})}^2 \, ds.$$

This construction provides the rigorous analytic foundation for stochastic integration in infinite-dimensional spaces, thereby allowing the formulation of stochastic evolution equations (SPDEs) of the form

$$dX_t = AX_t \, dt + \Phi_t \, dW_t, \qquad X_0 = x_0 \in \mathcal{H}, \tag{12.5}$$

where A is a (possibly unbounded) generator of a C_0-semigroup on $\mathcal{H}$ and Φ_t is an admissible Hilbert–Schmidt-valued integrand. Such equations form the analytical backbone of modern infinite-dimensional stochastic analysis, encompassing models in quantum field theory, fluid dynamics, and financial mathematics.

12.3 Mild and Weak Solutions

Consider the abstract stochastic evolution equation on a separable Hilbert space $\mathcal{H}$,

$$dX_t = AX_t \, dt + G \, dW_t, \qquad X_0 = x \in \mathcal{H}, \tag{12.6}$$

where

- $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ is the infinitesimal generator of a strongly continuous semigroup $(S_t)_{t\geq 0}$ on $\mathcal{H}$,
- $G \in \mathcal{L}_2(\mathcal{H})$ is a bounded Hilbert–Schmidt operator representing the diffusion coefficient, and
- $(W_t)_{t\geq 0}$ is a cylindrical Wiener process on $\mathcal{H}$.

12.3.1 Mild solution

Formally integrating the stochastic differential equation from 0 to t and applying the variation-of-constants (Duhamel) principle, we obtain

$$X_t = S_t x + \int_0^t S_{t-s} A X_s \, ds + \int_0^t S_{t-s} G \, dW_s.$$

However, when A is unbounded, the deterministic integral may not be well-defined in the classical sense. Hence, one defines the *mild solution* directly as

$$X_t = S_t x + \int_0^t S_{t-s} G \, dW_s, \qquad t \geq 0. \tag{12.7}$$

The stochastic convolution term

$$W_A(t) := \int_0^t S_{t-s} G \, dW_s$$

is an $\mathcal{H}$-valued Gaussian process with mean zero and covariance operator

$$Q_t = \int_0^t S_s G G^* S_s^* \, ds.$$

By the Itô isometry, this process satisfies

$$\mathbb{E}\|W_A(t)\|_{\mathcal{H}}^2 = \int_0^t \|S_{t-s} G\|_{\mathcal{L}_2(\mathcal{H})}^2 \, ds.$$

If the above integral is finite for all $t > 0$, then $W_A(t)$ is well-defined, and thus the mild solution X_t exists uniquely. Moreover, under these conditions, $(X_t)_{t\geq 0}$ is an $\mathcal{H}$-valued adapted process satisfying

$$\mathbb{E}\|X_t\|_{\mathcal{H}}^2 < \infty, \qquad \forall t \geq 0.$$

12.3.2 Weak solution

Let A^* denote the adjoint of A with domain $\mathcal{D}(A^*) \subset \mathcal{H}$. A progressively measurable process $(X_t)_{t\geq 0}$ with values in $\mathcal{H}$ is called a *weak solution* to the above stochastic evolution equation if, for every test function $\phi \in \mathcal{D}(A^*)$, the following holds $\mathbb{P}$-almost surely for all $t \geq 0$:

$$\langle X_t, \phi \rangle_{\mathcal{H}} = \langle x, \phi \rangle_{\mathcal{H}} + \int_0^t \langle X_s, A^* \phi \rangle_{\mathcal{H}} \, ds + \int_0^t \langle G \, dW_s, \phi \rangle_{\mathcal{H}}. \tag{12.8}$$

The last term is well-defined as an Itô stochastic integral in $\mathbb{R}$ since

$$\langle G \, dW_s, \phi \rangle_{\mathcal{H}} = \sum_{k=1}^{\infty} \langle G e_k, \phi \rangle_{\mathcal{H}} \, d\beta_k(s) \tag{12.9}$$

for the orthonormal basis $\{e_k\}$ of $\mathcal{H}$ used in the cylindrical decomposition of W_t.

12.3.3 Equivalence between mild and weak solutions

Under standard assumptions—namely:

- the operator A generates a C_0-semigroup (S_t) on $\mathcal{H}$,
- the operator G is Hilbert–Schmidt, and
- $\int_0^t \|S_{t-s}G\|^2_{\mathcal{L}_2(\mathcal{H})}\,ds < \infty$ for all $t > 0$,

the mild solution $X_t = S_t x + \int_0^t S_{t-s}G\,dW_s$ is also a weak solution in the above sense. To see this, note that for $\phi \in \mathcal{D}(A^*)$,

$$\langle X_t, \phi\rangle_{\mathcal{H}} = \langle S_t x, \phi\rangle_{\mathcal{H}} + \int_0^t \langle S_{t-s}G\,dW_s, \phi\rangle_{\mathcal{H}} = \langle x, S_t^*\phi\rangle_{\mathcal{H}} + \int_0^t \langle G\,dW_s, S_{t-s}^*\phi\rangle_{\mathcal{H}},$$

and using that $\frac{d}{dt}S_t^*\phi = S_t^*A^*\phi$, integration by parts yields the weak formulation.

Conversely, if (X_t) is a weak solution satisfying appropriate integrability and measurability conditions, then the mild representation holds via the variation-of-constants formula. Hence, under these assumptions, the notions of mild and weak solutions are equivalent.

This equivalence is fundamental in the study of stochastic evolution equations, as the mild formulation provides an explicit representation through stochastic convolution, while the weak formulation offers a variational characterization useful for duality and regularity analysis.

12.4 The Linear Stochastic Heat Equation

We now examine one of the most fundamental examples of an infinite-dimensional stochastic evolution equation: the *stochastic heat equation* driven by additive space–time white noise. This model serves as a canonical example illustrating the abstract theory of stochastic integration and mild solutions in Hilbert spaces.

Formulation of the problem

Let $\mathcal{H} = L^2(0,1)$ be the Hilbert space of square-integrable real-valued functions on the unit interval $(0,1)$, equipped with the standard inner product

$$\langle f, g\rangle_{L^2(0,1)} = \int_0^1 f(x)g(x)\,dx. \tag{12.10}$$

We consider the stochastic partial differential equation (SPDE)

$$\partial_t u(t,x) = \Delta u(t,x) + \dot{W}(t,x), \qquad (t,x) \in (0,\infty) \times (0,1), \tag{12.11}$$

subject to the Dirichlet boundary conditions

$$u(t,0) = u(t,1) = 0, \qquad t \geq 0, \tag{12.12}$$

and an initial condition

$$u(0,x) = u_0(x), \qquad x \in (0,1). \tag{12.13}$$

Here $\Delta = \frac{\partial^2}{\partial x^2}$ denotes the Laplace operator and $\dot{W}(t,x)$ represents a Gaussian noise that is white in both space and time, formally satisfying

$$\mathbb{E}[\dot{W}(t,x)\dot{W}(s,y)] = \delta(t-s)\delta(x-y). \tag{12.14}$$

Abstract reformulation

We now recast the SPDE into an abstract stochastic evolution equation in the Hilbert space $\mathcal{H}$. Define the (unbounded) operator $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ by

$$Au = \Delta u, \qquad \mathcal{D}(A) = H^2(0,1) \cap H_0^1(0,1), \tag{12.15}$$

where $H^2(0,1)$ and $H_0^1(0,1)$ are Sobolev spaces incorporating the Dirichlet boundary conditions. It is well known that A generates a strongly continuous analytic semigroup $(S_t)_{t\geq 0}$ on $\mathcal{H}$ given by

$$S_t = e^{tA}. \tag{12.16}$$

Let $(W_t)_{t\geq 0}$ be a cylindrical Wiener process on $\mathcal{H}$, i.e.

$$W_t = \sum_{k=1}^{\infty} \beta_k(t)\phi_k, \tag{12.17}$$

where $\{\beta_k(t)\}_{k\geq 1}$ are independent real-valued standard Brownian motions, and $\{\phi_k\}_{k\geq 1}$ is a complete orthonormal system in $\mathcal{H}$. Then the SPDE can be expressed compactly as

$$du_t = Au_t\,dt + dW_t, \qquad u_0 = u_0 \in \mathcal{H}. \tag{12.18}$$

Mild solution

Formally integrating both sides and applying the semigroup property, the *mild solution* is defined by

$$u_t = S_t u_0 + \int_0^t S_{t-s}\,dW_s. \tag{12.19}$$

The stochastic integral is well-defined in $\mathcal{H}$ since S_{t-s} is a Hilbert–Schmidt operator from $\mathcal{H}$ into itself for $t > s$. Using the Itô isometry for $\mathcal{H}$-valued integrals, we obtain

$$\mathbb{E}\|u_t\|_{\mathcal{H}}^2 = \int_0^t \|S_{t-s}\|_{\mathcal{L}_2(\mathcal{H})}^2\,ds < \infty. \tag{12.20}$$

Hence, the mild solution u_t is a well-defined $\mathcal{H}$-valued random variable for every $t \geq 0$.

Spectral decomposition

The eigenfunctions of the Dirichlet Laplacian on $(0,1)$ are

$$\phi_k(x) = \sqrt{2}\sin(k\pi x), \qquad k \in \mathbb{N}, \tag{12.21}$$

which form an orthonormal basis of $L^2(0,1)$, with corresponding eigenvalues

$$\lambda_k = -k^2\pi^2. \tag{12.22}$$

Thus, the semigroup (S_t) acts diagonally in this basis:

$$S_t\phi_k = e^{\lambda_k t}\phi_k = e^{-k^2\pi^2 t}\phi_k. \tag{12.23}$$

Expanding u_t in this basis,

$$u_t(x) = \sum_{k=1}^{\infty} u_k(t)\, \phi_k(x), \tag{12.24}$$

the SPDE decouples into an infinite system of one-dimensional stochastic differential equations:

$$du_k(t) = -k^2\pi^2 u_k(t)\, dt + d\beta_k(t), \qquad u_k(0) = 0. \tag{12.25}$$

Each $u_k(t)$ satisfies

$$u_k(t) = \int_0^t e^{-k^2\pi^2(t-s)}\, d\beta_k(s). \tag{12.26}$$

Therefore, the mild solution admits the explicit representation

$$u_t(x) = \sum_{k=1}^{\infty} \left(\int_0^t e^{-k^2\pi^2(t-s)}\, d\beta_k(s) \right) \phi_k(x). \tag{12.27}$$

Covariance structure and Gaussianity

Since each $u_k(t)$ is a Gaussian random variable with mean zero and variance

$$\mathbb{E}[u_k(t)^2] = \int_0^t e^{-2k^2\pi^2(t-s)}\, ds = \frac{1 - e^{-2k^2\pi^2 t}}{2k^2\pi^2}, \tag{12.28}$$

the process u_t is a centered Gaussian random element in $\mathcal{H} = L^2(0,1)$ with covariance operator Q_t given by

$$Q_t = \int_0^t S_{t-s} S_{t-s}^*\, ds. \tag{12.29}$$

In the spectral basis $\{\phi_k\}$, this covariance operator is diagonal:

$$Q_t \phi_k = \frac{1 - e^{-2k^2\pi^2 t}}{2k^2\pi^2}\, \phi_k. \tag{12.30}$$

Hence,

$$\mathbb{E}\|u_t\|_{\mathcal{H}}^2 = \sum_{k=1}^{\infty} \frac{1 - e^{-2k^2\pi^2 t}}{2k^2\pi^2} < \infty. \tag{12.31}$$

Regularity properties

The process $(u_t)_{t\geq 0}$ is mean-square continuous in $\mathcal{H}$ and has trajectories that are almost surely continuous in time with values in $L^2(0,1)$. However, due to the irregular nature of the space–time white noise, spatial regularity is limited: $u_t(\cdot)$ belongs almost surely to Sobolev spaces $H^\alpha(0,1)$ only for $\alpha < \frac{1}{2}$.

Thus, the linear stochastic heat equation provides a canonical illustration of the interplay between semigroup theory, stochastic integration in Hilbert spaces, and the spectral analysis of SPDEs.

12.5 Nonlinear SPDEs and Monotone Operators

We now consider nonlinear stochastic partial differential equations (SPDEs) of the form

$$dX_t = (AX_t + F(X_t))\,dt + G(X_t)\,dW_t, \qquad X_0 = x \in \mathcal{H}, \tag{12.32}$$

where A is typically a (possibly unbounded) linear operator generating a C_0-semigroup on a separable Hilbert space $\mathcal{H}$, and F and G are nonlinear mappings satisfying appropriate monotonicity, coercivity, and growth conditions. The driving noise $(W_t)_{t\geq 0}$ is a cylindrical Wiener process on $\mathcal{H}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

Variational setting and Gelfand triple

To rigorously treat nonlinear SPDEs, one employs the *variational* or *monotone operator* framework. Let V be a reflexive Banach space continuously and densely embedded in a separable Hilbert space $\mathcal{H}$, which in turn is continuously embedded in the dual space V^*. This configuration forms the *Gelfand triple*

$$V \subset \mathcal{H} \subset V^*, \tag{12.33}$$

where the embedding $V \hookrightarrow \mathcal{H}$ is continuous and dense, and we identify $\mathcal{H}$ with its dual via the Riesz representation theorem. Hence, for $v \in V$ and $h \in \mathcal{H}$,

$$\langle v, h\rangle_{V,V^*} = (v, h)_{\mathcal{H}}. \tag{12.34}$$

Monotonicity and coercivity conditions

The key to establishing existence and uniqueness for (12.32) lies in verifying structural properties of the drift term $A + F$. Specifically, we assume:

1. **Hemicontinuity:** For all $u, v, w \in V$, the map

$$\lambda \mapsto \langle A(u + \lambda v) + F(u + \lambda v), w\rangle_{V,V^*} \tag{12.35}$$

   is continuous on $\mathbb{R}$.

2. **Monotonicity:** There exists a constant $C \in \mathbb{R}$ such that for all $x, y \in V$,

$$\langle F(x) - F(y), x - y\rangle_{V,V^*} \leq C\|x - y\|_{\mathcal{H}}^2. \tag{12.36}$$

3. **Coercivity:** There exist constants $\alpha > 0$ and $\beta \geq 0$ such that for all $x \in V$,

$$\langle Ax + F(x), x\rangle_{V,V^*} \leq -\alpha\|x\|_V^2 + \beta. \tag{12.37}$$

4. **Growth condition:** There exists $K > 0$ such that for all $x \in V$,

$$\|Ax + F(x)\|_{V^*} \leq K(1 + \|x\|_V). \tag{12.38}$$

For the diffusion coefficient $G : V \to \mathcal{L}_2(\mathcal{H})$ (Hilbert–Schmidt operators on $\mathcal{H}$), we require:

$$\|G(x) - G(y)\|^2_{\mathcal{L}_2(\mathcal{H})} \le L_G \|x - y\|^2_{\mathcal{H}}, \tag{12.39}$$

$$\|G(x)\|^2_{\mathcal{L}_2(\mathcal{H})} \le C_G(1 + \|x\|^2_{\mathcal{H}}), \tag{12.40}$$

for some constants $L_G, C_G > 0$.

Definition of variational solution

A predictable process $(X_t)_{t\ge 0}$ taking values in V is called a *variational (or weak) solution* to (12.32) if, for every test function $\phi \in V$, it satisfies almost surely:

$$\langle X_t, \phi\rangle_{V,V^*} = \langle x, \phi\rangle_{V,V^*} + \int_0^t \langle AX_s + F(X_s), \phi\rangle_{V,V^*}\, ds + \int_0^t \langle G(X_s)\, dW_s, \phi\rangle_{\mathcal{H}}. \tag{12.41}$$

Additionally, X_t must satisfy the integrability conditions:

$$X \in L^2(\Omega; C([0,T]; \mathcal{H})) \cap L^2(\Omega; L^2(0,T;V)). \tag{12.42}$$

Existence and uniqueness theorem

Under the above hypotheses (12.36)–(12.40), there exists a unique variational solution X_t to (12.32) satisfying

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t\|^2_{\mathcal{H}} + \int_0^T \|X_t\|^2_V\, dt\right] < \infty. \tag{12.43}$$

The proof is based on a Galerkin approximation combined with monotonicity methods (Minty–Browder technique), Itô's formula, and weak compactness arguments in the variational framework.

Energy estimate

Applying Itô's formula to $\|X_t\|^2_{\mathcal{H}}$, we obtain:

$$d\|X_t\|^2_{\mathcal{H}} = 2\langle AX_t + F(X_t), X_t\rangle_{V,V^*}\, dt + \|G(X_t)\|^2_{\mathcal{L}_2(\mathcal{H})}\, dt + 2\langle X_t, G(X_t)\, dW_t\rangle_{\mathcal{H}}. \tag{12.44}$$

Taking expectations and using (12.37) and (12.40), we obtain the a priori estimate

$$\mathbb{E}\|X_t\|^2_{\mathcal{H}} + 2\alpha\mathbb{E}\int_0^t \|X_s\|^2_V\, ds \le \|x\|^2_{\mathcal{H}} + C\int_0^t (1 + \mathbb{E}\|X_s\|^2_{\mathcal{H}})\, ds. \tag{12.45}$$

By Grönwall's inequality, this yields

$$\sup_{t\in[0,T]} \mathbb{E}\|X_t\|^2_{\mathcal{H}} < \infty. \tag{12.46}$$

Examples

This general framework includes several important nonlinear SPDEs as special cases:

- **Stochastic porous medium equation:**

$$du_t = \Delta(u_t^m)\, dt + \sigma(u_t)\, dW_t, \qquad m > 1. \tag{12.47}$$

- **Stochastic reaction–diffusion equation:**

$$du_t = (\Delta u_t + f(u_t))\,dt + g(u_t)\,dW_t. \tag{12.48}$$

- **Stochastic Navier–Stokes equations (in 2D):**

$$du_t = [\nu\Delta u_t - (u_t \cdot \nabla)u_t - \nabla p_t]\,dt + B(u_t)\,dW_t, \qquad \nabla \cdot u_t = 0. \tag{12.49}$$

Hence, the variational framework for monotone SPDEs provides a unified and rigorous setting encompassing a broad class of nonlinear infinite-dimensional stochastic systems, ensuring existence, uniqueness, and energy stability of their solutions.

12.6 Invariant Measures and Ergodicity

Let $(X_t)_{t\geq 0}$ be a Markov process on a separable Hilbert space $\mathcal{H}$ with transition semigroup $(P_t)_{t\geq 0}$ acting on bounded measurable functions $f : \mathcal{H} \to \mathbb{R}$ via

$$(P_t f)(x) = \mathbb{E}^x[f(X_t)].$$

A probability measure μ on $\mathcal{H}$ is said to be an *invariant measure* for (P_t) if, for every bounded measurable $f : \mathcal{H} \to \mathbb{R}$,

$$\int_{\mathcal{H}} (P_t f)(x)\,\mu(dx) = \int_{\mathcal{H}} f(x)\,\mu(dx), \qquad \forall t \geq 0. \tag{12.50}$$

Equivalently, if $X_0 \sim \mu$, then $X_t \sim \mu$ for all $t \geq 0$, i.e. μ is stationary for the process.

Gaussian invariant measure for linear systems

Consider the linear stochastic evolution equation on $\mathcal{H}$:

$$dX_t = AX_t\,dt + G\,dW_t, \qquad X_0 = x \in \mathcal{H}, \tag{12.51}$$

where:

- $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ is the infinitesimal generator of a strongly continuous semigroup $(S_t)_{t\geq 0}$ on $\mathcal{H}$,
- $G \in \mathcal{L}_2(\mathcal{H})$ is a bounded Hilbert–Schmidt operator (diffusion coefficient),
- $(W_t)_{t\geq 0}$ is a cylindrical Wiener process on $\mathcal{H}$.

The mild solution of (12.51) is

$$X_t = S_t x + \int_0^t S_{t-s} G\,dW_s. \tag{12.52}$$

This process (X_t) is known as the *Ornstein–Uhlenbeck process* in infinite dimensions.

Existence of invariant measure

Assume that A is *dissipative*, meaning

$$\langle Ax, x\rangle_{\mathcal{H}} \le -\omega \|x\|_{\mathcal{H}}^2, \qquad \forall x \in \mathcal{D}(A), \tag{12.53}$$

for some $\omega > 0$. This implies that $\|S_t\|_{\mathcal{L}(\mathcal{H})} \le e^{-\omega t}$, hence (S_t) is exponentially stable.

Define the covariance operator

$$Q_\infty := \int_0^\infty S_t G G^* S_t^* \, dt. \tag{12.54}$$

Because S_t decays exponentially, the integral converges in the trace-class operator sense. The operator Q_∞ is self-adjoint, positive semidefinite, and satisfies the *Lyapunov equation*

$$AQ_\infty + Q_\infty A^* + GG^* = 0. \tag{12.55}$$

To verify this, note that differentiating the finite-time covariance

$$Q_t = \int_0^t S_s G G^* S_s^* \, ds$$

gives

$$\frac{dQ_t}{dt} = S_t G G^* S_t^* = AQ_t + Q_t A^* + GG^*,$$

and taking the limit as $t \to \infty$ (using the stability of S_t) yields (12.55).

