LEONARD GROSS’S WORK IN INFINITE-DIMENSIONAL ANALYSIS AND HEAT KERNEL ANALYSIS

BRIAN C. HALL*

Abstract. This paper describes a certain part of Leonard Gross’s work in infinite-dimensional analysis, connected to the Gross Ergodicity Theorem. I then look at ways in which Gross’s work helped to create a new subject within (mostly) finite-dimensional analysis, a subject which may be called “harmonic analysis with respect to heat kernel measure.” This subject transfers to Lie groups certain constructions on $\mathbb{R}^n$ that involves a Gaussian measure. On the Lie group, the role of the Gaussian measure is played by a heat kernel measure.

1. The Gross Ergodicity Theorem and its Consequences

The purpose of this article is to outline one part of Leonard Gross’s work in infinite-dimensional analysis and to show how that work lead to the development of a new discipline within (mostly) finite-dimensional analysis, a discipline which may be called “harmonic analysis with respect to heat kernel measure,” or “heat kernel analysis” more briefly. The story begins with what is now called the Gross Ergodicity Theorem, established in [21]. The result may be described as follows. Let $K$ be a connected compact Lie group equipped with a bi-invariant Riemannian metric. Let $W(K)$ denote the continuous path group, i.e., the group of continuous maps of $[0,1]$ into $K$ sending 0 to the identity $e$ in $K$. We consider on $W(K)$ the Wiener measure $\rho$. Now let $\mathcal{L}(K)$ denote the finite-energy loop group, i.e., the group of maps of $[0,1]$ into $K$ sending both 0 and 1 to $e$ and having one distributional derivative in $L^2$. Then the left action of $\mathcal{L}(K)$ on $W(K)$ leaves the Wiener measure quasi-invariant.

**Theorem 1.1.** Given $f \in L^2(W(K), \rho)$, suppose that for all $l \in \mathcal{L}(K)$, $f(l \cdot g) = f(g)$ for almost every $g$ in $W(K)$. Then there exists a measurable function $\phi$ on $K$ such that $f(g) = \phi(g(1))$ for almost every $g$ in $W(K)$.

That is to say, if a function on the path group is invariant under the left action of the loop group, then that function depends only on the endpoint of the path. The difficulty in this theorem comes in the mismatch between the levels of smoothness: The paths in $W(K)$ have to be continuous rather than finite energy, because the Wiener measure does not live on finite-energy paths (they form a set of measure zero). Meanwhile, the loops in $\mathcal{L}(K)$ have to be finite energy rather than continuous.

---

2000 Mathematics Subject Classification. Primary 60H30.

Key words and phrases. Gross Ergodicity Theorem, Taylor map, Hermite expansion, Segal–Bargmann transform.

* Supported in part by NSF Grant DMS-0555862.
than continuous, because only finite-energy loops act in a way that leaves the Wiener measure quasi-invariant. If only it were possible to use paths and loops of the same smoothness, the result would be easy. If two continuous paths \( g_1 \) and \( g_2 \) have the same endpoint, then there is a loop \( l \) such that \( lg_1 = g_2 \). However, the loop \( l \) will only be continuous rather than finite energy, so invariance of \( f \) under \( \mathcal{L}(K) \) does not imply (at least in any obvious way) that \( f(lg_1) = f(g_2) \).

From Theorem 1.1, Gross was able to deduce that the action of \( \mathcal{L}(K) \) on the continuous loop group is ergodic (i.e., all invariant subsets have measure 0 or measure 1). This ergodicity theorem has been used to prove that the pinned Wiener measure on the continuous loop group is equivalent to the heat kernel measure of Malliavin. Absolute continuity in one direction was established by Driver and Srimurthy in [13]; absolute continuity in the other direction follows easily from the Gross Ergodicity Theorem, as shown by Aida and Driver [1].

Gross’s proof of Theorem 1.1 begins by “linearizing” the problem, using the Itô map. Let \( W(\mathfrak{k}) \) denote the continuous path space (beginning at the origin) in the Lie algebra \( \mathfrak{k} \) of \( K \). Then to a path \( X(\cdot) \) in \( W(\mathfrak{k}) \) we associate a path \( \theta(X)(\cdot) \) in \( W(K) \) by solving the Stratonovich differential equation

\[
\frac{d\theta(t)}{dt} = \theta(t) \circ dX(t).
\]

This stochastic differential equation gives a map (defined almost everywhere) from \( W(\mathfrak{k}) \) to \( W(K) \), and the push-forward under \( \theta \) of the Wiener measure on \( W(\mathfrak{k}) \) coincides with the Wiener measure on \( W(K) \). Given, then, a function \( f \) on \( W(K) \), we can turn it into a function on \( W(\mathfrak{k}) \) by composing \( f \) with \( \theta \). Under this map, the loop-invariant functions (those satisfying \( f(l \cdot g) = f(g) \)) corresponds to functions on \( W(\mathfrak{k}) \) invariant under a certain action of the loop group \( \mathcal{L}(K) \). (The action of the loop group on \( W(\mathfrak{k}) \) has the form of a gauge transformations in Yang–Mills theory; more on this point below.)