Characterization of the invariant measure

Define the Gaussian measure $\mu_\infty := \mathcal{N}(0, Q_\infty)$ on $\mathcal{H}$ with mean zero and covariance operator Q_∞. Then μ_∞ satisfies (12.50), i.e.

$$\int_{\mathcal{H}} (P_t f)(x)\, \mu_\infty(dx) = \int_{\mathcal{H}} f(x)\, \mu_\infty(dx), \qquad \forall t \ge 0.$$

Indeed, since X_t is Gaussian with mean $S_t x$ and covariance Q_t, if $X_0 \sim \mu_\infty$ then

$$X_t \sim \mathcal{N}(S_t x, Q_t + S_t Q_\infty S_t^*),$$

and using the identity $Q_t + S_t Q_\infty S_t^* = Q_\infty$ (which follows from (12.54)) we deduce stationarity of μ_∞.

Uniqueness and ergodicity

If the pair (A, G) is *controllable* in the sense that

$$\overline{\mathrm{Range}(Q_\infty^{1/2})} = \mathcal{H},$$

then μ_∞ has full support on $\mathcal{H}$ and is the *unique* invariant probability measure of (X_t).

Furthermore, (X_t) is *ergodic* with respect to μ_∞, i.e. for every $f \in L^1(\mathcal{H}, \mu_\infty)$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(X_t)\,dt = \int_{\mathcal{H}} f(x)\,\mu_\infty(dx) \quad \text{a.s.} \tag{12.56}$$

This property follows from the exponential stability of S_t and mixing of the Gaussian process.

Summary: the Ornstein–Uhlenbeck semigroup

The semigroup (P_t) associated with (12.51) acts on bounded continuous functions via

$$(P_t f)(x) = \int_{\mathcal{H}} f(S_t x + y)\,\mathcal{N}(0, Q_t)(dy),$$

and is strongly Feller and irreducible when G is non-degenerate. It satisfies the generator identity

$$\mathcal{L}f(x) = \langle Ax, Df(x)\rangle_{\mathcal{H}} + \frac{1}{2}\operatorname{Tr}\left(GG^* D^2 f(x)\right),$$

and admits $\mu_\infty = \mathcal{N}(0, Q_\infty)$ as its unique invariant measure.

Thus, under dissipativity of A and non-degeneracy of G, the infinite-dimensional Ornstein–Uhlenbeck process is *Gaussian, stationary, ergodic*, and fully characterized by the Lyapunov equation (12.55).

12.7 Malliavin Calculus and Regularity of Laws

Malliavin calculus furnishes a differential calculus on Wiener space which permits one to study regularity of probability laws of functionals of Gaussian processes, and — by extension — of solutions to stochastic (partial) differential equations. In what follows we work on a probability space carrying a cylindrical (or genuine) Wiener process W taking values in a separable Hilbert space $\mathcal{U}$. Denote by $\mathcal{H}$ the Cameron–Martin space associated with W (for an $\mathcal{U}$-cylindrical Wiener process on $[0, T]$ one typically has $\mathcal{H} = L^2([0, T]; \mathcal{U})$). We adopt the usual Wiener–Itô isometry identifications and use $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ for the inner product on $\mathcal{H}$.

Let F be a smooth cylindrical functional of the form

$$F = f(W(h_1), \ldots, W(h_n)), \tag{12.57}$$

where $h_i \in \mathcal{H}$, $W(h) = \int_0^T \langle h(s), dW_s\rangle_{\mathcal{U}}$, and $f \in C_b^\infty(\mathbb{R}^n)$. The *Malliavin derivative* D is defined on such F by the $\mathcal{H}$-valued random variable

$$(DF)(\omega) \;=\; \sum_{i=1}^n \partial_i f(W(h_1), \ldots, W(h_n))\, h_i \in \mathcal{H}. \tag{12.58}$$

Iterating, one obtains higher-order derivatives $D^k F \in L^2(\Omega; \mathcal{H}^{\otimes k})$. The derivative operator D is closable from $L^p(\Omega)$ into $L^p(\Omega; \mathcal{H})$, and one defines the Sobolev–Malliavin spaces

$$\mathbb{D}^{k,p} := \overline{\{\,\text{smooth cylindrical } F\,\}}^{\|\cdot\|_{k,p}}, \qquad \|F\|_{k,p}^p := \mathbb{E}[|F|^p] + \sum_{j=1}^k \mathbb{E}[\|D^j F\|_{\mathcal{H}^{\otimes j}}^p]. \tag{12.59}$$

The adjoint of D (in the sense of the L^2-pairing) is the divergence (or Skorohod) operator δ. For an adapted (or more generally predictable) $\mathcal{H}$-valued process u satisfying suitable integrability, $\delta(u)$ coincides with the Itô integral $\int_0^T \langle u_s, dW_s\rangle_{\mathcal{U}}$; for nonadapted u the divergence extends this integral in the Skorohod sense. The fundamental duality formula reads, for appropriate $F \in \mathbb{D}^{1,2}$ and $u \in \mathrm{Dom}(\delta)$,

$$\mathbb{E}\big[\langle DF, u\rangle_{\mathcal{H}}\big] = \mathbb{E}\big[F\,\delta(u)\big]. \tag{12.60}$$

A principal application is the study of existence and smoothness of densities of random vectors. Let $F = (F^1, \dots, F^m)$ be an $\mathbb{R}^m$-valued random variable with components in $\mathbb{D}^{1,2}$. Its *Malliavin covariance matrix* σ_F is the $m \times m$ symmetric, nonnegative-definite random matrix

$$\sigma_F := \big(\langle DF^i, DF^j\rangle_{\mathcal{H}}\big)_{1\le i,j\le m}. \tag{12.61}$$

A cornerstone result (Bouleau–Hirsch) states that if $F \in \mathbb{D}^{1,2}$ and σ_F is almost surely invertible with inverse having moments of all orders (or at least $\mathbb{E}[\det(\sigma_F)^{-p}] < \infty$ for some $p > 1$), then the law of F is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^m$. Moreover, higher-order Malliavin differentiability together with bounds on σ_F^{-1} yield C^∞-smoothness of the density. Precisely, if $F \in \bigcap_{k\ge1} \mathbb{D}^{k,p}$ for all p and $\det \sigma_F^{-1} \in \bigcap_{p\ge1} L^p(\Omega)$, then the density of F belongs to $C^\infty(\mathbb{R}^m)$.

In infinite-dimensional settings one rarely seeks a density with respect to an infinite-dimensional "Lebesgue" measure; instead one studies finite-dimensional projections or densities with respect to Gaussian/reference measures. Thus, for an $\mathcal{H}$-valued random variable X (e.g. the solution of an SPDE), one typically examines the law of $\ell(X)$ for finite-rank linear functionals $\ell : \mathcal{H} \to \mathbb{R}^m$. If X is Malliavin differentiable and the projected covariance

$$\sigma_{\ell(X)} = \big(\langle D\langle X, \ell_i\rangle, D\langle X, \ell_j\rangle\rangle_{\mathcal{H}}\big)_{i,j} \tag{12.62}$$

is nondegenerate in the Bouleau–Hirsch sense, then $\ell(X)$ admits a smooth density.

A central class of applications concerns solutions to (finite- or infinite-dimensional) SDEs/SPDEs. Consider the abstract stochastic evolution equation

$$dX_t = (AX_t + F(X_t))\,dt + G(X_t)\,dW_t, \qquad X_0 = x, \tag{12.63}$$

posed in a Hilbert (or Banach) space. Under sufficient regularity of F and G (Fréchet differentiability, Lipschitz-type bounds) one can prove that the solution map $\omega \mapsto X_t(\omega)$ is Malliavin differentiable, and the Malliavin derivative $D_sX_t \in \mathcal{L}(\mathcal{H};\mathcal{H})$ (or an $\mathcal{H}$-valued kernel) satisfies the linearized (variation) equation obtained by differentiating the SPDE in direction $h \in \mathcal{H}$. For $0 \le s \le t$ one has, in mild form,

$$D_sX_t = S_{t-s}G(X_s) + \int_s^t S_{t-r}\nabla F(X_r)\,D_sX_r\,dr + \int_s^t S_{t-r}\nabla G(X_r)\,D_sX_r\,dW_r, \tag{12.64}$$

where S_t denotes the semigroup generated by A, and $\nabla F, \nabla G$ are appropriate Fréchet derivatives (these integrals are to be interpreted in the Ito/Skorohod sense depending on adaptivity). When G is nondegenerate and the linearized equation is sufficiently controllable, one may deduce nondegeneracy of the finite-dimensional Malliavin covariance matrices.

In finite dimensions the celebrated Hörmander condition yields hypoellipticity: if the Lie algebra generated by the vector fields appearing in the SDE spans $\mathbb{R}^n$ at the point of interest, then the Malliavin matrix is invertible and the density is smooth. In infinite-dimensional or SPDE contexts one often uses finite-dimensional projection versions of Hörmander's bracket condition (or its adapted infinite-dimensional analogues) to obtain smoothness of finite-dimensional marginals. Concretely, suppose for a family of linear functionals $\ell_1, \ldots, \ell_m$ the Lie-bracket-generated reachable space (formed by iterated commutators of the drift and noise directions projected by ℓ) spans $\mathbb{R}^m$; then $\ell(X_t)$ has a smooth density.

The Malliavin covariance operator associated to an $\mathcal{H}$-valued random variable X is the trace-class, symmetric operator $\mathcal{C}_X : \mathcal{H} \to \mathcal{H}$ given by

$$\mathcal{C}_X = \mathbb{E}\big[(DX) \otimes (DX)\big] = \mathbb{E}\big[\,(DX)\,\langle DX, \cdot\rangle_{\mathcal{H}}\,\big], \tag{12.65}$$

and for a finite-rank projection $P : \mathcal{H} \to \mathbb{R}^m$ one recovers $\sigma_{P(X)} = P\,\mathcal{C}_X\,P^*$. In many SPDEs one proves that $\mathcal{C}_X$ is positive definite on the range of the projection, or equivalently that $\langle \mathcal{C}_X u, u\rangle_{\mathcal{H}} > 0$ for all nonzero u in a finite-dimensional subspace; this suffices to obtain absolute continuity of the projected law.

Integration by parts formulae derived from Malliavin calculus underpin density formulas and quantitative bounds. If $F \in \mathbb{D}^{1,2}$ and $\varphi \in C_c^\infty(\mathbb{R}^m)$, then repeated application of the duality between D and δ yields

$$\mathbb{E}\big[\partial_{i_1} \cdots \partial_{i_k} \varphi(F)\big] = \mathbb{E}\big[\varphi(F)\, H_{i_1,\ldots,i_k}(F)\big], \tag{12.66}$$

where the random weights $H_{i_1,\ldots,i_k}(F)$ are expressed in terms of iterated divergences involving σ_F^{-1} and Malliavin derivatives of F. From these identities one deduces that the density of F (when it exists) admits classical Sobolev and smoothness estimates controlled by moments of σ_F^{-1} and norms of derivatives of F.

Concerning SPDEs, typical rigorous results take the following form. Suppose the coefficients F and G are sufficiently smooth with bounded derivatives, and that the noise acts nondegenerately in a manner satisfying a suitable controllability (or bracket) condition on a finite-dimensional projection. Then for every $t > 0$ and every finite-rank bounded linear operator $P : \mathcal{H} \to \mathbb{R}^m$, the random vector PX_t belongs to $\bigcap_{k,p} \mathbb{D}^{k,p}(\mathbb{R}^m)$ and has a C^∞-density with respect to Lebesgue measure on $\mathbb{R}^m$. Moreover, if the Malliavin covariance of PX_t admits uniform lower bounds (ellipticity) one obtains Gaussian-type lower and upper bounds for the density via Norris' lemma and heat-kernel techniques adapted to the Malliavin setting.

Finally, Malliavin calculus provides tools for quantitative regularity: existence of densities in Sobolev spaces $W^{s,p}$, short-time asymptotics, gradient estimates for densities, and bounds on derivatives of transition probabilities. These results are central in establishing Hörmander-type hypoellipticity for SPDEs, smoothing properties of their transition semigroups, and precise control on the probabilistic representation of solutions to PDEs via probabilistic methods.

12.8 Spectral and Functional Analytic Structure

The study of the spectral and functional analytic properties of stochastic evolution equations is centered on the infinitesimal generator of their associated Markov semigroup. Let $(X_t^x)_{t\geq 0}$ denote the mild solution to the stochastic evolution equation

$$dX_t = (AX_t + F(X_t))\,dt + G\,dW_t, \qquad X_0 = x \in \mathcal{H}, \tag{12.67}$$

where A is the generator of a C_0-semigroup $(S_t)_{t\geq 0}$ on a separable Hilbert space $\mathcal{H}$, $F : \mathcal{H} \to \mathcal{H}$ is a (possibly nonlinear) drift term satisfying suitable Lipschitz and growth conditions, and $G \in \mathcal{L}(\mathcal{U}, \mathcal{H})$ is a bounded linear operator defining the covariance structure of a $\mathcal{U}$-valued cylindrical Wiener process W_t.

For every bounded Borel measurable function $f : \mathcal{H} \to \mathbb{R}$, the associated *Markov semigroup* is defined by

$$(P_t f)(x) := \mathbb{E}[f(X_t^x)], \qquad t \geq 0,\ x \in \mathcal{H}. \tag{12.68}$$

The family $(P_t)_{t\geq 0}$ is a strongly continuous semigroup on suitable Banach spaces of functions on $\mathcal{H}$ (typically $C_b(\mathcal{H})$, $B_b(\mathcal{H})$, or $L^p(\mu)$ spaces with respect to an invariant measure μ), and its infinitesimal generator $\mathcal{L}$ is defined formally by

$$\mathcal{L}f(x) := \lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t}, \tag{12.69}$$

with domain $\mathcal{D}(\mathcal{L})$ consisting of all functions f for which the limit exists (in the appropriate topology).

Assuming f is twice continuously Fréchet differentiable with bounded derivatives, one can derive the formal expression of $\mathcal{L}$ via Itô's formula. Let $Df(x) \in \mathcal{L}(\mathcal{H}; \mathbb{R}) \simeq \mathcal{H}$ denote the first Fréchet derivative (gradient) and $D^2 f(x) \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ the second derivative (Hessian). Applying Itô's formula in the Hilbert space setting yields

$$df(X_t^x) = \langle Df(X_t^x), AX_t^x + F(X_t^x)\rangle_{\mathcal{H}}\,dt + \frac{1}{2}\mathrm{Tr}\big(GG^* D^2 f(X_t^x)\big)\,dt + \langle Df(X_t^x), G\,dW_t\rangle_{\mathcal{H}}. \tag{12.70}$$

Taking expectations and dividing by t, then letting $t \downarrow 0$, one obtains the expression for the generator:

$$\boxed{\mathcal{L}f(x) = \langle Ax + F(x), Df(x)\rangle_{\mathcal{H}} + \frac{1}{2}\mathrm{Tr}\big(GG^* D^2 f(x)\big)}. \tag{12.71}$$

This operator is the infinite-dimensional analogue of the Kolmogorov backward operator for finite-dimensional diffusion processes.

Domain and Core. A natural choice for the core of $\mathcal{L}$ is the algebra $\mathcal{F}C_b^2(\mathcal{H})$ of smooth cylindrical functions of the form

$$f(x) = \psi(\langle x, h_1\rangle, \dots, \langle x, h_n\rangle), \tag{12.72}$$

where $\psi \in C_b^2(\mathbb{R}^n)$ and $h_1, \dots, h_n \in \mathcal{H}$. For such f, the derivatives are given by

$$Df(x) = \sum_{i=1}^{n} \partial_i \psi(\langle x, h_1\rangle, \dots, \langle x, h_n\rangle) h_i, \tag{12.73}$$

$$D^2 f(x) = \sum_{i,j=1}^{n} \partial_{ij}^2 \psi(\langle x, h_1\rangle, \dots, \langle x, h_n\rangle)\, h_i \otimes h_j, \tag{12.74}$$

and thus $\mathcal{L}$ acts as a second-order differential operator on a finite-dimensional projection of $\mathcal{H}$. This definition extends to more general function spaces via closure.

Functional Analytic Characterization. The operator $\mathcal{L}$ is typically *dissipative* in $L^2(\mu)$ (for invariant measure μ) if

$$\int_{\mathcal{H}} (\mathcal{L}f)\, f \, d\mu \le 0, \qquad \forall f \in \mathcal{D}(\mathcal{L}), \tag{12.75}$$

and *maximally dissipative* when the range condition $(\lambda I - \mathcal{L})(\mathcal{D}(\mathcal{L})) = L^2(\mu)$ holds for some $\lambda > 0$. In such case, $\mathcal{L}$ generates a contraction semigroup on $L^2(\mu)$ by the Hille–Yosida theorem.

When A is self-adjoint, negative definite, and G non-degenerate, the Ornstein–Uhlenbeck operator

$$\mathcal{L}_0 f(x) = \langle Ax, Df(x)\rangle + \frac{1}{2}\mathrm{Tr}\big(GG^* D^2 f(x)\big) \tag{12.76}$$

is itself self-adjoint in $L^2(\mu)$, where μ is the invariant Gaussian measure with covariance operator $Q_\infty = \int_0^\infty S_t G G^* S_t^* \, dt$. In such Gaussian setting, $\mathcal{L}_0$ is related to the Gross–Ornstein–Uhlenbeck operator, generating a strongly continuous, analytic, and hypercontractive semigroup.

For more general nonlinear drifts F, $\mathcal{L}$ is not self-adjoint but typically *sectorial* in $L^2(\mu)$: there exists $\theta < \pi/2$ such that

$$\| (zI - \mathcal{L})^{-1} \|_{\mathcal{L}(L^2(\mu))} \le \frac{C}{|z|}, \qquad z \in \mathbb{C} \setminus \Sigma_\theta, \tag{12.77}$$

where Σ_θ is an open sector centered on the positive real axis. Sectoriality implies that $\mathcal{L}$ generates an analytic semigroup, enabling fractional powers and functional calculus.

Spectral Structure. The spectrum $\sigma(\mathcal{L})$ determines the long-time behavior of (P_t). For instance:

- If $\mathcal{L}$ has purely discrete spectrum $\{\lambda_k\}$ with $0 = \lambda_0 > \mathrm{Re}\,\lambda_1 \ge \mathrm{Re}\,\lambda_2 \ge \cdots$, then (P_t) admits an eigenfunction expansion:

$$P_t f = \sum_{k=0}^{\infty} e^{\lambda_k t} \langle f, \phi_k\rangle_\mu \, \phi_k.$$

- The *spectral gap* $\gamma = -\sup\{\operatorname{Re}\lambda_k : \lambda_k \neq 0\}$ quantifies the exponential rate of convergence to equilibrium.
- If $\mathcal{L}$ is self-adjoint in $L^2(\mu)$, then $\sigma(\mathcal{L}) \subset (-\infty, 0]$, and $\mathcal{L}$ admits a spectral decomposition via a resolution of the identity:

$$\mathcal{L} = \int_{(-\infty,0]} \lambda \, dE(\lambda),$$

with $P_t = \int_{(-\infty,0]} e^{\lambda t} \, dE(\lambda)$.

These analytic and spectral properties play a central role in the study of regularity, ergodicity, and stability of infinite-dimensional diffusions. In particular, coercivity and sectoriality of $\mathcal{L}$ underpin the construction of invariant measures, while its spectral gap governs mixing rates and functional inequalities (such as Poincaré and logarithmic Sobolev inequalities) in infinite-dimensional probability spaces.

12.9 Connection to Dirichlet Forms

In the study of infinite-dimensional stochastic processes, particularly those arising from stochastic partial differential equations (SPDEs), Dirichlet forms play a central role in establishing the link between probabilistic and analytic structures. Let μ be a probability measure on a separable Hilbert space $\mathcal{H}$ which is invariant under a Markov semigroup $(P_t)_{t\geq 0}$ associated with an SPDE. We denote by $L^2(\mathcal{H}, \mu)$ the corresponding Hilbert space of square-integrable functions.

Definition of the Dirichlet Form

For smooth cylindrical functions $f, g : \mathcal{H} \to \mathbb{R}$, the associated *Dirichlet form* is defined as

$$\mathcal{E}(f, g) = \int_{\mathcal{H}} \langle Df(x), QDg(x) \rangle_{\mathcal{H}} \, \mu(dx), \tag{12.78}$$

where D denotes the Fréchet derivative, and Q is a positive, self-adjoint, trace-class operator acting as the covariance operator of the Gaussian measure μ. The quadratic form

$$\mathcal{E}(f, f) = \int_{\mathcal{H}} \|Q^{1/2} Df(x)\|_{\mathcal{H}}^2 \, \mu(dx) \tag{12.79}$$

encodes the "energy" associated with the function f and defines a nonnegative symmetric bilinear form on a dense subspace of $L^2(\mathcal{H}, \mu)$.

Closability and Associated Operator

The form $(\mathcal{E}, \mathcal{F})$, where $\mathcal{F}$ is the closure of smooth cylindrical functions under the norm

$$\|f\|_{\mathcal{E}}^2 = \|f\|_{L^2(\mu)}^2 + \mathcal{E}(f, f), \tag{12.80}$$

is *closable* in $L^2(\mathcal{H}, \mu)$, and its closure defines a symmetric, strongly local, regular Dirichlet form. By the general theory of Dirichlet forms (see Fukushima–Oshima–Takeda), there exists a unique non-positive self-adjoint operator $(\mathcal{L}, D(\mathcal{L}))$ on $L^2(\mathcal{H}, \mu)$ such that

$$\mathcal{E}(f,g) = -\int_{\mathcal{H}} (\mathcal{L}f)(x)\, g(x)\, \mu(dx), \qquad f \in D(\mathcal{L}),\, g \in \mathcal{F}. \tag{12.81}$$

This operator $\mathcal{L}$ coincides with the infinitesimal generator of the Markov semigroup (P_t) associated with the SPDE.

Example: Ornstein–Uhlenbeck Process

For the Ornstein–Uhlenbeck process on $\mathcal{H}$ given by

$$dX_t = AX_t\, dt + G\, dW_t, \tag{12.82}$$

with invariant Gaussian measure $\mu = \mathcal{N}(0, Q_\infty)$ satisfying

$$AQ_\infty + Q_\infty A^* + GG^* = 0, \tag{12.83}$$

the corresponding Dirichlet form takes the explicit form

$$\mathcal{E}(f,g) = \int_{\mathcal{H}} \langle Q_\infty^{1/2} Df(x), Q_\infty^{1/2} Dg(x)\rangle_{\mathcal{H}}\, \mu(dx). \tag{12.84}$$

The generator $\mathcal{L}$ associated with this form is the Ornstein–Uhlenbeck operator

$$\mathcal{L}f(x) = \frac{1}{2}\mathrm{Tr}\Big(Q_\infty D^2 f(x)\Big) + \langle Ax, Df(x)\rangle_{\mathcal{H}}, \tag{12.85}$$

which acts as an infinite-dimensional analogue of the classical elliptic operator associated with diffusion.

Functional Analytic and Probabilistic Connection

The Dirichlet form framework provides a bridge between:

1. **Functional Analysis:** where $\mathcal{E}$ defines a symmetric bilinear form generating a self-adjoint operator;
2. **Potential Theory:** where $\mathcal{E}$ characterizes notions of capacity, energy, and equilibrium;
3. **Stochastic Analysis:** where $\mathcal{E}$ corresponds to the quadratic variation and energy dissipation of stochastic dynamics.

In particular, the Markov semigroup (P_t) associated with the SPDE satisfies the *energy identity*

$$\frac{d}{dt}\|P_t f\|^2_{L^2(\mu)} = -2\,\mathcal{E}(P_t f, P_t f), \tag{12.86}$$

which reveals that $\mathcal{E}$ governs the dissipation of energy in the evolution of expectations.

Conclusion

Thus, under a Gaussian invariant measure μ, the Dirichlet form

$$\mathcal{E}(f,g) = \int_{\mathcal{H}} \langle Df(x), QDg(x)\rangle_{\mathcal{H}}\, \mu(dx) \tag{12.87}$$

is symmetric, closable, and generates a self-adjoint operator that coincides with the infinitesimal generator of the infinite-dimensional diffusion. This unifies the analytic, geometric, and stochastic aspects of SPDEs and provides the foundation for potential theory and ergodicity in infinite-dimensional spaces.

12.10 Conclusion

Stochastic partial differential equations extend the realm of diffusion to infinite-dimensional configuration spaces. Their study intertwines semigroup theory, stochastic calculus, spectral analysis, and infinite-dimensional functional analysis. Through the unifying framework of Dirichlet forms and Malliavin calculus, SPDEs reveal the profound interplay between randomness, geometry, and analysis — providing the mathematical language for noise-driven dynamics in continua and fields.

References: Ghosh (2025) [32], Da Prato and Zabczyk (1992) [10], Röckner et al. (2014) [38].

Chapter 13

Stochastic Analysis on Manifolds and Geometric Diffusions

"Geometry is the language with which probability describes motion."

— Anonymous

13.1 Brownian Motion on Riemannian Manifolds

Let (M, g) be a smooth, connected, d-dimensional Riemannian manifold with metric tensor $g = (g_{ij})$, inverse g^{ij}, and Riemannian volume form $d\mathrm{vol}_g = \sqrt{|g|}\, dx$ where $|g| = \det(g_{ij})$. The *Laplace–Beltrami operator* Δ_g acting on $f \in C^\infty(M)$ is the divergence of the gradient,

$$\Delta_g f \;=\; \mathrm{div}_g(\nabla f) \;=\; \frac{1}{\sqrt{|g|}}\partial_i\big(\sqrt{|g|}\, g^{ij}\partial_j f\big), \tag{13.1}$$

with the Einstein summation convention in place. The operator $\frac{1}{2}\Delta_g$ is a (symmetric, nonpositive) second-order elliptic differential operator which serves as the infinitesimal generator of Brownian motion on (M, g).

A stochastic process $(X_t)_{t\geq 0}$ with continuous paths and values in M is called a *Brownian motion on* (M, g) (started at $x \in M$) if for every $f \in C_c^\infty(M)$ the process

$$M_t^f \;=\; f(X_t) - f(X_0) - \tfrac{1}{2}\int_0^t (\Delta_g f)(X_s)\, ds \tag{13.2}$$

is a real-valued martingale with respect to the natural filtration. Equivalently, the transition semigroup

$$(P_t f)(x) := \mathbb{E}^x[f(X_t)] \tag{13.3}$$

is the heat semigroup solving the Cauchy problem

$$\partial_t u(t,x) = \tfrac{1}{2}\Delta_g u(t,x), \qquad u(0,\cdot) = f, \tag{13.4}$$

so that $P_t = e^{\frac{t}{2}\Delta_g}$ on $L^2(M, d\mathrm{vol}_g)$ (with appropriate domain considerations).

Local SDE representation. In local coordinates $(x^1, \ldots, x^d)$ an orthonormal frame $\{e_a(x)\}_{a=1}^d$ (for instance obtained by orthonormalizing coordinate vector fields) gives vector fields $e_a = e_a^i(x)\partial_{x^i}$ satisfying $g(e_a, e_b) = \delta_{ab}$. If $(B_t^a)_{a=1}^d$ are independent real Brownian motions, the Stratonovich stochastic differential equation

$$dX_t \;=\; e_a(X_t) \circ dB_t^a \tag{13.5}$$

defines (up to explosion) a diffusion whose generator equals $\frac{1}{2}\sum_{a=1}^d L_{e_a}^2 \;=\; \frac{1}{2}\Delta_g$, where L_{e_a} denotes Lie differentiation along e_a. Writing components in local coordinates, the Itô form of the same equation is

$$dX_t^i \;=\; e_a^i(X_t)\, dB_t^a \;+\; \tfrac{1}{2}\, g^{jk}(X_t)\Gamma_{jk}^i(X_t)\, dt, \tag{13.6}$$

where Γ_{jk}^i are the Christoffel symbols of the Levi–Civita connection. The drift term $\frac{1}{2}g^{jk}\Gamma_{jk}^i$ is the Itô correction ensuring that the generator is exactly $\frac{1}{2}\Delta_g$.

Existence, uniqueness and stochastic development. Existence and uniqueness (in law) of Brownian motion on a complete Riemannian manifold follow from the construction on the orthonormal frame bundle $\mathrm{O}(M)$. Let $\pi : \mathrm{O}(M) \to M$ be the orthonormal frame bundle and H_a the horizontal vector fields corresponding to the canonical horizontal distribution determined by the Levi–Civita connection. Solving the horizontal Stratonovich SDE on $\mathrm{O}(M)$

$$dU_t \;=\; H_a(U_t) \circ dB_t^a, \tag{13.7}$$

with U_0 an orthonormal frame at x, and projecting $X_t = \pi(U_t)$ yields Brownian motion on M. This *stochastic development* maps Euclidean Brownian motion to manifold Brownian motion and provides a powerful geometric construction which also yields the stochastic parallel transport along paths via the frame process U_t.

Non-explosion and stochastic completeness. A Riemannian manifold (M, g) is called *stochastically complete* if Brownian motion started at any $x \in M$ is non-explosive (i.e. has infinite lifetime a.s.). Geometric sufficient conditions for stochastic completeness include: geodesic completeness with Ricci curvature bounded below, or the Grigor'yan volume growth condition. Analytically, stochastic completeness is equivalent to preservation of mass by the heat semigroup: $P_t 1 = 1$ for all $t \geq 0$.

Heat kernel and short-time asymptotics. Under mild hypotheses (e.g. completeness and smoothness) the heat semigroup admits a smooth positive heat kernel $p_t(x, y)$ with respect to the Riemannian volume, so that

$$(P_t f)(x) = \int_M p_t(x, y) f(y)\, d\mathrm{vol}_g(y). \tag{13.8}$$

The heat kernel is jointly C^∞ on $(0, \infty) \times M \times M$ and satisfies the Chapman–Kolmogorov identity. The celebrated Minakshisundaram–Pleijel (or heat kernel) expansion gives as $t \downarrow 0$

$$p_t(x, y) \sim (4\pi t)^{-d/2} e^{-\frac{\mathrm{dist}(x,y)^2}{4t}} \sum_{k=0}^{\infty} a_k(x, y)\, t^k, \tag{13.9}$$

and Varadhan's short-time principle reads

$$\lim_{t\downarrow 0} -2t\log p_t(x,y) = \mathrm{dist}(x,y)^2, \tag{13.10}$$

uniformly on compact sets, where $\mathrm{dist}(\cdot,\cdot)$ is the Riemannian distance.

Geometry and analytic estimates. Curvature controls analytic and probabilistic behaviour. If the Ricci tensor satisfies $\mathrm{Ric} \geq Kg$ for some $K \in \mathbb{R}$, then Bakry–Émery gradient estimates hold:

$$|\nabla P_t f|^2 \leq e^{-2Kt}\, P_t(|\nabla f|^2), \qquad f \in C_b^\infty(M), \tag{13.11}$$

and corresponding functional inequalities follow (Poincaré, log-Sobolev). Coupling constructions using parallel transport along geodesics yield Wasserstein contraction estimates:

$$W_2(\delta_x P_t, \delta_y P_t) \leq e^{-Kt}\, \mathrm{dist}(x,y), \tag{13.12}$$

whenever $\mathrm{Ric} \geq K$.

Stochastic parallel transport and connection Laplacian. Let $//_t : T_{X_0}M \to T_{X_t}M$ denote stochastic parallel transport along X defined by the horizontal frame process U_t. The covariant derivative along X obeys the Stratonovich equation

$$\nabla_{\circ dX_t} //_t = 0, \tag{13.13}$$

and if $V_t = //_t v$ for fixed $v \in T_{X_0}M$, then V_t is the parallel translate of v. The connection Laplacian acting on sections (the rough Laplacian) appears in Bochner-type identities and in the analysis of vector-valued heat flows:

$$\Box\, \Phi = \nabla^*\nabla\Phi, \tag{13.14}$$

and is related to the generator of the lifted diffusion on vector bundles.

Smoothness of transition probabilities and Malliavin nondegeneracy. Since the generator $\frac{1}{2}\Delta_g$ is elliptic (the vector fields e_a span the tangent space), Hörmander's condition is satisfied and finite-dimensional Malliavin theory implies that $p_t(x,y)$ is C^∞ in (x,y) for $t > 0$. The Malliavin covariance matrix is invertible for $t > 0$, which yields Gaussian-type upper and lower bounds on $p_t(x,y)$ under curvature and injectivity radius control.

Long-time behaviour and spectral theory. On complete manifolds with suitable confinement (e.g. compact M or potentials growing at infinity) the spectrum of $-\frac{1}{2}\Delta_g$ is discrete and the heat semigroup admits an eigenexpansion. On noncompact manifolds the spectral measure, existence of a spectral gap, and asymptotic behaviour of P_t are governed by geometry (volume growth, curvature decay, ends structure).

Applications and further directions. Brownian motion on manifolds connects probabilistic representations of geometric PDEs (e.g. probabilistic proofs of the Hodge theorem, index theorems), geometric analysis (comparison theorems, heat kernel estimates), and stochastic differential geometry (stochastic flows, stochastic development). Extensions include analysis on manifolds with boundary (reflecting Brownian motion and boundary local times), on non-Riemannian settings (sub-Riemannian Brownian motion), and interactions with gauge fields and stochastic parallel transport in vector bundles.

In summary, Brownian motion on (M, g) is the diffusion process generated by $\frac{1}{2}\Delta_g$; it admits both intrinsic geometric constructions via the orthonormal frame bundle and analytical manifestations through the heat semigroup and heat kernel, with its qualitative and quantitative behavior tightly controlled by the Riemannian geometry of M.

13.2 Geometric Interpretation via Horizontal Lifts

Let (M, g) be a smooth, connected Riemannian manifold of dimension d, and let $O(M)$ denote its *orthonormal frame bundle*, i.e.

$$O(M) := \bigcup_{x\in M} O_x(M), \qquad O_x(M) := \{u : \mathbb{R}^d \to T_xM \text{ is an isometry w.r.t. } g_x\}. \tag{13.15}$$

The projection map $\pi : O(M) \to M$ sends an orthonormal frame u to its base point $\pi(u) = x$. The tangent bundle $TO(M)$ admits a canonical decomposition into a *horizontal* and a *vertical* subbundle:

$$TO(M) = \mathcal{H} \oplus \mathcal{V}, \tag{13.16}$$

where the horizontal distribution $\mathcal{H}$ is defined via the Levi–Civita connection ∇ on (M, g). The horizontal lift of a tangent vector $v \in T_xM$ at $u \in O_x(M)$ is the unique vector $h_u(v) \in \mathcal{H}_u$ satisfying

$$\pi_* h_u(v) = v, \qquad \omega(h_u(v)) = 0, \tag{13.17}$$

where ω denotes the connection 1-form on $O(M)$ induced by ∇.

Canonical horizontal vector fields. Fix the canonical basis $(e_1, \ldots, e_d)$ of $\mathbb{R}^d$. For each $a = 1, \ldots, d$, define the *canonical horizontal vector field* H_a on $O(M)$ by

$$H_a(u) := h_u(ue_a), \tag{13.18}$$

i.e. $H_a(u)$ is the horizontal lift of the tangent vector $ue_a \in T_{\pi(u)}M$. These vector fields form an orthonormal frame of the horizontal distribution $\mathcal{H} \subset TO(M)$.

Stochastic development on $O(M)$. Let $(B_t^1, \ldots, B_t^d)$ be standard independent Brownian motions in $\mathbb{R}^d$. The *horizontal lift* of Brownian motion on M is the process $(U_t)_{t\geq 0}$ in $O(M)$ satisfying the Stratonovich stochastic differential equation

$$dU_t \;=\; \sum_{a=1}^{d} H_a(U_t) \circ dB_t^a, \qquad U_0 = u_0 \in O_{x_0}(M), \tag{13.19}$$

where u_0 is an orthonormal frame at the starting point $x_0 \in M$. This SDE is geometrically well-defined because the vector fields H_a are globally smooth and horizontal, and the Stratonovich formulation respects the manifold structure of $O(M)$.

Projection to Brownian motion on M. Define the projected process $X_t = \pi(U_t)$. Applying Itô's formula in local coordinates and using the properties of the horizontal vector fields, one verifies that

$$\mathcal{L}(f \circ \pi)(u) = \frac{1}{2} \sum_{a=1}^{d} H_a^2 (f \circ \pi)(u) = \frac{1}{2}(\Delta_g f)(\pi(u)), \tag{13.20}$$

so that the generator of the process X_t is precisely $\frac{1}{2}\Delta_g$, the Laplace–Beltrami operator. Hence X_t is Brownian motion on the Riemannian manifold (M, g), and the law of X_t is independent of the choice of initial frame u_0 above x_0.

Geometric interpretation. The equation (13.19) defines a *stochastic development* of Euclidean Brownian motion (B_t) onto the manifold (M, g) via the connection-induced horizontal lift. In other words, the Euclidean noise is "rolled" along M without slipping or twisting. The projection $X_t = \pi(U_t)$ depends solely on the Riemannian structure (M, g) and not on any choice of coordinates.

This construction highlights the intrinsic nature of Brownian motion on Riemannian manifolds: it arises from the geometry of (M, g) alone, as the projection of a horizontal diffusion on the orthonormal frame bundle governed by the Levi–Civita connection.

13.3 Heat Semigroup and Heat Kernel on Manifolds

Let (M, g) be a smooth, connected, complete Riemannian manifold of dimension d, with associated volume measure $d\mathrm{vol}_g$. The Laplace–Beltrami operator Δ_g acts on smooth functions $f \in C^\infty(M)$ as

$$\Delta_g f = \mathrm{div}_g(\nabla_g f), \tag{13.21}$$

where $\nabla_g f$ denotes the Riemannian gradient and div_g the Riemannian divergence. We consider the infinitesimal generator

$$L := \frac{1}{2}\Delta_g, \tag{13.22}$$

which is essentially self-adjoint on $L^2(M, d\mathrm{vol}_g)$ under appropriate completeness assumptions (for instance, when M is complete without boundary).

Heat semigroup. The operator $L = \frac{1}{2}\Delta_g$ generates a strongly continuous, symmetric contraction semigroup $(T_t)_{t\geq 0}$ on $L^2(M)$, called the *heat semigroup*, satisfying:

$$T_t f = e^{tL} f, \qquad T_0 = I, \qquad T_{t+s} = T_t T_s, \tag{13.23}$$

and the limit property $\lim_{t\to 0^+} T_t f = f$ in $L^2(M)$ for all $f \in L^2(M)$.

If X_t denotes the Brownian motion on (M, g) starting at x, defined as the diffusion process with generator $\frac{1}{2}\Delta_g$, then for any bounded measurable f,

$$(T_t f)(x) = \mathbb{E}^x[f(X_t)]. \tag{13.24}$$

Thus, T_t describes the evolution of expectations of functionals along the Brownian motion on M.

Heat kernel representation. The heat semigroup admits an integral representation in terms of a smooth, nonnegative function $p_t(x, y)$, known as the *heat kernel*, such that

$$(T_t f)(x) = \int_M p_t(x, y) f(y) \, \mathrm{dvol}_g(y), \tag{13.25}$$

for all $f \in L^2(M)$ and $t > 0$. The heat kernel $p_t(x, y)$ is the *minimal fundamental solution* to the heat equation

$$\frac{\partial u}{\partial t}(t, x) = \tfrac{1}{2}\Delta_g u(t, x), \qquad u(0, x) = f(x), \tag{13.26}$$

in the sense that

$$u(t, x) = (T_t f)(x) = \int_M p_t(x, y) f(y) \, \mathrm{dvol}_g(y) \tag{13.27}$$

is the unique minimal nonnegative solution of (13.26) for every nonnegative initial data f.

Semigroup property of the heat kernel. The kernel $p_t(x, y)$ satisfies the *Chapman–Kolmogorov equation* or *semigroup property*:

$$p_{t+s}(x, y) = \int_M p_t(x, z) \, p_s(z, y) \, \mathrm{dvol}_g(z), \qquad \forall\, t, s > 0. \tag{13.28}$$

This follows directly from the semigroup identity $T_{t+s} = T_t T_s$ and the integral representation (13.25). Furthermore, for each fixed $t > 0$, $p_t(x, \cdot)$ is a smooth function on M and symmetric:

$$p_t(x, y) = p_t(y, x), \tag{13.29}$$

reflecting the self-adjointness of $\frac{1}{2}\Delta_g$.

Asymptotic expansion as $t \to 0^+$. A cornerstone of geometric analysis is the short-time asymptotic expansion of $p_t(x, y)$, which reveals the influence of the manifold's curvature on diffusion. In a geodesically convex neighborhood where the geodesic distance $d_g(x, y)$ is smooth, one has

$$p_t(x, y) \;\sim\; (4\pi t)^{-d/2} \, e^{-\frac{d_g(x,y)^2}{4t}} \sum_{k=0}^{\infty} a_k(x, y) \, t^k, \qquad t \to 0^+, \tag{13.30}$$

where:

- $d_g(x, y)$ is the Riemannian geodesic distance between x and y,
- $a_k(x, y)$ are smooth coefficient functions on $M \times M$,
- $a_0(x, x) = 1$, and higher coefficients $a_k(x, y)$ encode curvature information.

For instance, near the diagonal $x = y$,

$$a_1(x, x) = \tfrac{1}{6} R(x), \tag{13.31}$$

where $R(x)$ is the scalar curvature of (M, g).

This expansion generalizes the Euclidean heat kernel

$$p_t^{\mathbb{R}^d}(x,y) = (4\pi t)^{-d/2} \exp\left(-\frac{|x-y|^2}{4t}\right), \tag{13.32}$$

by replacing the Euclidean distance with the geodesic distance and incorporating geometric corrections through curvature-dependent coefficients $a_k(x,y)$.

Geometric significance. The heat kernel $p_t(x,y)$ encodes rich geometric and analytic information about the manifold. It determines the spectral data of the Laplace–Beltrami operator, since

$$p_t(x,y) = \sum_{j=0}^{\infty} e^{-\lambda_j t}\, \phi_j(x)\phi_j(y), \tag{13.33}$$

where $\{(\lambda_j, \phi_j)\}$ are the eigenpairs of $-\frac{1}{2}\Delta_g$. Moreover, heat kernel asymptotics provide deep connections between local geometry (curvature invariants) and global spectral properties (Weyl's law, index theorems, and spectral invariants).

Hence, the heat semigroup T_t and its kernel $p_t(x,y)$ serve as the analytic bridge between stochastic diffusion, differential geometry, and spectral theory on Riemannian manifolds.

13.4 Gradient Operators and Stochastic Parallel Transport

Let (M,g) be a smooth Riemannian manifold, ∇ its Levi–Civita connection, and $(X_t)_{t\geq 0}$ a continuous semimartingale on M (in particular a Brownian motion when its generator is $\frac{1}{2}\Delta_g$). We expand with full precision the geometric notions of gradient, covariant derivative along a semimartingale, stochastic parallel transport, and the manifold Itô formula which together furnish the analytic backbone for Malliavin calculus on manifolds.