After using the Itô map to transfer the problem to the linear path space \( W(\mathfrak{k}) \), Gross uses the so-called chaos expansion, that is, the expansion of a function on \( W(\mathfrak{k}) \) as a sum of multiple Wiener integrals. This expansion is the infinite-dimensional counterpart to the expansion of a function on \( L^2(\mathbb{R}^n, \text{Gauss}) \) as a sum of Hermite polynomials. In general, the \( n \)th “coefficient” in the chaos expansion is function on the \( n \)-simplex with values in the tensor-product space \( \mathfrak{k}^\otimes n \). As shown in Theorem 5.1 of [21], for a loop-invariant function on \( W(\mathfrak{k}) \), these tensor-valued functions are all constants. Furthermore, the constants fit together to form an element of the tensor algebra over \( \mathfrak{k} \) that is orthogonal to the ideal \( J \) generated by elements of the form \( XY - YX - [X,Y] \). From this, Gross was able to then show that loop-invariant functions are actually endpoint functions.

A consequence of Gross’s proof is a sort of “Hermite expansion” for the compact group \( K \). In infinite-dimensional terms, one starts with a function \( \phi \) on \( K \), forms the “endpoint function” \( f(X) = \phi(\theta(X)(1)) \). That is to say, an endpoint function is a function \( f \) on \( W(\mathfrak{k}) \) such that the value of \( f(X) \) depends only on the endpoint of the Itô map \( \theta \) applied to \( X \). If one performs the chaos expansion of the function \( f \) associated to a function \( \phi \) on \( K \), one obtains an element of the space \( J^\perp \). The squared \( L^2 \) norm of the function \( \phi \) on \( K \) can be expressed as a sum of squares
of coefficients of the element of $J^\perp$. Furthermore, if $K$ is simply connected, then Gross shows that every element of $J^\perp$ arises from some function $\phi$ on $K$. The space $J^\perp$ may be thought of as a (completion of) the universal enveloping algebra of the Lie algebra $\mathfrak{k}$. The Hermite expansion on $K$ should be thought of as a generalization of the identification of $L^2(\mathbb{R}^n, \text{Gauss})$ with the Fock space of symmetric tensors. When $\mathbb{R}^n$ is replaced by the Lie group $K$, the Fock space of symmetric tensors is replaced by the universal enveloping algebra.

Of course, once it is known that there is a map of this sort (from functions on $K$ to elements of $J^\perp$), it is possible to describe this map in purely finite-dimensional terms. In Gross’s original paper [21], the finite-dimensional description is given in terms of how the map intertwines certain “annihilation operators.” (See also [22, 23].) O. Hijab then gave a more explicit finite-dimensional description of the map in terms of derivatives of heat kernels. This lead to purely finite-dimensional proofs of the properties of the $J^\perp$ expansion for $K$ by Hijab [41, 42] and B. Driver [9]. The most difficult part of these proofs is establishing that the map is onto $J^\perp$ in the case that $K$ is simply connected.

The $J^\perp$ expansion established by Gross is analogous to the expansion of a function in $L^2(\mathbb{R}^n, \text{Gauss})$ as a sum of Hermite polynomials. Under this analogy, the Gaussian measure on $\mathbb{R}^n$ is replaced by a heat kernel measure on the compact group $K$. The Hermite polynomials on $\mathbb{R}^n$ are then replaced by logarithmic-type derivatives of the heat kernel. Specifically, we may identify the $J^\perp$ with (a completion of) the universal enveloping algebra of the Lie algebra $\mathfrak{k}$ of $K$. We think of the enveloping algebra as the algebra of left-invariant differential operators on $K$.

Given a left-invariant operator $\alpha$ on $K$ (an element of $J^\perp$), we associate to it the function $(\alpha \rho_t)/\rho_t$, where $\rho_t$ is the heat kernel (at the identity) on $K$. In the $\mathbb{R}^n$ case, the heat kernel is just a Gaussian and such logarithmic-type derivatives of the heat kernel are the usual Hermite polynomials.