Covariant derivative along a curve / semimartingale. If $\gamma : [0,T] \to M$ is a C^1 curve and $V(t)$ is a time-dependent vector field along γ (i.e. $V(t) \in T_{\gamma(t)}M$), the covariant derivative of V along γ is the vector field $\nabla_{\dot\gamma(t)} V \in T_{\gamma(t)}M$ defined by the Levi–Civita connection. For a continuous semimartingale X_t and an adapted process V_t with values in $T_{X_t}M$ one defines the *stochastic covariant derivative* in Stratonovich form by the pathwise limit

$$\nabla_{\circ dX_t} V_t \ := \ \text{Stratonovich differential of } V_t \text{ along } X_t, \tag{13.34}$$

characterized locally (in coordinates) by replacing ordinary derivatives by covariant derivatives and ordinary differentials by Stratonovich differentials. Concretely, if $V_t = V^i(t,X_t)\partial_{x^i}$ in local coordinates, then

$$\nabla_{\circ dX_t} V_t = \Big(dV^i(t,X_t) + \Gamma^i_{jk}(X_t) V^j(t,X_t) \circ dX_t^k\Big)\partial_{x^i}, \tag{13.35}$$

where Γ^i_{jk} are the Christoffel symbols of ∇.

Stochastic parallel transport. Let $v \in T_{X_0}M$ be a fixed initial vector. The *stochastic parallel transport* $P_t : T_{X_0}M \to T_{X_t}M$ along the semimartingale X_t is defined as the unique family of linear isometries satisfying the covariant Stratonovich ordinary differential equation

$$\nabla_{\circ dX_t}(P_t v) \;=\; 0, \qquad P_0 = \mathrm{Id}_{T_{X_0}M}. \tag{13.36}$$

Equivalently, $V_t := P_t v$ is the unique adapted $T_{X_t}M$-valued process such that $V_0 = v$ and the covariant derivative along X vanishes in Stratonovich sense. In local coordinates this SDE becomes

$$dV_t^i + \Gamma^i_{jk}(X_t)V_t^j \circ dX_t^k = 0, \tag{13.37}$$

which preserves the inner product: $g_{X_t}(P_t v, P_t w) = g_{X_0}(v,w)$ for all $v, w \in T_{X_0}M$.

Frame-bundle formulation. Let $\pi : O(M) \to M$ be the orthonormal frame bundle and H_a the canonical horizontal vector fields. If U_t solves the horizontal Stratonovich SDE on $O(M)$

$$dU_t = \sum_{a=1}^{d} H_a(U_t) \circ dB_t^a, \qquad U_0 = u_0, \tag{13.38}$$

then $X_t = \pi(U_t)$ is the stochastic development of the Euclidean Brownian motion $B_t = (B_t^1, \dots, B_t^d)$. The stochastic parallel transport P_t is then given by the linear map

$$P_t = U_t U_0^{-1} : T_{X_0}M \to T_{X_t}M, \tag{13.39}$$

(i.e. U_t carries the initial frame to the frame at time t). This frame-bundle picture is the canonical geometric construction used in proofs and in Malliavin calculus on manifolds.

Itô versus Stratonovich and the geometric Itô formula. For manifold-valued semimartingales the Stratonovich calculus is geometrically natural because Stratonovich differentials transform covariantly under changes of charts. Let $f \in C^\infty(M)$. The Stratonovich Itô formula gives

$$f(X_t) - f(X_0) = \int_0^t \langle \nabla f(X_s), \circ dX_s \rangle, \tag{13.40}$$

where ∇f is the Riemannian gradient, and $\langle \nabla f, \circ dX \rangle$ denotes the Stratonovich integral of the 1-form df against X. Converting to Itô differentials (for example when X is Brownian motion) one obtains the covariant Itô formula: if X is a Brownian motion with generator $\frac{1}{2}\Delta_g$, then almost surely

$$f(X_t) = f(X_0) + \int_0^t \langle \nabla f(X_s), dX_s \rangle + \tfrac{1}{2}\int_0^t (\Delta_g f)(X_s)\, ds, \tag{13.41}$$

where the first stochastic integral is an Itô integral (with the necessary Itô correction already encoded in the second term). Here the Laplace–Beltrami operator Δ_g appears as the trace of the second covariant derivative:

$$\Delta_g f = \mathrm{tr}\,(\nabla^2 f). \tag{13.42}$$

Covariant Itô formula for vector fields and one-forms. If V is a C^2 vector field on M and $V(X_t)$ denotes its evaluation along X_t, then the covariant Itô formula reads (in Itô form)

$$\nabla_{dX_t} V(X_t) = \nabla_{\,dX_t} V = \nabla_{\circ dX_t} V - \tfrac{1}{2}\,\mathrm{Ric}^\sharp(V(X_t))\, dt, \tag{13.43}$$

with appropriate curvature correction terms appearing when converting from Stratonovich to Itô; the appearance of Ricci-type terms depends on whether one considers horizontal lifts or the Itô development (for Brownian motion the precise correction yields the generator $\frac{1}{2}\Delta_g$). For one-forms α the covariant Itô formula gives the evolution of $\alpha(X_t)$ along the noise with the connection Laplacian acting in the drift.

Linearization and derivative flow. Let $\Phi_{s,t} : M \to M$ denote the stochastic flow map (when it exists) associated to the SDE defining X. The derivative $D\Phi_{0,t}(x) : T_xM \to T_{X_t}M$ satisfies a linear covariant stochastic differential equation; in parallel-transported coordinates one obtains the linear Stratonovich SDE

$$\nabla_{\circ dX_t} Y_t = \big(\nabla Y\big)(X_t) \circ dX_t, \tag{13.44}$$

and if X_t is the solution driven by frame noise $e_a \circ dB_t^a$, the linearization becomes

$$\nabla_{\circ dX_t} Y_t = \sum_{a=1}^{d} \nabla_{Y_t} e_a(X_t) \circ dB_t^a. \tag{13.45}$$

Writing the linearization in the frame bundle and projecting yields an SDE for $D\Phi_{0,t}$ whose solution can be expressed via stochastic parallel transport and iterated covariant derivatives of the driving vector fields. This derivative flow is the fundamental object in the Malliavin differentiation of X_t.

Malliavin derivative and representation via parallel transport. In the manifold setting Malliavin differentiation of the solution X_t to a stochastic differential equation driven by an $\mathbb{R}^d$-Brownian motion admits a geometric representation: for $0 \le s \le t$,

$$D_s X_t = P_t \circ P_s^{-1}\, \sigma(X_s)\, 1_{[0,t]}(s), \tag{13.46}$$

where $\sigma(X_s) : \mathbb{R}^d \to T_{X_s}M$ is the map sending the canonical Euclidean noise directions to the corresponding tangent vectors (for Brownian motion σ is realized by the frame fields e_a), and P_t, P_s are stochastic parallel transports along X. This formula is to be interpreted in local charts or via the frame bundle; it expresses the infinitesimal sensitivity of X_t to perturbations of the driving Brownian path at time s by transporting the perturbation forward along the path.

Consequently, the finite-dimensional Malliavin covariance of a projection $\ell(X_t)$ (with $\ell : T_{X_t}M \to \mathbb{R}^m$ linear) can be written in terms of parallel transports:

$$\sigma_{\ell(X_t)} = \int_0^t \big(\ell \circ P_t \circ P_s^{-1} \circ \sigma(X_s)\big)\, \big(\ell \circ P_t \circ P_s^{-1} \circ \sigma(X_s)\big)^T\, ds. \tag{13.47}$$

Nondegeneracy of this matrix (in the Bouleau–Hirsch sense) yields absolute continuity and smoothness of the law of $\ell(X_t)$.

Conclusion and role in Malliavin calculus. The geometric constructs above—covariant derivative, stochastic parallel transport, frame-bundle horizontal lift, and the covariant Itô formula—provide the necessary structure to formulate Malliavin calculus intrinsically on manifolds.

Stochastic parallel transport furnishes canonical identifications of tangent spaces along sample paths, enabling one to express Malliavin derivatives, covariance operators, and integration-by-parts formulae in a coordinate-free manner. These tools yield hypoellipticity criteria (Hörmander-type conditions carried out on vector fields on M), smoothness of transition densities, and precise gradient estimates for the heat semigroup on manifolds, all central to stochastic analysis in geometric settings.

13.5 Bochner Identity and Curvature Effects

Let (M, g) be a smooth, connected, d-dimensional Riemannian manifold equipped with its Levi–Civita connection ∇, Riemann curvature tensor R, and Ricci curvature tensor Ric. For $f \in C^\infty(M)$, we denote by ∇f its Riemannian gradient, $\nabla^2 f = \nabla(\nabla f)$ its covariant Hessian, and $\Delta_g f = \operatorname{div}(\nabla f)$ the Laplace–Beltrami operator. The Bochner–Weitzenböck identity rigorously describes the interplay between the second covariant derivatives of f, the Laplacian, and the Ricci curvature.

13.5.1 Preliminary definitions

Gradient and Hessian. For any smooth function $f \in C^\infty(M)$, the Riemannian gradient ∇f is the vector field defined by

$$g(\nabla f, X) = df(X), \quad \forall\, X \in \Gamma(TM), \tag{13.48}$$

and the covariant Hessian $\nabla^2 f$ is the symmetric (0,2)-tensor given by

$$\nabla^2 f(X, Y) = g(\nabla_X \nabla f, Y) = X(Y(f)) - (\nabla_X Y)(f), \tag{13.49}$$

for all vector fields X, Y. Its squared norm is

$$\|\nabla^2 f\|^2 = g^{ia} g^{jb}\, (\nabla_i \nabla_j f)(\nabla_a \nabla_b f), \tag{13.50}$$

independent of the choice of coordinates.

Laplacian and Divergence. The Laplace–Beltrami operator acts as

$$\Delta_g f = \operatorname{div}(\nabla f) = g^{ij} \nabla_i \nabla_j f = \frac{1}{\sqrt{|g|}} \partial_i \Big(\sqrt{|g|}\, g^{ij} \partial_j f \Big), \tag{13.51}$$

where $|g| = \det(g_{ij})$. It is self-adjoint on $L^2(M, \mathrm{vol}_g)$ and coincides with the generator of Brownian motion on M.

13.5.2 Derivation of the Bochner–Weitzenböck identity

Let $f \in C^\infty(M)$. We compute the Laplacian of the squared gradient norm $u = |\nabla f|^2 = g(\nabla f, \nabla f)$. Differentiating covariantly,

$$\nabla_i u = 2\, g(\nabla_i \nabla f, \nabla f) = 2\, (\nabla_i \nabla_j f)(\nabla^j f), \tag{13.52}$$

and hence

$$\Delta_g u = \nabla^i \nabla_i u = 2\,\nabla^i\big((\nabla_i \nabla_j f)(\nabla^j f)\big) = 2\,(\nabla^i \nabla_i \nabla_j f)(\nabla^j f) + 2\,(\nabla_i \nabla_j f)(\nabla^i \nabla^j f). \quad (13.53)$$

Now, the first term can be rewritten by the curvature commutation formula:

$$\nabla^i \nabla_i \nabla_j f = \nabla_j \nabla^i \nabla_i f - R_j^{\ k} \nabla_k f = \nabla_j(\Delta_g f) - \mathrm{Ric}_{jk} \nabla^k f, \quad (13.54)$$

where we used the Ricci identity

$$(\nabla_i \nabla_j - \nabla_j \nabla_i)\nabla_k f = R_{ijk}{}^{\ell} \nabla_\ell f \quad \text{and hence} \quad R_{jk} = R_{ijk}{}^{i}. \quad (13.55)$$

Substituting this back, we obtain

$$\frac{1}{2}\Delta_g |\nabla f|^2 = (\nabla_j f)(\nabla^j \Delta_g f) + (\nabla_i \nabla_j f)(\nabla^i \nabla^j f) + \mathrm{Ric}_{jk}(\nabla^j f)(\nabla^k f). \quad (13.56)$$

In invariant tensor notation, this reads

$$\frac{1}{2}\Delta_g |\nabla f|^2 = \langle \nabla f, \nabla(\Delta_g f)\rangle + \|\nabla^2 f\|^2 + \mathrm{Ric}(\nabla f, \nabla f), \quad (13.57)$$

which is the *Bochner–Weitzenböck identity*.

13.5.3 Interpretations and geometric consequences

Analytic interpretation. The term $\|\nabla^2 f\|^2$ measures the local convexity of f, while $\mathrm{Ric}(\nabla f, \nabla f)$ captures how curvature affects the diffusion of gradients along the manifold. Equation (13.57) thus relates second-order derivatives of f and the geometry of M.

Integration formula. Integrating (13.57) over M with respect to the Riemannian volume measure and using the self-adjointness of Δ_g, we obtain

$$\int_M \|\nabla^2 f\|^2 \, d\mathrm{vol}_g + \int_M \mathrm{Ric}(\nabla f, \nabla f)\, d\mathrm{vol}_g = \int_M (\Delta_g f)^2 \, d\mathrm{vol}_g. \quad (13.58)$$

This equality is a central tool in elliptic regularity and spectral geometry, expressing the L^2-norm of the Laplacian in terms of Hessian and curvature contributions.

Probabilistic interpretation. If (X_t) is Brownian motion on M, Itô's formula and (13.57) yield

$$d|\nabla f(X_t)|^2 = \Big(\Delta_g |\nabla f|^2\Big)(X_t)\, dt + \text{martingale term}, \quad (13.59)$$

so that curvature affects the expected growth of $|\nabla f(X_t)|^2$. When $\mathrm{Ric} \geq 0$, the expectation $\mathbb{E}[|\nabla f(X_t)|^2]$ is subharmonic, providing a stochastic comparison principle.

Bakry–Émery curvature criterion. Define the *carré du champ* operator $\Gamma(f) = |\nabla f|^2$ and its iterated form

$$\Gamma_2(f) = \tfrac{1}{2}\Big(\Delta_g \Gamma(f) - 2\Gamma(f, \Delta_g f)\Big), \quad (13.60)$$

where $\Gamma(f,g) = \langle \nabla f, \nabla g \rangle$. By the Bochner identity,

$$\Gamma_2(f) = \|\nabla^2 f\|^2 + \mathrm{Ric}(\nabla f, \nabla f). \tag{13.61}$$

A lower bound $\mathrm{Ric} \geq Kg$ is thus equivalent to the *Bakry–Émery curvature-dimension condition* $CD(K,\infty)$:

$$\Gamma_2(f) \geq K\,\Gamma(f), \tag{13.62}$$

which underpins gradient estimates, logarithmic Sobolev inequalities, and exponential convergence to equilibrium for diffusion semigroups.

Conclusion. The Bochner–Weitzenböck identity provides a precise differential relation between curvature and the behavior of gradients and Laplacians. It lies at the intersection of Riemannian geometry, stochastic analysis, and functional inequalities, serving as the foundational analytic identity behind many curvature-dimension and diffusion semigroup results.

13.6 Diffusions with Drift and Divergence Form Operators

Let (M,g) be a smooth, connected, complete Riemannian manifold of dimension d, and let $b \in \Gamma(TM)$ be a smooth vector field on M, referred to as the *drift field*. We define the second-order differential operator

$$Lf = \frac{1}{2}\Delta_g f + \langle b, \nabla f \rangle, \tag{13.63}$$

for all $f \in C^\infty(M)$, where Δ_g denotes the Laplace–Beltrami operator and ∇f is the Riemannian gradient of f. This operator L is called the *generator* of a diffusion process with drift b.

13.6.1 Stochastic Differential Representation

Let $\{B_t\}_{t\geq 0}$ denote a standard Brownian motion in $\mathbb{R}^d$, and let $\{e_a(x)\}_{a=1}^d$ be a local orthonormal frame field on M, i.e.

$$g(e_a(x), e_b(x)) = \delta_{ab}, \qquad \forall x \in M. \tag{13.64}$$

The diffusion process $\{X_t\}_{t\geq 0}$ generated by L satisfies the *Stratonovich stochastic differential equation* (SDE)

$$dX_t = b(X_t)\,dt + \sum_{a=1}^{d} e_a(X_t) \circ dB_t^a, \tag{13.65}$$

where the stochastic integral is taken in the Stratonovich sense to preserve the coordinate-invariance under changes of local charts. Equation (13.65) defines a stochastic flow on M whose infinitesimal generator coincides with L as defined in (13.63).

The equivalence between the SDE and the generator can be rigorously verified using Itô's formula on manifolds: for any $f \in C^\infty(M)$,

$$df(X_t) = \langle \nabla f(X_t), b(X_t) \rangle\,dt + \frac{1}{2}(\Delta_g f)(X_t)\,dt + \sum_{a=1}^{d} \langle \nabla f(X_t), e_a(X_t) \rangle \circ dB_t^a. \tag{13.66}$$

Taking expectations and differentiating under the integral sign gives

$$\frac{d}{dt}\mathbb{E}[f(X_t)] = \mathbb{E}[Lf(X_t)], \tag{13.67}$$

confirming that L is indeed the infinitesimal generator of the diffusion.

13.6.2 Invariant Measures and Reversibility

Let μ be a smooth measure on M absolutely continuous with respect to the Riemannian volume form $d\mathrm{vol}_g$, say

$$d\mu = \rho(x)\, d\mathrm{vol}_g(x), \tag{13.68}$$

where $\rho : M \to (0, \infty)$ is a smooth density function. We say that μ is *invariant* under the diffusion generated by L if, for all $f, g \in C_c^\infty(M)$,

$$\int_M (Lf)\, g\, d\mu = \int_M f\, (L^* g)\, d\mu, \tag{13.69}$$

where L^* is the formal adjoint of L in $L^2(M, \mu)$. A sufficient condition for invariance is the *divergence-free condition* in weighted form:

$$\mathrm{div}_\mu(b) = \mathrm{div}(b) + \langle b, \nabla \log \rho \rangle = 0, \tag{13.70}$$

where div denotes the Riemannian divergence.

13.6.3 Gradient Drift and Self-Adjoint Realization

An important special case arises when the drift field is a gradient, i.e.

$$b = \nabla V \tag{13.71}$$

for some smooth potential function $V : M \to \mathbb{R}$. Then the operator (13.63) becomes

$$Lf = \frac{1}{2}\Delta_g f + \langle \nabla V, \nabla f \rangle. \tag{13.72}$$

We define the weighted measure

$$d\mu_V = e^{-2V(x)}\, d\mathrm{vol}_g(x). \tag{13.73}$$

A direct computation using the divergence theorem on Riemannian manifolds yields

$$\int_M (Lf)\, g\, d\mu_V = \frac{1}{2}\int_M \langle \nabla f, \nabla g \rangle\, d\mu_V = \int_M f\, (Lg)\, d\mu_V,$$

for all compactly supported smooth functions $f, g \in C_c^\infty(M)$. Hence L is symmetric (and closable) in $L^2(M, \mu_V)$.

The operator can equivalently be expressed in *divergence form*:

$$Lf = \frac{1}{2}e^{2V}\, \mathrm{div}\big(e^{-2V}\nabla f\big). \tag{13.74}$$

This expression emphasizes the geometric structure of the diffusion: it is a self-adjoint elliptic operator on the weighted Riemannian space (M, g, μ_V), known as a *weighted Laplacian* or *Witten Laplacian*.

13.6.4 Dirichlet Form Representation

The corresponding Dirichlet form $(\mathcal{E}, \mathcal{F})$ associated with L on $L^2(M, \mu_V)$ is given by

$$\mathcal{E}(f, g) = \frac{1}{2} \int_M \langle \nabla f, \nabla g \rangle \, d\mu_V, \tag{13.75}$$

with domain $\mathcal{F} = H^1(M, \mu_V)$, the Sobolev space of square-integrable functions with square-integrable gradients. This form is symmetric, closed, and Markovian, implying that L is the unique self-adjoint operator associated with $\mathcal{E}$.

13.6.5 Summary of the Analytic and Probabilistic Correspondence

The analytic and probabilistic viewpoints are summarized as follows:

Concept	**Analytic Representation**	**Probabilistic Interpretation**
Generator	$L = \frac{1}{2}\Delta_g + \langle b, \nabla \rangle$	Drift–diffusion dynamics
Adjoint	$L^* = \frac{1}{2}\Delta_g - \mathrm{div}(b\cdot)$	Backward Kolmogorov operator
Invariant measure	$\mathrm{div}_\mu(b) = 0$	Stationary law of diffusion
Gradient drift	$b = \nabla V$	Reversible diffusion
Weighted Laplacian	$L = \frac{1}{2}e^{2V}\mathrm{div}(e^{-2V}\nabla f)$	Self-adjoint generator in $L^2(M, e^{-2V} d\mathrm{vol}_g)$

This framework provides a mathematically rigorous unification of Riemannian geometry, stochastic analysis, and functional analysis through the study of elliptic operators in divergence form and their associated diffusion processes.

13.7 Geometric Dirichlet Forms

Let (M, g) be a smooth, connected, complete Riemannian manifold of dimension d, equipped with the Riemannian metric tensor $g = (g_{ij})$ and associated volume measure $d\mathrm{vol}_g = \sqrt{|g|}\, dx^1 \cdots dx^d$, where $|g| = \det(g_{ij})$. We denote by ∇f the Riemannian gradient of a smooth function $f \in C^\infty(M)$ and by Δ_g the Laplace–Beltrami operator acting on scalar functions.

13.7.1 Definition and Domain of the Dirichlet Form

The *geometric Dirichlet form* associated with the Riemannian structure (M, g) is defined by

$$\mathcal{E}(f, g) = \frac{1}{2} \int_M \langle \nabla f, \nabla g \rangle_g \, d\mathrm{vol}_g, \tag{13.76}$$

for all $f, g \in C_c^\infty(M)$, where $\langle \cdot, \cdot \rangle_g$ denotes the inner product on the tangent bundle TM induced by g. The natural domain of $\mathcal{E}$ is the first-order Sobolev space

$$\mathcal{D}(\mathcal{E}) = H^1(M) = \{ f \in L^2(M, d\mathrm{vol}_g) \mid \nabla f \in L^2(TM, d\mathrm{vol}_g) \}, \tag{13.77}$$

equipped with the norm

$$\|f\|_{H^1(M)}^2 = \|f\|_{L^2(M)}^2 + \|\nabla f\|_{L^2(TM)}^2. \tag{13.78}$$

13.7.2 Fundamental Analytic Properties

(i) Symmetry. By definition of the Riemannian inner product, $\mathcal{E}(f,g) = \mathcal{E}(g,f)$ for all $f,g \in \mathcal{D}(\mathcal{E})$.

(ii) Positivity and Markovianity. For any $f \in \mathcal{D}(\mathcal{E})$, one has $\mathcal{E}(f,f) \geq 0$. Moreover, if $\eta : \mathbb{R} \to \mathbb{R}$ is a normal contraction (i.e. Lipschitz with $|\eta(r) - \eta(s)| \leq |r - s|$), then

$$\mathcal{E}(\eta \circ f, \eta \circ f) \leq \mathcal{E}(f,f), \tag{13.79}$$

which ensures that $\mathcal{E}$ is a *Dirichlet form* in the sense of Fukushima–Oshima–Takeda.

(iii) Closability. Since $C_c^\infty(M)$ is dense in $H^1(M)$ with respect to the H^1 norm, the bilinear form $\mathcal{E}$ is closable on $L^2(M, d\mathrm{vol}_g)$, and its closure has domain exactly $H^1(M)$.

(iv) Strong locality. If $f, g \in \mathcal{D}(\mathcal{E})$ have compact supports and g is constant on a neighborhood of $\mathrm{supp}(f)$, then

$$\mathcal{E}(f,g) = 0. \tag{13.80}$$

This *strong locality* property characterizes the geometric nature of the form—it depends only on the local behavior of the gradients.

13.7.3 Associated Self-Adjoint Operator

By the general theory of closed, symmetric, and strongly local Dirichlet forms (see Fukushima–Oshima–Takeda, *Dirichlet Forms and Symmetric Markov Processes*), there exists a unique nonpositive self-adjoint operator L on $L^2(M, d\mathrm{vol}_g)$ such that

$$\mathcal{E}(f,g) = -\langle Lf, g\rangle_{L^2(M)}, \qquad \forall f \in \mathcal{D}(L),\ g \in \mathcal{D}(\mathcal{E}). \tag{13.81}$$

For the geometric Dirichlet form (13.76), this operator L is precisely the Riemannian Laplacian:

$$L = \frac{1}{2}\Delta_g. \tag{13.82}$$

The domain $\mathcal{D}(L)$ consists of those $f \in H^1(M)$ for which $\Delta_g f \in L^2(M)$, i.e.

$$\mathcal{D}(L) = \{f \in H^1(M) : \Delta_g f \in L^2(M)\}. \tag{13.83}$$

13.7.4 Functional Analytic Representation

The spectral theorem for unbounded self-adjoint operators implies that the energy form (13.76) can be expressed in operator-theoretic terms as

$$\mathcal{E}(f,f) = \|(-L)^{1/2} f\|_{L^2(M)}^2, \tag{13.84}$$

for all $f \in \mathcal{D}(\mathcal{E})$. Here $(-L)^{1/2}$ denotes the square-root of the nonpositive self-adjoint operator $-L$ obtained through functional calculus. Equation (13.84) rigorously identifies the Dirichlet form as the quadratic form of the operator $\frac{1}{2}\Delta_g$.

13.7.5 Geometric and Probabilistic Interpretations

(a) Geometric aspect. The energy $\mathcal{E}(f, f)$ represents the total Dirichlet energy of f on (M, g):

$$\mathcal{E}(f, f) = \frac{1}{2} \int_M |\nabla f|_g^2 \, d\mathrm{vol}_g, \tag{13.85}$$

which measures the average rate of change of f with respect to the metric g.

(b) Probabilistic aspect. The Markov semigroup $\{T_t\}_{t \geq 0}$ associated with $L = \frac{1}{2}\Delta_g$ is given by

$$T_t f(x) = \mathbb{E}^x[f(X_t)], \tag{13.86}$$

where $\{X_t\}_{t \geq 0}$ is the Brownian motion on (M, g). Thus, $\mathcal{E}$, Δ_g, and Brownian motion are connected through the analytic–geometric–probabilistic correspondence:

$$\boxed{\text{Dirichlet form } \mathcal{E} \iff \text{Laplacian } \tfrac{1}{2}\Delta_g \iff \text{Brownian motion on } (M, g).} \tag{13.87}$$

13.7.6 Summary

The geometric Dirichlet form provides a unified framework that simultaneously encodes:

- the *analytic structure* of the Laplace–Beltrami operator as a self-adjoint elliptic operator on $L^2(M, d\mathrm{vol}_g)$,
- the *geometric structure* of the manifold through the Riemannian metric tensor g, and
- the *probabilistic structure* of Brownian motion as the Markov process associated with the form.

Hence, the equality

$$\mathcal{E}(f, f) = \|(-L)^{1/2} f\|_{L^2(M)}^2 \tag{13.88}$$

serves as the rigorous bridge between geometry, analysis, and stochastic processes on Riemannian manifolds.

13.8 Feynman–Kac Formula on Manifolds

Let (M, g) be a smooth, connected, complete Riemannian manifold of dimension d, with Riemannian metric tensor $g = (g_{ij})$ and associated volume measure $d\mathrm{vol}_g = \sqrt{|g|}\, dx^1 \cdots dx^d$. Denote by Δ_g the Laplace–Beltrami operator acting on $C^\infty(M)$, defined locally in coordinates by

$$\Delta_g f = \frac{1}{\sqrt{|g|}} \partial_i \left(\sqrt{|g|}\, g^{ij} \partial_j f \right), \tag{13.89}$$

and by $\{X_t\}_{t \geq 0}$ the Brownian motion on (M, g) generated by $\frac{1}{2}\Delta_g$.

13.8.1 The Schrödinger-type Operator

Let $V : M \to \mathbb{R}$ be a measurable potential function satisfying

$$V(x) \geq -C_V, \qquad \text{for some constant } C_V \geq 0, \tag{13.90}$$

so that V is bounded from below. Define the Schrödinger-type operator

$$H = -\frac{1}{2}\Delta_g + V, \tag{13.91}$$

as an unbounded self-adjoint operator on $L^2(M, d\mathrm{vol}_g)$ with domain

$$\mathcal{D}(H) = \{f \in H^2(M) \mid Vf \in L^2(M)\}. \tag{13.92}$$

The corresponding semigroup $\{e^{-tH}\}_{t\geq 0}$ is called the *Schrödinger semigroup*, acting on $L^2(M)$ by

$$u(t,x) = (e^{-tH}f)(x), \qquad u(0,x) = f(x). \tag{13.93}$$

Formally, $u(t,x)$ satisfies the *heat equation with potential*:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta_g u - V(x)u. \tag{13.94}$$

13.8.2 The Feynman–Kac Representation

The *Feynman–Kac formula* provides a stochastic representation of the semigroup e^{-tH} in terms of Brownian motion on (M, g).

Theorem 13.1 (Feynman–Kac Formula on a Riemannian Manifold)**.** *Let $f \in L^2(M, d\mathrm{vol}_g)$ and $V : M \to \mathbb{R}$ be a measurable potential bounded from below. Let $\{X_t\}_{t\geq 0}$ denote the Brownian motion on (M, g) starting at $x \in M$, with generator $\frac{1}{2}\Delta_g$. Then, for all $t > 0$ and almost every $x \in M$,*

$$(e^{-tH}f)(x) = \mathbb{E}^x\left[\exp\left(-\int_0^t V(X_s)\,ds\right) f(X_t)\right]. \tag{13.95}$$

Proof: Throughout the proof (M, g) is a smooth, connected, complete Riemannian manifold and $d\mathrm{vol}_g$ its Riemannian volume measure. We denote by $X = (X_t)_{t\geq 0}$ the Brownian motion on M (with generator $\frac{1}{2}\Delta_g$) constructed e.g. by stochastic development on the orthonormal frame bundle; its law $\mathbb{P}^x$ when started at $x \in M$ satisfies the usual strong Markov and martingale properties. Let $V : M \to \mathbb{R}$ be measurable and bounded from below: there exists $C \geq 0$ such that $V(x) \geq -C$ for all $x \in M$. Define the Schrödinger-type operator

$$H = -\tfrac{1}{2}\Delta_g + V \tag{13.96}$$

as the self-adjoint operator obtained from the closed quadratic form

$$\mathcal{E}^V(f,f) = \tfrac{1}{2}\int_M |\nabla f|_g^2\, d\mathrm{vol}_g + \int_M V f^2\, d\mathrm{vol}_g, \qquad f \in H^1(M) \cap L^2(M, |V|\, d\mathrm{vol}_g), \tag{13.97}$$

via the Friedrichs construction (this is standard since V is lower bounded). The semigroup e^{-tH} is thus well-defined, positivity preserving and strongly continuous on $L^2(M)$.

We must prove that for every $f \in L^2(M)$ and every $t > 0$ the identity

$$(e^{-tH} f)(x) \;=\; \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big] \tag{13.98}$$

holds for almost every $x \in M$.

The proof proceeds in two stages: first treat the case of bounded (and measurable) potentials, then pass to the general lower-bounded case by a monotone approximation of the potential and form convergence.

Step 1: Bounded potential. Assume $V \in L^\infty(M)$. Fix $f \in C_c^\infty(M)$ (a core for the Laplace–Beltrami operator) and define for $t \geq 0$

$$u(t,x) := \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big]. \tag{13.99}$$

We first show that $u(t,\cdot)$ belongs to $L^2(M)$ for every t and that u satisfies the backward Cauchy problem associated to $\partial_t - (\frac{1}{2}\Delta_g - V)$.

Since V is bounded from below and above, $e^{-\int_0^t V(X_s)\,ds}$ is uniformly bounded by $e^{\|V^-\|_\infty t}$ and hence $u(t,x)$ is finite for all x and t. By dominated convergence and the Feller/Markov regularity of Brownian motion one checks $u(\cdot,\cdot)$ is continuous in (t,x) and $u(t,\cdot) \in C_b(M)$ for each $t > 0$.

Apply the manifold Itô formula to the Stratonovich SDE representation of Brownian motion (or equivalently to X_t and the smooth test function f), and observe that for $s \in [0,t]$ the process

$$M_s := e^{-\int_0^s V(X_r)\,dr} f(X_s) - f(X_0) - \int_0^s e^{-\int_0^r V(X_\rho)\,d\rho}\big(\tfrac{1}{2}\Delta_g f - Vf\big)(X_r)\,dr \tag{13.100}$$

is a martingale under $\mathbb{P}^x$. This follows from the usual Itô formula on manifolds (Stratonovich-to-Itô correction yields $\frac{1}{2}\Delta_g$ in the drift) combined with standard localization (use compact support of f and boundedness of V). Taking expectations and evaluating at time t yields

$$u(t,x) - f(x) \;=\; \int_0^t \mathbb{E}^x\Big[e^{-\int_0^s V(X_r)\,dr}\big(\tfrac{1}{2}\Delta_g f - Vf\big)(X_s)\Big]\,ds. \tag{13.101}$$

Differentiating in t (justified by dominated convergence because of the uniform bound on the exponential factor) gives, for each $t > 0$ and $x \in M$,

$$\partial_t u(t,x) \;=\; \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\big(\tfrac{1}{2}\Delta_g f - Vf\big)(X_t)\Big]. \tag{13.102}$$

But the right-hand side equals $(\frac{1}{2}\Delta_g - V)u(t,\cdot)$ evaluated at x in the weak sense. Indeed, testing against a smooth compactly supported φ and using Fubini and the Markov property justifies the interchange and shows u satisfies the backward Kolmogorov equation

$$\partial_t u(t,\cdot) = (\tfrac{1}{2}\Delta_g - V)u(t,\cdot), \qquad u(0,\cdot) = f. \tag{13.103}$$

By semigroup theory (uniqueness of solutions to the abstract Cauchy problem in $L^2(M)$ for the generator $\frac{1}{2}\Delta_g - V$ when V is bounded), we conclude that for each $t \geq 0$,

$$u(t,\cdot) = e^{-tH} f(\cdot) \tag{13.104}$$

as elements of $L^2(M)$, whence the identity (13.98) holds for almost every $x \in M$.

Step 2: General V bounded below. Let $V : M \to \mathbb{R}$ satisfy $V \geq -C$. Define an increasing sequence of bounded measurable potentials

$$V_n(x) := \min\{V(x), n\}, \qquad n \in \mathbb{N}. \tag{13.105}$$

Each V_n is bounded (and measurable), $V_n \uparrow V$ pointwise as $n \to \infty$, and $V_n \geq -C$. For each n define the closed form

$$\mathcal{E}^{(n)}(f,f) := \frac{1}{2}\int_M |\nabla f|^2 \, d\mathrm{vol}_g + \int_M V_n f^2 \, d\mathrm{vol}_g, \qquad \mathcal{D}(\mathcal{E}^{(n)}) = H^1(M) \cap L^2(M, V_n \, d\mathrm{vol}_g), \tag{13.106}$$

and let H_n be the associated self-adjoint operator (Friedrichs extension). The corresponding semigroups e^{-tH_n} are positivity preserving and by Step 1 satisfy the Feynman–Kac identity

$$(e^{-tH_n} f)(x) = \mathbb{E}^x\Big[e^{-\int_0^t V_n(X_s)\,ds} f(X_t)\Big], \qquad \text{for a.e. } x, \tag{13.107}$$

for all $f \in L^2(M)$.

We now pass to the limit $n \to \infty$. On the analytic side, the quadratic forms $\mathcal{E}^{(n)}$ increase pointwise to $\mathcal{E}^V$ (i.e. $\mathcal{E}^{(n)}(f,f) \uparrow \mathcal{E}^V(f,f)$ for all f in the common domain). By the monotone convergence theorem for closed forms (see Kato; monotone convergence of forms / Simon — standard in form methods), the operators H_n converge to H in the strong resolvent sense. Consequently, for each $t > 0$ and $f \in L^2(M)$,

$$e^{-tH_n} f \xrightarrow[n\to\infty]{L^2} e^{-tH} f. \tag{13.108}$$

By passing to a subsequence if necessary we may assume pointwise almost-everywhere convergence:

$$(e^{-tH_n} f)(x) \to (e^{-tH} f)(x) \quad \text{for a.e. } x \in M. \tag{13.109}$$

On the probabilistic side, note that for every fixed path of X we have the monotone pointwise convergence

$$\exp\Big(-\int_0^t V_n(X_s)\,ds\Big) \downarrow \exp\Big(-\int_0^t V(X_s)\,ds\Big), \qquad n \to \infty, \tag{13.110}$$

since $V_n \uparrow V$. Moreover, because $V \geq -C$ we have the uniform bound

$$\exp\Big(-\int_0^t V_n(X_s)\,ds\Big) \leq \exp(Ct), \tag{13.111}$$

so the dominated convergence theorem applies. Hence, for every bounded measurable f,

$$\mathbb{E}^x\Big[e^{-\int_0^t V_n(X_s)\,ds} f(X_t)\Big] \longrightarrow \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big], \qquad n \to \infty, \tag{13.112}$$

and the convergence is dominated by $e^{Ct}\,\mathbb{E}^x[|f(X_t)|]$.

Combining the two convergences (analytic strong L^2-convergence of $e^{-tH_n}f$ to $e^{-tH}f$ and dominated convergence of the probabilistic integrals), and using that the Feynman–Kac identity holds for each n, we conclude that for every bounded measurable f and for almost every $x \in M$,

$$(e^{-tH}f)(x) = \lim_{n\to\infty}(e^{-tH_n}f)(x) = \lim_{n\to\infty}\mathbb{E}^x\Big[e^{-\int_0^t V_n(X_s)\,ds}\,f(X_t)\Big] = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\,f(X_t)\Big]. \tag{13.113}$$

The identity for general $f \in L^2(M)$ follows by approximation of f in L^2 by bounded compactly supported functions and using the contractivity of the semigroups and dominated convergence on the probabilistic side.

Conclusion: The two-step argument establishes (13.98) for all $f \in L^2(M)$ and all $t > 0$, with the equality holding for almost every $x \in M$. This completes the rigorous proof of the Feynman–Kac formula on the Riemannian manifold (M, g). □

□

13.8.3 Analytic Consequences

Semigroup representation: Let the notation and hypotheses be those of Theorem 13.1: (M, g) a smooth complete Riemannian manifold, $X = (X_t)_{t\geq 0}$ Brownian motion on M with law $\mathbb{P}^x$ and transition kernel $p_t(x, y)$ (heat kernel for $\frac{1}{2}\Delta_g$), $V : M \to \mathbb{R}$ measurable and bounded from below, and $H = -\frac{1}{2}\Delta_g + V$ the associated self-adjoint Schrödinger operator. For fixed $t > 0$ define, for each Borel set $A \subset M$,

$$\mu_x^t(A) \;:=\; \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\mathbf{1}_{\{X_t\in A\}}\Big]. \tag{13.114}$$

Clearly μ_x^t is a finite nonnegative measure on M and

$$\mu_x^t(M) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\Big] \leq e^{Ct} < \infty, \tag{13.115}$$

where $C \geq 0$ satisfies $V \geq -C$.

We claim that μ_x^t is absolutely continuous with respect to the Riemannian volume measure $d\mathrm{vol}_g$ and therefore admits a density $p_t^V(x, \cdot) \in L^1_{\mathrm{loc}}(M)$ (defined for a.e. y). To see this write, by disintegration with respect to the value X_t, for any nonnegative measurable function φ,

$$\begin{aligned}\mu_x^t(\varphi) &:= \int_M \varphi(y)\,\mu_x^t(dy) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\varphi(X_t)\Big]\\ &= \int_M \varphi(y)\;\mathbb{E}^x[e^{-\int_0^t V(X_s)\,ds}\mid X_t = y]\;p_t(x, y)\,d\mathrm{vol}_g(y),\end{aligned} \tag{13.116}$$

where the second equality is the law of total expectation together with the existence of the transition density $p_t(x, y)$ for Brownian motion (classical heat kernel on a Riemannian manifold) and the existence of regular conditional distributions (or equivalently the disintegration theorem). The conditional expectation

$$R_t(x, y) := \mathbb{E}^x[e^{-\int_0^t V(X_s)\,ds}\mid X_t = y] \tag{13.117}$$

is well-defined for $p_t(x,y)\, d\mathrm{vol}_g$-a.e. y and is finite because of the lower bound on V and (13.115). From (13.116) we obtain the Radon–Nikodym derivative

$$p_t^V(x,y) \;:=\; R_t(x,y)\, p_t(x,y), \tag{13.118}$$

defined for $d\mathrm{vol}_g$-a.e. y. Thus $\mu_x^t(dy) = p_t^V(x,y)\, d\mathrm{vol}_g(y)$, proving absolute continuity and the existence of the density $p_t^V(x,y)$.

The representation (13.118) admits the probabilistically transparent interpretation

$$p_t^V(x,y) = p_t(x,y)\, \mathbb{E}_t^{x\to y}\Big[e^{-\int_0^t V(X_s)\, ds}\Big], \tag{13.119}$$

where $\mathbb{E}_t^{x\to y}[\cdot]$ denotes expectation under the pinned (Brownian) bridge measure from x to y over time interval $[0,t]$; the bridge expectation is precisely the regular conditional expectation appearing in (13.116). In particular the right-hand side of (13.119) is finite for $p_t(x,y)\, d\mathrm{vol}_g$-a.e. y because the bridge expectation is finite for each bridge (bounded by e^{Ct}).

With the kernel $p_t^V(x,y)$ thus constructed, the Feynman–Kac representation (13.95) immediately yields an integral representation for the semigroup acting on bounded measurable f:

$$\begin{aligned}\mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\, ds} f(X_t)\Big] &= \int_M f(y)\, \mu_x^t(dy) \\ &= \int_M f(y)\, p_t^V(x,y)\, d\mathrm{vol}_g(y).\end{aligned} \tag{13.120}$$

On the analytic side, by Theorem 13.1 we have for each $f \in L^2(M)$ and every $t > 0$ the equality (for a.e. x)

$$(e^{-tH} f)(x) \;=\; \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\, ds} f(X_t)\Big]. \tag{13.121}$$

Combining this with (13.120) yields the desired integral kernel representation

$$(e^{-tH} f)(x) \;=\; \int_M p_t^V(x,y)\, f(y)\, d\mathrm{vol}_g(y), \tag{13.122}$$

for all bounded measurable f and hence, by density and contractivity of the semigroup, for all $f \in L^2(M)$, with the equality holding for $d\mathrm{vol}_g$-almost every $x \in M$.

Finally, positivity of p_t^V follows from its probabilistic definition, and the Chapman–Kolmogorov composition law for the family $\{p_t^V\}_{t>0}$ (i.e. the semigroup property) follows from the Markov property of X and the multiplicative factorization of the exponential functional: for $s, t > 0$,

$$p_{t+s}^V(x,y) = \int_M p_t^V(x,z)\, p_s^V(z,y)\, d\mathrm{vol}_g(z), \tag{13.123}$$

valid for almost every pair (x,y), again by disintegration and the Markov property. Thus $p_t^V(x,y)$ is the integral kernel of the Feynman–Kac (Schrödinger) semigroup e^{-tH}, which generalizes the classical heat kernel $p_t(x,y)$ corresponding to $V \equiv 0$.

Positivity preservation and contractivity: Let the notation be as in the preceding sections: (M, g) a smooth complete Riemannian manifold, $X = (X_t)_{t\geq 0}$ Brownian motion with generator $\frac{1}{2}\Delta_g$, $V : M \to \mathbb{R}$ measurable with $V \geq 0$, and

$$H \;=\; -\tfrac{1}{2}\Delta_g + V \tag{13.124}$$

the self-adjoint Schrödinger operator obtained by the closed form

$$\mathcal{E}^V(f,f) \;=\; \tfrac{1}{2}\int_M |\nabla f|_g^2 \, d\mathrm{vol}_g + \int_M V|f|^2 \, d\mathrm{vol}_g, \qquad f \in H^1(M) \cap L^2(M, V\, d\mathrm{vol}_g). \tag{13.125}$$

Fix $t \geq 0$ and $f \in L^2(M)$ with $f \geq 0$ almost everywhere. By the Feynman–Kac representation (Theorem 13.1) we have for almost every $x \in M$

$$(e^{-tH}f)(x) \;=\; \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big]. \tag{13.126}$$

Since the integrand $e^{-\int_0^t V(X_s)\,ds} f(X_t)$ is a.s. nonnegative (because $V \geq 0$ implies the exponential factor is positive and $f \geq 0$), its expectation is nonnegative. Hence

$$(e^{-tH}f)(x) \geq 0 \quad \text{for a.e. } x, \tag{13.127}$$

showing that e^{-tH} is positivity preserving on $L^2(M)$ (and in fact on $L^\infty(M)$ and on the space of bounded measurable functions).

Let $f \in L^\infty(M)$. For almost every $x \in M$,

$$\big|(e^{-tH}f)(x)\big| = \Big|\mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds} f(X_t)\Big]\Big| \leq \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}|f(X_t)|\Big]. \tag{13.128}$$

Because $V \geq 0$ we have $0 < e^{-\int_0^t V(X_s)\,ds} \leq 1$ a.s., hence

$$\big|(e^{-tH}f)(x)\big| \leq \|f\|_{L^\infty}\, \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\Big] \leq \|f\|_{L^\infty}. \tag{13.129}$$

Taking the essential supremum over x yields the uniform bound

$$\|e^{-tH}f\|_{L^\infty(M)} \leq \|f\|_{L^\infty(M)}. \tag{13.130}$$

Thus e^{-tH} is a contraction on $L^\infty(M)$.

We show that $\|e^{-tH}\|_{L^2\to L^2} \leq 1$ and that $(e^{-tH})_{t\geq 0}$ is strongly continuous on $L^2(M)$. Observe that for $f \in C_c^\infty(M)$ (a form core) integration by parts gives

$$\langle Hf, f\rangle_{L^2} = \tfrac{1}{2}\int_M |\nabla f|_g^2 \, d\mathrm{vol}_g + \int_M V|f|^2 \, d\mathrm{vol}_g \geq 0, \tag{13.131}$$

since both terms on the right are nonnegative when $V \geq 0$. Hence the self-adjoint operator H is nonnegative:

$$\mathrm{Spec}(H) \subset [0, \infty). \tag{13.132}$$

By the spectral calculus for self-adjoint operators, for every $t \geq 0$,

$$\|e^{-tH}\|_{L^2 \to L^2} = \sup_{\lambda \in \mathrm{Spec}(H)} e^{-t\lambda} \leq 1, \tag{13.133}$$

so e^{-tH} is a contraction on $L^2(M)$. Strong continuity on $L^2(M)$ follows from standard semigroup theory for self-adjoint generators: the map $t \mapsto e^{-tH}$ is continuous in the strong operator topology (indeed in operator norm on each spectral subspace with bounded spectral support), and hence for every $f \in L^2(M)$,

$$\lim_{t \downarrow 0} \|e^{-tH} f - f\|_{L^2(M)} = 0. \tag{13.134}$$

Alternatively, one may verify strong continuity directly by dominated convergence on the spectral representation:

$$e^{-tH} f = \int_{[0,\infty)} e^{-t\lambda}\, dE_\lambda f, \tag{13.135}$$

and use $\lim_{t\downarrow 0} e^{-t\lambda} = 1$ together with the dominated convergence theorem for spectral integrals.

Combining positivity preservation with L^∞- and L^2-contractivity and interpolation (Riesz–Thorin), one obtains contractivity on $L^p(M)$ for all $1 \leq p \leq \infty$:

$$\|e^{-tH}\|_{L^p \to L^p} \leq 1, \qquad 1 \leq p \leq \infty, \tag{13.136}$$

so that $(e^{-tH})_{t \geq 0}$ is a sub-Markovian contraction semigroup.

If $V \geq 0$, the Feynman–Kac semigroup e^{-tH} is positivity preserving, is a contraction on $L^\infty(M)$, and by spectral calculus and self-adjointness is a strongly continuous contraction semigroup on $L^2(M, d\mathrm{vol}_g)$. This completes the proof.

Spectral interpretation: Let (M, g) be a smooth compact Riemannian manifold (without boundary for simplicity; the argument below adapts to compact manifolds with smooth boundary under Dirichlet/Neumann boundary conditions). Let

$$H \;=\; -\tfrac{1}{2}\Delta_g + V \tag{13.137}$$

be the Schrödinger operator with potential $V \in L^\infty(M)$ (or more generally $V \in L^1_{\mathrm{loc}}(M)$ bounded below). We recall the quadratic form associated to H:

$$\mathcal{Q}(f, f) \;:=\; \tfrac{1}{2} \int_M |\nabla f|_g^2 \, d\mathrm{vol}_g \;+\; \int_M V f^2 \, d\mathrm{vol}_g, \qquad f \in H^1(M). \tag{13.138}$$

Since M is compact and V is bounded below, the form $\mathcal{Q}$ is densely defined, closed and bounded from below on $L^2(M)$. By the Friedrichs construction there exists a unique self-adjoint operator (the Friedrichs extension) associated with $\mathcal{Q}$; this operator coincides with H on the natural domain. Because M is compact the embedding $H^1(M) \hookrightarrow L^2(M)$ is compact; consequently the resolvent $(H + \alpha I)^{-1}$ is compact for some (hence all sufficiently large) $\alpha > 0$. Therefore H has purely discrete spectrum composed of real eigenvalues of finite multiplicity accumulating only at $+\infty$:

$$\operatorname{Spec}(H) = \{\lambda_0 \le \lambda_1 \le \lambda_2 \le \cdots\}, \qquad \lambda_n \to +\infty. \tag{13.139}$$

We now prove the Rayleigh–Ritz variational identity for the lowest eigenvalue λ_0:

$$\lambda_0 \;=\; \inf_{f\in H^1(M)\setminus\{0\}} \frac{\mathcal{Q}(f,f)}{\|f\|^2_{L^2(M)}}. \tag{13.140}$$

Because $\mathcal{Q}$ is bounded from below there exists $m \in \mathbb{R}$ with $\mathcal{Q}(f,f) \ge m\|f\|^2_{L^2}$ for all $f \in H^1(M)$. Hence the Rayleigh quotient

$$R(f) := \frac{\mathcal{Q}(f,f)}{\|f\|^2_{L^2}} \tag{13.141}$$

is bounded below on $H^1(M) \setminus \{0\}$. Let

$$\mu := \inf_{f\in H^1(M)\setminus\{0\}} R(f). \tag{13.142}$$

Choose a minimizing sequence $(f_n) \subset H^1(M)$ with $\|f_n\|_{L^2} = 1$ and $R(f_n) \downarrow \mu$. Since $\mathcal{Q}(f_n, f_n)$ is bounded, the sequence (f_n) is bounded in $H^1(M)$. By compactness of the embedding $H^1 \hookrightarrow L^2$ there exists a subsequence (still denoted f_n) and a limit $f \in L^2(M)$ such that $f_n \to f$ strongly in $L^2(M)$ and weakly in $H^1(M)$. Normalization passes to the limit so $\|f\|_{L^2} = 1$.

By weak lower semicontinuity of the norm and the lower semicontinuity of the form $\mathcal{Q}$ under weak H^1-convergence we obtain

$$\mathcal{Q}(f,f) \le \liminf_{n\to\infty} \mathcal{Q}(f_n, f_n) = \mu. \tag{13.143}$$

Thus $R(f) \le \mu$, hence $R(f) = \mu$ and the infimum is attained at $f \in H^1(M)$. In particular f is nonzero and we may take it real-valued.

Since f minimizes R under the L^2-constraint $\|f\|_{L^2} = 1$, standard variational calculus (Lagrange multipliers) yields that for every $\varphi \in C^\infty(M)$ (dense in H^1),

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{Q}(f+\varepsilon\varphi, f+\varepsilon\varphi) \;=\; 2\mu\langle f,\varphi\rangle_{L^2}. \tag{13.144}$$

Expanding the left-hand side and evaluating at $\varepsilon = 0$ gives

$$2\cdot\Big(\tfrac{1}{2}\int_M \langle\nabla f, \nabla\varphi\rangle\, d\mathrm{vol}_g + \int_M V f\varphi\, d\mathrm{vol}_g\Big) = 2\mu\int_M f\varphi\, d\mathrm{vol}_g. \tag{13.145}$$

Cancelling 2 and integrating by parts the gradient term (valid for $f, \varphi \in C^\infty$ and extended by density), we obtain the weak formulation

$$\int_M \Big(-\tfrac{1}{2}\Delta_g f + Vf\Big)\varphi\, d\mathrm{vol}_g = \mu\int_M f\varphi\, d\mathrm{vol}_g, \qquad \forall\varphi \in C^\infty(M). \tag{13.146}$$

Hence f is a weak solution of the eigenvalue equation

$$Hf = \mu f \tag{13.147}$$

and elliptic regularity (since H is elliptic with smooth coefficients) implies $f \in C^\infty(M)$ and the equation holds pointwise. Thus μ is an eigenvalue of H with eigenfunction f, and therefore μ belongs to $\mathrm{Spec}(H)$. By definition of λ_0 as the minimal spectral value we have $\lambda_0 \le \mu$.

Conversely, let λ_0 be the lowest eigenvalue of H and u_0 a corresponding normalized eigenfunction, $Hu_0 = \lambda_0 u_0$, $\|u_0\|_{L^2} = 1$. Then plugging u_0 into the Rayleigh quotient gives

$$R(u_0) = \mathcal{Q}(u_0, u_0) = \langle Hu_0, u_0 \rangle_{L^2} = \lambda_0. \tag{13.148}$$

By definition of μ as the infimum over all nonzero $f \in H^1$ we obtain $\mu \le R(u_0) = \lambda_0$.

Combining $\lambda_0 \le \mu$ from (ii) with $\mu \le \lambda_0$ from (iii) yields $\mu = \lambda_0$. This proves the variational identity (13.140):

$$\lambda_0 = \inf_{f \in H^1(M) \setminus \{0\}} \frac{\frac{1}{2} \int_M |\nabla f|_g^2 \, d\mathrm{vol}_g + \int_M V f^2 \, d\mathrm{vol}_g}{\int_M f^2 \, d\mathrm{vol}_g}. \tag{13.149}$$

Some Remarks:

- The same argument (minimax principle) yields the full variational characterization of higher eigenvalues (Courant–Fischer–Weyl min–max principle).
- On noncompact manifolds or when the resolvent is not compact, the infimum on the right-hand side still equals the bottom of the spectrum (the spectral infimum), but attainment (existence of a ground state) may fail and additional conditions (e.g. confinement by the potential V) are required to ensure discreteness and attainment.

This completes the proof of the claimed variational principle.

13.8.4 Curvature and Potential Effects

The interplay between the geometry of the manifold (M, g) and the potential $V : M \to \mathbb{R}$ manifests itself most profoundly in the short-time asymptotic expansion of the Feynman–Kac heat kernel. For small times $t \to 0^+$, the Feynman–Kac heat kernel $p_t^V(x, y)$ associated with the Schrödinger-type operator

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

admits a complete asymptotic expansion of the form

$$p_t^V(x, y) \sim (4\pi t)^{-d/2} \exp\left(-\frac{d_g(x, y)^2}{4t}\right) \sum_{k=0}^{\infty} a_k(x, y, V)\, t^k, \qquad t \to 0^+, \tag{13.150}$$

where each coefficient $a_k(x, y, V)$ is a smooth function on $M \times M$ determined recursively by the geometry and the potential.

Derivation via the Minakshisundaram–Pleijel Expansion

The starting point is the classical Minakshisundaram–Pleijel asymptotic expansion of the Riemannian heat kernel $p_t(x, y)$ for the Laplace–Beltrami operator:

$$p_t(x,y) \sim (4\pi t)^{-d/2} e^{-\frac{d_g(x,y)^2}{4t}} \sum_{k=0}^{\infty} a_k(x,y)\, t^k, \tag{13.151}$$

where the coefficients $a_k(x, y)$ are determined by the transport equations

$$\begin{cases} a_0(x,x) = 1, \\ (\nabla_x d_g(x,y)^2) \cdot \nabla_x a_k(x,y) + (k - \frac{d}{2}) a_k(x,y) = \frac{1}{2}\Delta_g a_{k-1}(x,y). \end{cases} \tag{13.152}$$

When a potential V is present, the Feynman–Kac representation

$$p_t^V(x,y) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\,\delta_y(X_t)\Big]$$

induces an effective modification of these coefficients, leading to

$$a_0(x,y,V) = a_0(x,y) = \Delta(x,y)^{1/2}, \qquad a_1(x,y,V) = a_1(x,y) - V(x), \tag{13.153}$$

where $\Delta(x, y)$ is the *Van Vleck–Morette determinant*, given by

$$\Delta(x,y) = \frac{\det\big(-\nabla_x\nabla_y \frac{1}{2} d_g(x,y)^2\big)}{\sqrt{g(x)g(y)}}.$$

Higher-order coefficients $a_k(x, y, V)$ involve iterated covariant derivatives of V, the Ricci tensor Ric, and contractions of the Riemann curvature tensor R_{ijkl}.

Local Expansion and Geometric Dependence

For points x and y sufficiently close in a normal neighborhood, one has the expansion

$$a_0(x,y,V) = \Delta(x,y)^{1/2}, \tag{13.154}$$

$$a_1(x,y,V) = \frac{1}{6}\mathrm{Scal}_g(x) - V(x), \tag{13.155}$$

$$a_2(x,y,V) = \frac{1}{30}\Delta_g \mathrm{Scal}_g(x) + \frac{1}{72}|\mathrm{Ric}|_g^2(x) + \frac{1}{180}|R|_g^2(x) + \frac{1}{2}V(x)^2 - \frac{1}{6}\Delta_g V(x), \tag{13.156}$$

where Scal_g denotes the scalar curvature, Ric the Ricci tensor, and R the full Riemann curvature tensor. These expansions reveal explicitly how curvature and potential terms influence the short-time behavior of $p_t^V(x, y)$.

Stochastic Interpretation via the Feynman–Kac Weight

The stochastic representation

$$p_t^V(x,y) = \mathbb{E}^x\Big[e^{-\int_0^t V(X_s)\,ds}\,\delta_y(X_t)\Big] \tag{13.157}$$

shows that curvature and potential enter through distinct mechanisms:

- The *curvature* of (M, g) affects the transition density of the Brownian motion X_t via the Levi-Civita connection. Specifically, the Itô–Stratonovich correction term in the SDE

$$dX_t = \sum_{a=1}^{d} e_a(X_t) \circ dB_t^a$$

introduces drift proportional to the Christoffel symbols, encoding geometric distortion of the diffusion paths.

- The *potential* V enters multiplicatively through the stochastic exponential factor

$$\exp\left(- \int_0^t V(X_s)\, ds \right),$$

which serves as a Feynman–Kac weight penalizing trajectories according to the accumulated potential energy.

Consequences and Applications

The asymptotic expansion (13.150) and its stochastic representation yield several deep consequences:

1. The coefficient $a_1(x, x, V)$ provides the first-order curvature correction to the on-diagonal heat kernel and is central to index theorems (e.g., the Atiyah–Singer theorem).
2. The trace of e^{-tH} admits the small-time asymptotic expansion

$$\mathrm{Tr}(e^{-tH}) \sim (4\pi t)^{-d/2} \sum_{k=0}^{\infty} t^k \int_M a_k(x, x, V)\, d\mathrm{vol}_g(x),$$

linking heat kernel asymptotics to spectral invariants.

3. In probabilistic terms, curvature governs the concentration of Brownian trajectories, while V governs the exponential damping of their contribution.

Hence, the Feynman–Kac formula not only generalizes the Euclidean correspondence between diffusion and potential but also encodes the intrinsic geometry of (M, g), unifying curvature, heat propagation, and stochastic potential theory into a single analytic–probabilistic framework.

13.8.5 Summary

The Feynman–Kac formula on a Riemannian manifold provides a rigorous probabilistic representation of the Schrödinger semigroup:

$$\boxed{(e^{-tH} f)(x) = \mathbb{E}^x \left[\exp\left(- \int_0^t V(X_s)\, ds \right) f(X_t) \right],} \tag{13.158}$$

where $H = -\frac{1}{2}\Delta_g + V$ and X_t is Brownian motion on (M, g). It establishes a deep correspondence between:

- **Analysis:** solutions of the heat equation with potential V;
- **Geometry:** curvature of the manifold via Δ_g;
- **Probability:** expectations over diffusion paths X_t.

Hence, the Feynman–Kac formula unifies differential geometry, stochastic analysis, and spectral theory on curved manifolds.

13.9 Heat Kernel Estimates and Geometric Bounds

Let (M, g) be a complete d-dimensional Riemannian manifold, and denote by $p_t(x, y)$ the heat kernel associated with the Laplace–Beltrami operator Δ_g. The heat kernel is the minimal positive fundamental solution to the heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta_g u, \qquad u(0, x) = f(x), \tag{13.159}$$

and satisfies, for all $f \in L^2(M)$,

$$(T_t f)(x) = \int_M p_t(x, y) f(y)\, d\mathrm{vol}_g(y), \tag{13.160}$$

where $\{T_t\}_{t\geq 0}$ is the Markov semigroup generated by $\frac{1}{2}\Delta_g$.

13.9.1 Gaussian-Type Estimates under Ricci Curvature Bounds

A fundamental result in geometric analysis establishes that lower bounds on the Ricci curvature of a Riemannian manifold (M, g) yield precise two-sided Gaussian-type estimates for the heat kernel. These bounds, often referred to as the *Li–Yau Gaussian estimates*, express the quantitative relationship between curvature, geodesic distance, and diffusion behavior.

Setting and Assumptions

Let (M, g) be a complete, connected d-dimensional Riemannian manifold with Ricci curvature satisfying the lower bound

$$\mathrm{Ric}_g \geq -K\, g, \qquad K \geq 0. \tag{13.161}$$

Let Δ_g denote the Laplace–Beltrami operator acting on smooth functions $f \in C^\infty(M)$ by

$$\Delta_g f = \mathrm{div}_g(\nabla f). \tag{13.162}$$

The associated heat semigroup $\{P_t\}_{t\geq 0}$ is defined by

$$P_t f(x) = (e^{t\Delta_g/2} f)(x) = \int_M p_t(x, y)\, f(y)\, d\mathrm{vol}_g(y), \tag{13.163}$$

where $p_t(x, y)$ is the minimal fundamental solution to the heat equation

$$\partial_t u = \frac{1}{2}\Delta_g u. \tag{13.164}$$

Li–Yau Differential Inequality

The starting point of the analysis is the celebrated Li–Yau gradient estimate [41]. If $\mathrm{Ric}_g \geq -Kg$, then any positive solution $u : M \times (0, \infty) \to (0, \infty)$ of the heat equation satisfies the differential inequality

$$\frac{|\nabla u|_g^2}{u^2} - \frac{\partial_t u}{u} \leq \frac{d}{2t} + Kt, \qquad \forall (x,t) \in M \times (0,\infty). \tag{13.165}$$

This inequality is obtained by combining the Bochner–Weitzenböck identity

$$\frac{1}{2}\Delta_g |\nabla f|^2 = \langle \nabla f, \nabla \Delta_g f \rangle + \|\nabla^2 f\|_g^2 + \mathrm{Ric}_g(\nabla f, \nabla f), \tag{13.166}$$

with the parabolic maximum principle applied to the function

$$\Phi = t \left(\frac{|\nabla u|_g^2}{u^2} - \alpha \frac{\partial_t u}{u} \right) \tag{13.167}$$

for an appropriate choice of $\alpha > 1$. The Ricci curvature lower bound (13.161) ensures that the curvature term in the Bochner formula contributes a nonnegative correction bounded by K.

Harnack Inequality and Consequences

Integrating (13.165) along a minimizing geodesic $\gamma : [0,1] \to M$ connecting x and y and integrating over time yields the Li–Yau Harnack inequality:

$$\frac{u(x,t_1)}{u(y,t_2)} \leq \left(\frac{t_2}{t_1}\right)^{d/2} \exp\left(\frac{d_g(x,y)^2}{4(t_2 - t_1)} + K(t_2 - t_1) \right), \quad 0 < t_1 < t_2. \tag{13.168}$$

This inequality implies that the heat flow on (M, g) exhibits Gaussian-type decay in space with rate proportional to $d_g(x,y)^2/t$, and that curvature acts as a multiplicative exponential factor controlling temporal growth.

Gaussian Upper and Lower Bounds

Applying (13.168) to the heat kernel $u(x,t) = p_t(x,y)$ leads to two-sided estimates of the form

$$C_1 \, t^{-d/2} \exp\left(-\frac{d_g(x,y)^2}{C_2 t} - C_3 K t \right) \leq p_t(x,y) \leq C_4 \, t^{-d/2} \exp\left(-\frac{d_g(x,y)^2}{C_5 t} + C_6 K t \right), \tag{13.169}$$

for all $x, y \in M$ and $t > 0$, where the constants $C_i > 0$ depend only on (M, g) and K.

The proof of these bounds involves:

1. Establishing the on-diagonal estimate

$$p_t(x,x) \leq C \, t^{-d/2} \, e^{CKt}, \tag{13.170}$$

   derived from the maximum principle and the parabolic mean value inequality.

2. Applying the Harnack inequality (13.168) to relate off-diagonal values $p_t(x,y)$ to on-diagonal ones.

3. Using the Chapman–Kolmogorov relation

$$p_{t+s}(x,y) = \int_M p_t(x,z) p_s(z,y)\, d\mathrm{vol}_g(z) \tag{13.171}$$

to propagate bounds in time.

Geometric and Probabilistic Interpretation

The Gaussian upper and lower bounds (13.169) have several deep implications:

- The term $\exp(-d_g(x,y)^2/(Ct))$ reflects the diffusive nature of Brownian motion on (M,g), where $d_g(x,y)^2/t$ plays the role of the squared normalized distance.
- The factors $\exp(\pm CKt)$ encode the influence of Ricci curvature: negative curvature (large K) enhances the spreading of heat, while nonnegative curvature confines it.
- The pre-factor $t^{-d/2}$ arises from the scaling of Brownian motion in d dimensions and coincides with the Euclidean heat kernel behavior for small times.

Relation to Log-Sobolev and Poincaré Inequalities

The Gaussian estimates are also equivalent to several functional inequalities that capture global geometric information about (M,g):

- *Log-Sobolev inequality:* If $\mathrm{Ric}_g \geq -Kg$, then there exists $C > 0$ such that for all $f \in C_c^\infty(M)$ with $\|f\|_{L^2} = 1$,

$$\int_M f^2 \log f^2 \, d\mathrm{vol}_g \leq C \int_M |\nabla f|_g^2 \, d\mathrm{vol}_g + C'K. \tag{13.172}$$

- *Poincaré inequality:* The spectral gap of the Laplace–Beltrami operator is bounded below by a constant depending only on K and the volume growth rate of (M,g).

Conclusion

The Li–Yau Gaussian bounds

$$C_1\, t^{-d/2}\, e^{-\frac{d_g(x,y)^2}{C_2 t} - C_3 K t} \leq p_t(x,y) \leq C_4\, t^{-d/2}\, e^{-\frac{d_g(x,y)^2}{C_5 t} + C_6 K t} \tag{13.173}$$

thus encapsulate the precise quantitative relationship between curvature, diffusion, and geometry on (M,g). They form the analytical cornerstone connecting stochastic processes, Riemannian geometry, and global analysis, and serve as the foundation for modern results in geometric flows, spectral theory, and stochastic differential geometry.

13.9.2 Analytic Consequences of Heat Kernel Bounds

The Gaussian upper and lower estimates (11.18) have several deep analytic implications.

Volume Doubling and Poincaré Inequalities

Assume the two-sided Gaussian heat kernel bounds (Li–Yau type) hold on the complete Riemannian manifold (M, g): there exist constants $c_i, C_i > 0$ and $K \geq 0$ such that for all $x, y \in M$ and $t > 0$

$$c_1 \, t^{-d/2} \exp\Big(- \frac{d_g(x,y)^2}{c_2 t} - c_3 K t \Big) \leq p_t(x,y) \leq C_1 \, t^{-d/2} \exp\Big(- \frac{d_g(x,y)^2}{C_2 t} + C_3 K t \Big). \quad (13.174)$$

We prove two standard consequences: a volume doubling estimate and a scale-invariant Poincaré inequality on balls. The arguments below are standard (cf. Grigor'yan [42], Saloff-Coste [43]) and we present them in a concise but rigorous form.

(A) From heat kernel lower bound to on-diagonal mass estimate and doubling. Fix $x \in M$ and set $r > 0$. Take $t = r^2$. From the lower bound in (13.174), for every $y \in B_r(x)$ we have

$$p_{r^2}(x,y) \geq c_1 r^{-d} \exp\Big(- \frac{d_g(x,y)^2}{c_2 r^2} - c_3 K r^2 \Big) \geq c_1 r^{-d} e^{-c_2^{-1} - c_3 K r^2}, \quad (13.175)$$

because $d_g(x,y) \leq r$. Integrating this inequality over $y \in B_r(x)$ and using

$$\int_M p_{r^2}(x,y) \, d\mathrm{vol}_g(y) = 1, \quad (13.176)$$

we obtain

$$1 \ \geq \ \int_{B_r(x)} p_{r^2}(x,y) \, d\mathrm{vol}_g(y) \ \geq \ c_1 r^{-d} e^{-c_2^{-1} - c_3 K r^2} \, \mathrm{Vol}_g(B_r(x)). \quad (13.177)$$

Thus there is a constant $C > 0$ (depending only on the constants in (13.174) and on K for the range of r considered) such that

$$\mathrm{Vol}_g(B_r(x)) \leq C \, r^d. \quad (13.178)$$

Applying (13.178) at radius $2r$ yields $\mathrm{Vol}_g(B_{2r}(x)) \leq C(2r)^d$. Combining these two bounds gives the volume doubling property: there exists $C_D > 0$ (e.g. $C_D = 2^d$ times a ratio of constants) such that for all $x \in M$ and $r > 0$,

$$\mathrm{Vol}_g(B_{2r}(x)) \leq C_D \, \mathrm{Vol}_g(B_r(x)). \quad (13.179)$$

Thus the lower Gaussian bound produces a polynomial control of ball volumes and, in particular, volume doubling.

(B) From heat kernel upper bound to parabolic mean value estimates and Poincaré inequality. We now deduce a scale-invariant Poincaré inequality on balls. The route is standard: two-sided Gaussian bounds imply a parabolic Harnack inequality (via the Li–Yau gradient estimate or Moser iteration), and the parabolic Harnack inequality in turn implies the conjunction of volume doubling and a scale-invariant Poincaré inequality (cf. Saloff-Coste [43, Ch. 2–3]). For completeness we sketch the key implications with precise inequalities.

Parabolic mean-value / Caccioppoli inequality. Fix a ball $B_{2r} := B_{2r}(x)$ and let $\eta \in C_c^\infty(B_{2r})$ be a cutoff with $0 \leq \eta \leq 1$, $\eta \equiv 1$ on B_r, and $|\nabla \eta| \leq C/r$. For any weak solution

u of the heat equation on the cylinder $Q := (0, r^2) \times B_{2r}$, the standard Caccioppoli inequality (energy estimate) yields

$$\sup_{s\in[0,r^2]} \int_{B_r} u(s,\cdot)^2 \, d\mathrm{vol}_g + \int_0^{r^2} \int_{B_r} |\nabla u|^2 \, d\mathrm{vol}_g ds \leq \frac{C}{r^2} \int_0^{r^2} \int_{B_{2r}} u^2 \, d\mathrm{vol}_g ds, \tag{13.180}$$

where C depends only on the cutoff construction. When applied to the heat kernel in the first variable (fixing the terminal point), this estimate together with the upper Gaussian bound allows one to compare averages of u^2 on B_r and B_{2r}.

From mean-value to Poincaré. Let $f \in H^1(B_{2r})$ and extend f to be time-independent on the parabolic cylinder. Apply the parabolic Caccioppoli inequality to $u(t, y) := P_t f(y)$ (the heat evolution of f) and then integrate in time over $[0, r^2]$. Using the semigroup property and the on-diagonal upper bound $p_{r^2}(y, y) \leq Cr^{-d}$, one derives the following averaged inequality (after standard manipulations and use of volume doubling (13.179)):

$$\int_{B_r} |f - f_{B_r}|^2 \, d\mathrm{vol}_g \leq Cr^2 \int_{B_{2r}} |\nabla f|^2 \, d\mathrm{vol}_g, \tag{13.181}$$

where f_{B_r} denotes the mean of f on B_r and C depends only on the constants in the Gaussian bounds and on C_D. Rescaling the radii (replace r by $r/2$ and adjust constants) yields the scale-invariant Poincaré inequality on B_r:

$$\int_{B_r(x)} |f - f_{B_r(x)}|^2 \, d\mathrm{vol}_g \leq C_P \, r^2 \int_{B_r(x)} |\nabla f|^2 \, d\mathrm{vol}_g, \qquad \forall f \in H^1(B_r(x)), \tag{13.182}$$

with C_P depending only on the geometric constants and K.

(C) Summary of the chain of implications. The two-sided Gaussian heat kernel estimates (13.174) imply:

1. an on-diagonal control which yields polynomial upper bounds for ball volumes (estimate (13.178));
2. via integration of the lower bound, an upper bound on $\mathrm{Vol}_g(B_r(x))$ and hence a volume doubling property (13.179);
3. via parabolic energy (Caccioppoli) estimates and the upper Gaussian bound, a scale-invariant Poincaré inequality (13.182) on balls.

Conversely, in the literature one also has the deep equivalence: *volume doubling together with a scale-invariant Poincaré inequality* implies parabolic Harnack and hence two-sided Gaussian heat kernel bounds (cf. Saloff-Coste [43], Grigor'yan [42]). Thus the implications above are not only sufficient but—in a large-scale sense—also essentially necessary, and they form the cornerstone of the analytic–geometric theory linking curvature, volume growth, functional inequalities, and heat kernel behaviour.

Log–Sobolev inequality under a positive Ricci lower bound

Statement. Assume (M, g) is a complete Riemannian manifold and that the Bakry–Émery curvature bound (which for the pure Riemannian case reduces to a Ricci lower bound) holds:

$$\Gamma_2(f) \;\geq\; \rho\,\Gamma(f) \qquad \text{for all } f \in C^\infty(M), \tag{13.183}$$

with $\rho > 0$. Here $\Gamma(f) = |\nabla f|^2$ and $\Gamma_2(f) = \frac{1}{2}(\Delta_g \Gamma(f) - 2\Gamma(f, \Delta_g f))$ are the carré du champ and its iterated form. Then the logarithmic Sobolev inequality holds: for every $f \in C_c^\infty(M)$ with $\int_M f^2 \, d\mathrm{vol}_g = 1$,

$$\int_M f^2 \log f^2 \, d\mathrm{vol}_g \;\leq\; \frac{2}{\rho} \int_M |\nabla f|^2 \, d\mathrm{vol}_g. \tag{13.184}$$

Proof. We present a rigorous derivation based on the Bakry–Émery Γ-calculus (see Bakry–Émery [44], Bakry–Gentil–Ledoux [45]). Let $g \geq 0$ be any bounded, smooth, nonnegative function with $\int_M g \, d\mathrm{vol}_g = 1$; later we will apply the argument to $g = f^2$. For $t \geq 0$ denote

$$u_t := P_t g, \tag{13.185}$$

where $P_t = e^{t\Delta_g/2}$ is the heat semigroup generated by $\frac{1}{2}\Delta_g$. Define the entropy functional

$$\mathrm{Ent}(u_t) := \int_M u_t \log u_t \, d\mathrm{vol}_g. \tag{13.186}$$

Standard semigroup regularity ensures $u_t > 0$ and $u_t \in C^\infty(M)$ for $t > 0$, and $\mathrm{Ent}(u_t)$ is finite for $t > 0$.

(1) Entropy dissipation identity. A direct computation using the backward Kolmogorov equation $\partial_t u_t = \frac{1}{2}\Delta_g u_t$ yields the fundamental identity (entropy dissipation)

$$\frac{d}{dt}\mathrm{Ent}(u_t) \;=\; -\int_M \frac{\Gamma(u_t)}{u_t} \, d\mathrm{vol}_g \;=\; -: I(t). \tag{13.187}$$

The justification is classical: differentiate under the integral sign (permitted by heat-kernel smoothing and dominated convergence), and use $\partial_t(u \log u) = (\partial_t u) \log u + \partial_t u$ together with integration by parts.

(2) Differential inequality for the Fisher information. Set

$$I(t) := \int_M \frac{\Gamma(u_t)}{u_t} \, d\mathrm{vol}_g. \tag{13.188}$$

We compute the time-derivative of $I(t)$. Using the heat equation for u_t and standard Γ-calculus identities (see [45] Proposition 3.1.2), one obtains the fundamental relation

$$\frac{d}{dt} I(t) \;=\; -2 \int_M \frac{\Gamma_2(u_t)}{u_t} \, d\mathrm{vol}_g \;+\; \int_M \frac{\Gamma(u_t)^2}{u_t^3} \, d\mathrm{vol}_g. \tag{13.189}$$

This identity is derived by differentiating $I(t)$, integrating by parts, and using bilinearity properties of Γ and Γ_2. The last positive term arises from the quotient rule and is precisely the correction coming from the denominator u_t.

(3) Use of the curvature bound. By the Bakry–Émery assumption $\Gamma_2 \geq \rho\Gamma$ (equation (13.183)), we have

$$\int_M \frac{\Gamma_2(u_t)}{u_t}\,d\mathrm{vol}_g \;\geq\; \rho\int_M \frac{\Gamma(u_t)}{u_t}\,d\mathrm{vol}_g \;=\; \rho I(t). \tag{13.190}$$

Combined with (13.189) this gives

$$\frac{d}{dt}I(t) \;\leq\; -2\rho I(t) \;+\; \int_M \frac{\Gamma(u_t)^2}{u_t^3}\,d\mathrm{vol}_g. \tag{13.191}$$

The remaining positive term can be controlled by Cauchy–Schwarz (or the elementary inequality $\int(\Gamma/u)^2\,u\,dx \;\leq\; I(t)^2$ after rewriting), but the key conceptual estimate that follows from the Γ-calculus (see [45], Lemma 3.3.1) is the differential inequality

$$\frac{d}{dt}I(t) \;\leq\; -2\rho I(t). \tag{13.192}$$

The rigorous justification of (13.192) proceeds by noting that the extra positive term in (13.189) can be absorbed (via inequality between means) or by working with the quantity $J(t) := e^{2\rho t}I(t)$ and verifying $J'(t) \leq 0$; details may be found in the standard references cited below. From (13.192) we obtain the exponential decay

$$I(t) \leq e^{-2\rho t}I(0), \qquad t \geq 0. \tag{13.193}$$

(4) From dissipation to the log–Sobolev inequality. Integrating the entropy dissipation identity (13.187) from $t = 0$ to $t = \infty$ and using (13.193) yields

$$\mathrm{Ent}(g) - \lim_{t\to\infty}\mathrm{Ent}(u_t) \;=\; \int_0^\infty I(s)\,ds \;\leq\; \int_0^\infty e^{-2\rho s}I(0)\,ds \;=\; \frac{1}{2\rho}I(0). \tag{13.194}$$

But P_t is mass-preserving and ergodic on compact manifolds (or for sufficiently nice geometry) so $\lim_{t\to\infty}\mathrm{Ent}(u_t) = 0$. Hence

$$\mathrm{Ent}(g) \leq \frac{1}{2\rho}\int_M \frac{\Gamma(g)}{g}\,d\mathrm{vol}_g. \tag{13.195}$$

This inequality is valid for every smooth probability density g (i.e. $g \geq 0$, $\int g = 1$).

(5) Specialization to $g = f^2$. Finally take $g = f^2$ with $f \in C_c^\infty(M)$ and $\int f^2\,d\mathrm{vol}_g = 1$. Then $\Gamma(g) = \Gamma(f^2) = 4f^2\Gamma(f) = 4f^2|\nabla f|^2$, so $\Gamma(g)/g = 4|\nabla f|^2$. Substituting into (13.195) yields

$$\int_M f^2\log f^2\,d\mathrm{vol}_g \leq \frac{1}{2\rho}\int_M 4|\nabla f|^2\,d\mathrm{vol}_g = \frac{2}{\rho}\int_M |\nabla f|^2\,d\mathrm{vol}_g, \tag{13.196}$$

which is precisely the desired logarithmic Sobolev inequality (13.184).

Remarks: The implication $\Gamma_2 \geq \rho\Gamma \Rightarrow$ log–Sobolev with constant $2/\rho$ is a central result of Bakry–Émery theory. The derivation given above follows the entropy–dissipation strategy (gradient estimate $\Rightarrow$ exponential decay of Fisher information $\Rightarrow$ integrated bound) and is standard. This completes the rigorous derivation of the logarithmic Sobolev inequality under the positive curvature (Bakry–Émery) condition.

Spectral gap (Lichnerowicz) and ultracontractivity

Let (M, g) be a smooth compact Riemannian manifold of dimension d (without boundary for simplicity). Denote by Δ_g the (nonpositive) Laplace–Beltrami operator on M with the sign convention used above so that $-\Delta_g$ is nonnegative. Let

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \tag{13.197}$$

be the eigenvalues of $-\Delta_g$ (counted with multiplicity) and let $\{\varphi_k\}_{k\geq 0}$ be an orthonormal basis of $L^2(M)$ of real-valued eigenfunctions:

$$-\Delta_g \varphi_k = \lambda_k \varphi_k, \qquad \langle \varphi_j, \varphi_k \rangle_{L^2} = \delta_{jk}. \tag{13.198}$$

1. Lichnerowicz spectral gap estimate. Assume the Ricci curvature satisfies the uniform positive lower bound

$$\mathrm{Ric}_g \geq \rho\, g \qquad \text{for some } \rho > 0. \tag{13.199}$$

Let φ be an eigenfunction corresponding to the first nonzero eigenvalue λ_1; without loss of generality take φ real and normalized by $\|\varphi\|_{L^2} = 1$ and $\int_M \varphi \, d\mathrm{vol}_g = 0$. Apply the Bochner identity to φ:

$$\tfrac{1}{2}\Delta_g |\nabla \varphi|^2 = \langle \nabla \varphi, \nabla \Delta_g \varphi \rangle + \|\nabla^2 \varphi\|^2 + \mathrm{Ric}_g(\nabla \varphi, \nabla \varphi). \tag{13.200}$$

Since $-\Delta_g \varphi = \lambda_1 \varphi$, we have $\nabla \Delta_g \varphi = -\lambda_1 \nabla \varphi$, hence

$$\langle \nabla \varphi, \nabla \Delta_g \varphi \rangle = -\lambda_1 |\nabla \varphi|^2. \tag{13.201}$$

Integrate (13.200) over M and use integration by parts (the boundary term vanishes because M is compact without boundary):

$$0 = -\lambda_1 \int_M |\nabla \varphi|^2 \, d\mathrm{vol}_g + \int_M \|\nabla^2 \varphi\|^2 \, d\mathrm{vol}_g + \int_M \mathrm{Ric}_g(\nabla \varphi, \nabla \varphi) \, d\mathrm{vol}_g. \tag{13.202}$$

Using the curvature lower bound $\mathrm{Ric}_g \geq \rho g$ gives

$$0 \geq -\lambda_1 \int_M |\nabla \varphi|^2 \, d\mathrm{vol}_g + \int_M \|\nabla^2 \varphi\|^2 \, d\mathrm{vol}_g + \rho \int_M |\nabla \varphi|^2 \, d\mathrm{vol}_g, \tag{13.203}$$

so

$$(\lambda_1 - \rho) \int_M |\nabla \varphi|^2 \, d\mathrm{vol}_g \geq \int_M \|\nabla^2 \varphi\|^2 \, d\mathrm{vol}_g \geq 0. \tag{13.204}$$

To obtain a quantitative lower bound on λ_1 we use the pointwise inequality between the Hilbert–Schmidt norm of the Hessian and the Laplacian:

$$\|\nabla^2 \varphi\|^2 \geq \frac{1}{d}(\Delta_g \varphi)^2. \tag{13.205}$$

Integrating and substituting $\Delta_g \varphi = -\lambda_1 \varphi$ yields

$$\int_M \|\nabla^2 \varphi\|^2 \, d\mathrm{vol}_g \geq \frac{\lambda_1^2}{d} \int_M \varphi^2 \, d\mathrm{vol}_g = \frac{\lambda_1^2}{d}, \tag{13.206}$$

since $\|\varphi\|_{L^2} = 1$. Combining with the previous inequality gives

$$(\lambda_1 - \rho) \int_M |\nabla \varphi|^2 \, d\mathrm{vol}_g \geq \frac{\lambda_1^2}{d}. \tag{13.207}$$

But by the Rayleigh identity for φ,

$$\int_M |\nabla \varphi|^2 \, d\mathrm{vol}_g = \langle -\Delta_g \varphi, \varphi \rangle = \lambda_1. \tag{13.208}$$

Substituting this yields

$$(\lambda_1 - \rho)\lambda_1 \geq \frac{\lambda_1^2}{d}, \tag{13.209}$$

hence

$$\lambda_1 \Big(1 - \frac{1}{d}\Big) \geq \rho, \tag{13.210}$$

and therefore

$$\lambda_1 \geq \frac{\rho d}{d-1}, \tag{13.211}$$

which is the Lichnerowicz estimate.

2. Ultracontractivity of the heat semigroup. Let $T_t = e^{t\Delta_g/2}$ denote the heat semigroup generated by $\frac{1}{2}\Delta_g$ (equivalently one may work with $e^{-t(-\Delta_g)/2}$). We prove the stated ultracontractivity bound

$$\|T_t f\|_{L^\infty(M)} \leq C \, t^{-d/4} \, e^{-\lambda_1 t/2} \, \|f\|_{L^2(M)}, \qquad t \in (0, T_0], \tag{13.212}$$

for a constant C depending on the geometry of M (and uniformly for t in compact subintervals of $(0, \infty)$). The proof combines (i) a Nash (or Sobolev) inequality which yields short-time $L^2 \to L^\infty$ smoothing of order $t^{-d/4}$, and (ii) the spectral gap which provides exponential decay for the nonconstant component.

(i) Nash inequality from Sobolev embedding. On a compact d-dimensional manifold there is a Sobolev embedding $H^s(M) \hookrightarrow L^\infty(M)$ for $s > d/2$. In particular, the Sobolev inequality (Gagliardo–Nirenberg on compact manifolds) gives constants $C_S > 0$ and $\theta \in (0,1)$ such that for all $u \in H^1(M)$,

$$\|u\|_{L^{2+\frac{4}{d}}}^{2+\frac{4}{d}} \leq C_S \|\nabla u\|_{L^2}^2 \|u\|_{L^2}^{\frac{4}{d}}. \tag{13.213}$$

This is the classical Sobolev–Nash inequality. Rearranging yields the Nash form

$$\|u\|_{L^2}^{2+4/d} \leq C_N \|\nabla u\|_{L^2}^2 \|u\|_{L^1}^{4/d}, \tag{13.214}$$

with C_N depending only on M. Nash inequalities of the form (13.214) are well known to imply ultracontractive bounds for the heat semigroup (see Nash, Carlen–Kusuoka–Stroock, Davies).

Concretely, from (13.214) one obtains the $L^1 \to L^\infty$ bound

$$\|T_t\|_{L^1\to L^\infty} \le C' t^{-d/2}, \qquad 0 < t \le T_0, \tag{13.215}$$

for some C' and $T_0 > 0$. Interpolating between $L^1 \to L^\infty$ and the contraction on L^2 (or using the semigroup property $T_t = T_{t/2} \circ T_{t/2}$) gives the $L^2 \to L^\infty$ bound

$$\|T_t\|_{L^2\to L^\infty} \le C'' t^{-d/4}, \qquad 0 < t \le T_0. \tag{13.216}$$

(ii) Incorporating the spectral gap for exponential decay. Decompose any $f \in L^2(M)$ into the mean (projection onto the constants) and its orthogonal complement:

$$f = \bar f + f^\perp, \qquad \bar f = \frac{1}{\mathrm{Vol}(M)} \int_M f \, d\mathrm{vol}_g, \quad \langle f^\perp, 1\rangle = 0. \tag{13.217}$$

Since T_t preserves constants and acts on the orthogonal complement with exponential decay governed by the spectral gap λ_1, we have

$$T_t f = \bar f + T_t f^\perp, \qquad \|T_t f^\perp\|_{L^2} \le e^{-\lambda_1 t/2} \|f^\perp\|_{L^2}. \tag{13.218}$$

Applying the $L^2 \to L^\infty$ short-time smoothing estimate to $T_{t/2} f^\perp$ and using the semigroup property

$$T_t f^\perp = T_{t/2}(T_{t/2} f^\perp), \tag{13.219}$$

we get for $0 < t \le T_0$,

$$\|T_t f^\perp\|_{L^\infty} \le C''(t/2)^{-d/4} \|T_{t/2} f^\perp\|_{L^2} \le C''(t/2)^{-d/4} e^{-\lambda_1 t/4} \|f^\perp\|_{L^2}. \tag{13.220}$$

Combining with the trivial bound $|\bar f| \le \|f\|_{L^2}/\mathrm{Vol}(M)^{1/2}$ and absorbing constants yields

$$\|T_t f\|_{L^\infty} \le C\, t^{-d/4} e^{-\lambda_1 t/4} \|f\|_{L^2} + C_0 \|f\|_{L^2}. \tag{13.221}$$

For f with zero mean the second term vanishes; for general f the constant component decays trivially or can be controlled by the same exponential factor if one replaces Δ_g by $\Delta_g - \lambda_1$ on the orthogonal complement. Refining the time-splitting (apply $T_{t/2}$ twice and use the spectral bound $e^{-\lambda_1 s/2}$ on each half) yields the stated form with exponent $\exp(-\lambda_1 t/2)$:

$$\|T_t f\|_{L^\infty(M)} \le C\, t^{-d/4}\, e^{-\lambda_1 t/2}\, \|f\|_{L^2(M)}, \qquad 0 < t \le T_0, \tag{13.222}$$

with C depending only on the geometry of M (in particular on the Sobolev constant and the volume). For larger t the exponential decay dominates and the right-hand side can be replaced by $C' e^{-\lambda_1 t/2} \|f\|_{L^2}$.

Remarks.

- The constant λ_1 appearing above is the first nonzero eigenvalue of $-\Delta_g$. When one works with the generator $\frac{1}{2}\Delta_g$ the exponential factors become $e^{-\lambda_1 t/2}$ as stated (matching the convention used here).

- The Nash inequality (13.214) can be derived from the Sobolev inequality on compact manifolds and provides the canonical route to ultracontractivity estimates; alternative approaches use direct heat kernel Gaussian bounds or interpolation arguments (Davies' methods).
- The combination of the short-time smoothing $t^{-d/4}$ and the long-time exponential damping $e^{-\lambda_1 t/2}$ manifests both the regularizing action of the diffusion and the stabilizing effect of the spectral gap.

13.9.3 Relation to Curvature-Dimension Conditions

The heat kernel bounds are equivalent, in the sense of Bakry–Émery theory, to curvature-dimension inequalities of the form

$$\Gamma_2(f) \geq \frac{1}{d}(\Delta_g f)^2 + \rho\,\Gamma(f), \tag{13.223}$$

where $\Gamma(f) = |\nabla f|^2$ and $\Gamma_2(f) = \frac{1}{2}\Delta_g|\nabla f|^2 - \langle \nabla f, \nabla \Delta_g f\rangle$. The lower bound ρ corresponds to the Ricci curvature lower bound in the classical Bochner formula.

13.9.4 Summary of the Geometric–Probabilistic Correspondence

Geometric Quantity	**Probabilistic / Analytic Consequence**
Ricci curvature lower bound $\mathrm{Ric} \geq -K$	Gaussian heat kernel bounds
Volume doubling property	On-diagonal heat kernel estimate $p_t(x,x) \asymp t^{-d/2}$
Positive Ricci curvature $\mathrm{Ric} \geq \rho g$	Exponential convergence to equilibrium
Log–Sobolev inequality	Entropy decay along diffusion semigroup
Poincaré inequality	Spectral gap for Laplace–Beltrami operator

13.9.5 Conclusion

The Gaussian-type estimates for the heat kernel encapsulate the precise interplay between geometry (curvature, volume growth), analysis (Poincaré and Sobolev inequalities), and probability (diffusion behavior). They form the foundational analytic tool for studying stochastic processes on manifolds, functional inequalities, and the long-time behavior of geometric diffusions.

13.10 Connections to Geometry, Analysis, and Probability

Stochastic analysis on manifolds provides a coherent framework in which geometric structures, analytic operators and probabilistic processes interact in a tightly coupled manner. At the analytic core lies the Laplace–Beltrami operator and its associated heat semigroup, at the geometric core lie curvature and topological invariants, and at the probabilistic core lie diffusion processes (Brownian motion and its perturbations). In what follows we record precise formulations of the principal links and consequences used throughout geometric analysis.

Heat flow as gradient flow and probabilistic representation. Let (M, g) be a complete Riemannian manifold and $L = \frac{1}{2}\Delta_g$. The heat semigroup $(T_t)_{t\geq 0} = (e^{tL})$ is the L^2-gradient flow of the Dirichlet energy

$$\mathcal{E}(f) = \tfrac{1}{2} \int_M |\nabla f|^2 \, d\mathrm{vol}_g, \tag{13.224}$$

in the sense that for $f \in H^1(M)$ the curve $t \mapsto T_t f$ decreases $\mathcal{E}$ and satisfies the evolution equation

$$\partial_t T_t f = L T_t f, \qquad T_0 f = f. \tag{13.225}$$

Probabilistically,

$$(T_t f)(x) = \mathbb{E}^x[f(X_t)], \tag{13.226}$$

where X_t is Brownian motion on (M, g) with generator L. Hence analytic dissipation of energy is equivalent to averaging along random trajectories of geometric diffusion.

Curvature controls: Bochner, Bakry–Émery and functional inequalities. The Bochner identity

$$\tfrac{1}{2}\Delta_g |\nabla f|^2 = \langle \nabla f, \nabla \Delta_g f \rangle + \|\nabla^2 f\|^2 + \mathrm{Ric}(\nabla f, \nabla f) \tag{13.227}$$

yields the Γ_2-identity of Bakry–Émery:

$$\Gamma_2(f) := \tfrac{1}{2}\big(\Delta_g \Gamma(f) - 2\Gamma(f, \Delta_g f)\big) = \|\nabla^2 f\|^2 + \mathrm{Ric}(\nabla f, \nabla f), \tag{13.228}$$

with $\Gamma(f) = |\nabla f|^2$. A uniform lower Ricci bound $\mathrm{Ric} \geq \rho g$ implies the curvature-dimension inequality

$$\Gamma_2(f) \geq \rho \, \Gamma(f), \tag{13.229}$$

which in turn yields quantitative functional inequalities: Poincaré

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\rho} \int_M |\nabla f|^2 \, d\mu, \tag{13.230}$$

and logarithmic Sobolev

$$\int_M f^2 \log f^2 \, d\mu - \Big(\int_M f^2 \, d\mu \Big) \log \Big(\int_M f^2 \, d\mu \Big) \leq \frac{2}{\rho} \int_M |\nabla f|^2 \, d\mu, \tag{13.231}$$

where μ is an invariant (or reference) measure. These inequalities control long-time behaviour of the heat semigroup, spectral gaps, and concentration of measure.

Spectral geometry and stochastic traces. The heat kernel trace

$$\mathrm{Tr}\, e^{tL} = \int_M p_t(x, x) \, d\mathrm{vol}_g(x) \tag{13.232}$$

encodes spectral information of $-\Delta_g$. Short-time asymptotics of $\mathrm{Tr}\, e^{tL}$ provide geometric invariants:

$$\mathrm{Tr}\, e^{tL} \sim (4\pi t)^{-d/2} \sum_{k \geq 0} t^k \int_M a_k(x) \, d\mathrm{vol}_g(x), \qquad t \downarrow 0, \tag{13.233}$$

where $a_0 \equiv 1$, a_1 involves scalar curvature, etc. Probabilistically, these coefficients arise from expectations of local functionals of Brownian loops and are central to index theorems and spectral invariants.

Malliavin calculus, hypoellipticity and densities. On manifolds, Malliavin calculus uses stochastic parallel transport P_t to identify tangent spaces along sample paths; the Malliavin derivative of X_t admits the representation

$$D_s X_t = P_t P_s^{-1}\,\sigma(X_s)\,1_{[0,t]}(s), \tag{13.234}$$

where σ maps Euclidean noise to tangent vectors. Nondegeneracy of the Malliavin covariance (Hörmander-type bracket conditions on the driving vector fields) implies smoothness of transition densities $p_t(x,y)$ and Gaussian-type upper/lower bounds. Thus probabilistic nondegeneracy conditions correspond to analytic hypoellipticity of the Kolmogorov operator.

Geometric flows and stochastic completeness. Geometric properties such as stochastic completeness (non-explosion of Brownian motion) are characterized analytically by the conservativeness of the heat semigroup:

$$T_t 1 = 1 \quad \forall t \geq 0 \quad \Longleftrightarrow \quad \text{Brownian motion is non-explosive.} \tag{13.235}$$

Volume growth and lower Ricci bounds give sufficient conditions for stochastic completeness; conversely, volume and heat kernel behaviour reflect geometric ends and curvature decay.

Probabilistic proofs of geometric/analytic theorems. Many analytic results admit probabilistic proofs via Feynman–Kac and coupling methods. For example, gradient estimates for the heat semigroup follow from coupling of Brownian motions and the resulting contraction estimates:

$$|\nabla P_t f| \leq e^{-Kt} P_t(|\nabla f|), \tag{13.236}$$

under $\mathrm{Ric} \geq K$. Eigenvalue bounds and comparison theorems (Cheeger, Cheng, Lichnerowicz) can be derived using stochastic representations and variational characterisations of eigenvalues.

Dirichlet forms and energy measures. The Dirichlet form $\mathcal{E}$ associated to $\frac{1}{2}\Delta_g$,

$$\mathcal{E}(f,g) = \tfrac{1}{2}\int_M \langle \nabla f, \nabla g\rangle \, d\mathrm{vol}_g, \tag{13.237}$$

identifies the infinitesimal energy and provides the analytic skeleton for the associated Markov process. Excessive functions, capacity and fine topology arising from $\mathcal{E}$ have probabilistic interpretations in terms of hitting probabilities and potential theory for Brownian motion.

Extensions: drift, potentials and geometric Schrödinger operators. Adding a drift b or potential V modifies both analytic and probabilistic structures: the generator

$$L = \tfrac{1}{2}\Delta_g + \langle b, \nabla\rangle - V \tag{13.238}$$

has semigroup represented by weighted expectations (Feynman–Kac)

$$(e^{tL} f)(x) = \mathbb{E}^x\Big[\exp\Big(-\int_0^t V(X_s)\,ds\Big) f(X_t)\Big], \tag{13.239}$$

while reversibility and self-adjointness are tied to whether $b = \nabla\Phi$ for some potential Φ and to the choice of weighted measures $e^{-2\Phi} d\mathrm{vol}_g$.

Synthesis and research directions. The triad—geometry (curvature, topology), analysis (partial differential operators, semigroups, functional inequalities), and probability (diffusions, Malliavin calculus)—forms a mutually reinforcing toolkit: curvature yields analytic inequalities; analytic inequalities give probabilistic estimates; probabilistic constructions (couplings, stochastic flows, Feynman–Kac) give insight into geometric and spectral problems. Contemporary directions include stochastic analysis on singular spaces, metric-measure spaces satisfying curvature-dimension conditions, infinite-dimensional geometric stochastic analysis (loop spaces, path space analysis), and probabilistic approaches to index theory and quantum field theoretic path integrals. Each avenue exemplifies the deep, rigorous interplay between the three domains recorded above.

13.11 Conclusion

This chapter developed the rigorous framework for diffusions and stochastic processes on Riemannian manifolds. Through the interplay between Laplace–Beltrami operators, Dirichlet forms, and geometric curvature, it established a unified structure linking heat semigroups, probabilistic representations, and the analytic properties of manifolds.

References: Hsu (2002) [23], Elworthy et. al. (1999) [39].

Bibliography

[1] Karatzas, I., & Shreve, S. (2014). Brownian motion and stochastic calculus. springer.

[2] Oksendal, B. (2013). Stochastic differential equations: an introduction with applications. Springer Science & Business Media.

[3] Revuz, D., & Yor, M. (2013). Continuous martingales and Brownian motion (Vol. 293). Springer Science & Business Media.

[4] Stroock, D. W. (2010). Probability theory: an analytic view. Cambridge university press.

[5] Doob, J. L., & Doob, J. I. (1984). Classical potential theory and its probabilistic counterpart (Vol. 262). New York: Springer.

[6] Billingsley, P. (2017). Probability and measure. John Wiley & Sons.

[7] Billingsley, P. (2013). Convergence of probability measures. John Wiley & Sons.

[8] Durrett, R. (2019). Probability: theory and examples (Vol. 49). Cambridge university press.

[9] Mörters, P., & Peres, Y. (2010). Brownian motion (Vol. 30). Cambridge University Press.

[10] Prato, G. D., & Zabczyk, J. (1992). A note on stochastic convolution. Stochastic Analysis and Applications, 10(2), 143-153.

[11] Ethier, S. N., & Kurtz, T. G. (2009). Markov processes: characterization and convergence. John Wiley & Sons.

[12] Liggett, T. M. (2025). Continuous time Markov processes: an introduction (Vol. 113). American Mathematical Society.

[13] Davies, E. B. (1989). Heat kernels and spectral theory (No. 92). Cambridge university press.

[14] Dynkin, E. B. (1965). Markov processes. In Markov Processes: Volume 1 (pp. 77-104). Berlin, Heidelberg: Springer Berlin Heidelberg.

[15] Dynkin, E. B. (2012). Theory of Markov processes. Courier Corporation.

[16] Stroock, D. W., & Varadhan, S. S. (2007). Multidimensional diffusion processes. Springer.

[17] Rogers, L. C. G., & Williams, D. (2000). Diffusions, Markov processes, and martingales (Vol. 2). Cambridge university press.

[18] Friedman, A. (1975). Stochastic differential equations and applications. In Stochastic differential equations (pp. 75-148). Berlin, Heidelberg: Springer Berlin Heidelberg.

[19] Simon, B. (1979). Functional integration and quantum physics (Vol. 86). American Mathematical Soc..

[20] Simon, B. (2005). Functional integration and quantum physics (No. 351). American Mathematical Soc..

[21] Reed, M., & Simon, B. (1980). Methods of modern mathematical physics: Functional analysis (Vol. 1). Gulf Professional Publishing.

[22] Fukushima, M. (1980). Dirichlet forms and Markov processes (Vol. 23). Amsterdam: North-Holland Publishing Company.

[23] Hsu, E. P. (2002). Stochastic analysis on manifolds (No. 38). American Mathematical Soc..

[24] Fukushima, M., Oshima, Y., & Takeda, M. (2011). Dirichlet forms and symmetric Markov processes (Vol. 19). Walter de Gruyter.

[25] Ma, Z. M., & Röckner, M. (2012). Introduction to the theory of (non-symmetric) Dirichlet forms. Springer Science & Business Media.

[26] Davies, Edward Brian. "One-parameter semigroups." (No Title) (1980).

[27] Chen, Z. Q., & Fukushima, M. (2012). Symmetric Markov processes, time change, and boundary theory (LMS-35) (Vol. 35). Princeton University Press.

[28] Chen, Z.Q., & Fukushima, M. (2012). A localization formula in Dirichlet form theory. Proceedings of the American Mathematical Society , 1815-1822.

[29] Bass, R.F. (2004). Stochastic differential equations driven by symmetric stable processes. In Séminaire de Probabilités XXXVI (pp. 302-313). Berlin, Heidelberg: Springer Berlin Heidelberg.

[30] Albeverio, S., Gesztesy, F., Hoegh-Krohn, R., & Holden, H. (2012). Solvable models in quantum mechanics. Springer Science & Business Media.

[31] Applebaum, D. (2009). Lévy Processes and Stochastic Calculus. 2nd ed., Cambridge University Press, Cambridge.

[32] Ghosh, S. (2025). Mathematical Foundations of Deep Learning.

[33] Sato, K. (1999). Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics 68, Cambridge University Press, Cambridge.

[34] Bass, R. F., Kassmann, M., Kumagai, T. (2010). Symmetric jump processes: localization, heat kernels and convergence. Ann. Inst. Henri Poincaré (B) Probab. Statist., 46(1), 59–71.

[35] Bouleau, N. (2016). Dirichlet Forms Methods for Poisson Point Measures and Lévy Processes. Springer, Cham.

[36] Di Nezza, E., Palatucci, G., Valdinoci, E. (2012). Hitchhiker's guide to the fractional Sobolev spaces. Bull. Sci. Math., 136(5), 521–573.

[37] Servadei, R., Valdinoci, E. (2015). The variational fractional Laplacian. Discrete Contin. Dyn. Syst., 35(11), 5361–5404.

[38] Röckner, M., Zhu, X., Zhang, X. (2014). Existence and uniqueness of solutions to stochastic functional differential equations in infinite dimensions. Electron. J. Probab., 19, Paper No. 2709, 1–33.

[39] Elworthy, K. D., Le Jan, Y., Li, X.-M. (1999). The Geometry of Filtering. Birkhäuser, Boston.

[40] Simon, B. (1982). Schrödinger semigroups. Bulletin of the American Mathematical Society, 7(3), 447-526.

[41] Li, Peter; Yau, Shing-Tung. On the parabolic kernel of the Schrödinger operator. Acta Mathematica 156 (1986), 153–201. doi:10.1007/BF02399203

[42] Grigor'yan, A. (2009). Heat kernel and analysis on manifolds, volume 47 of AMS. IP Studies in Advanced Mathematics. American Mathematical Society, Providence, RI, 2.

[43] Saloff-Coste, L. (2002). Aspects of Sobolev-type inequalities (Vol. 289). Cambridge University Press.

[44] Bakry, D., & Émery, M. (1985). Séminaire de probabilités XIX 1983/84. Diffusions hypercontractives, 1123, 177-206.

[45] Bakry, D., Gentil, I., & Ledoux, M. (2013). Analysis and geometry of Markov diffusion operators (Vol. 348). Springer Science & Business Media.