The Hermite expansion that came out of Gross’s work lead him to suggest to me that I look for an analog of the Segal–Bargmann transform on a compact Lie group. This work became my PhD thesis and resulted in the paper [25]. The classical Segal–Bargmann transform [5, 53] is a unitary map from $L^2(\mathbb{R}^n, \text{Gauss})$ onto the $L^2$-space of holomorphic functions on $\mathbb{C}^n$ with respect to a Gaussian measure. The transform itself can be described in terms of the heat operator for $\mathbb{R}^n$. In [25], I replace the $\mathbb{R}^n$ with a compact group $K$ and replace $\mathbb{C}^n$ with the “complexification” $K_\mathbb{C}$ of $K$. (For example, if $K = SU(n)$, the $K_\mathbb{C} = SL(n, \mathbb{C})$.) The Gaussian measures on $\mathbb{R}^n$ and $\mathbb{C}^n$ are then replaced by heat kernel measures and the heat operator on $\mathbb{R}^n$ is replaced by the heat operator on $K$. The resulting map is unitary from $L^2(K)$ (with respect to a heat kernel measure) onto the space of $L^2$ holomorphic functions on $K_\mathbb{C}$ (with respect to another heat kernel measure).

The paper [25] describes the Segal–Bargmann transform for $K$ in purely finite-dimensional terms. Nevertheless, the origin of the Hermite expansion for $K$ suggests that there could be a connection to the infinite-dimensional analysis in Gross’s paper [21]. Indeed, this turns out to be the case, as Gross demonstrated in a paper with P. Malliavin [24]. Gross and Malliavin show something similar to what we have for the Hermite expansion: The Segal–Bargmann transform for $K$ can be viewed as the Segal–Bargmann transform for an infinite-dimensional
linear space (the path space $W(k)$) applied to endpoint functions. That is, given a function $\phi$ on $K$, we consider as before the function $f$ on $W(k)$ given by $f(X) = \phi(\theta(X)(1))$. Then the Segal–Bargmann transform of $f$ is the function $F$ given by $F(Z) = \Phi(\theta_{\mathbb{C}}(Z)(1))$. Here $\theta_{\mathbb{C}}$ is the Itô map for $k_{\mathbb{C}}$ and $\Phi$ is the holomorphic function on $K_{\mathbb{C}}$ obtained by applying the Segal–Bargmann transform for $K$, in the sense of [25].

In [39], A. Sengupta and I extend the reasoning of [24] to more general functions of the Itô map. This eventually led to the development of a unitary Segal–Bargmann transform for the path group $W(K)$. (See also [6].) Furthermore, the paper [32] turns the reasoning in [24] around and uses the Segal–Bargmann transform for $K$ to give a new proof (one of several by now) of the Gross Ergodicity Theorem.

Meanwhile, in [12], Driver and I exploit the similarity between the action of $\mathcal{L}(K)$ on $W(\mathfrak{k})$ and gauge transformations to study the quantization of (1 + 1)-dimensional Yang–Mills theory. We develop a new form of the Segal–Bargmann transform for $K$ (see also [27]) and prove a generalization of the result of Gross and Malliavin in that setting. Our argument there draws on the analysis in [21], especially the use of the chaos expansion to understand endpoint functions. Our work in [12] was motivated by the work of Wren [60].

We have, then, three unitary maps that come out of Gross’s proof of Theorem 1.1. The maps are (1) the Hermite expansion, which maps functions on a compact group $K$ to the space $J_{\mathbb{C}}^\perp$, (2) the generalized Segal–Bargmann transform, mapping functions on $K$ to holomorphic functions on $K_{\mathbb{C}}$, and (3) the Taylor map, mapping a space of holomorphic functions on $K_{\mathbb{C}}$ to $J_{\mathbb{C}}^\perp$. The Taylor map can be defined as the composition of the Hermite expansion and the inverse Segal–Bargmann transform, but as shown in [9] it can be computed in a natural direct fashion in terms of the derivatives at the origin of a holomorphic function on $K_{\mathbb{C}}$.

2. Heat Kernel Analysis on Groups

The study of these maps and their generalizations has become a subject of its own, which may be called “heat kernel analysis on Lie groups.” This phrase should be understood not as analysis of heat kernel (though of course there is some of that involved) but rather as analysis with respect to a heat kernel measure. More specifically, heat kernel analysis means generalizing to the setting of Lie groups various results in analysis on Euclidean space that involve a Gaussian measure. In the Lie group setting, the Gaussian measure on $\mathbb{R}^n$ is replaced by a heat kernel measure on the group. The remainder of this article is devoted to explaining some of the developments over the last 15 years in this part of analysis. See also the papers [28, 30, 33] for more detailed exposition.

A major development in the study of the Taylor map is the observation by Driver and Gross that the proof of the isometricity of the Taylor map does not require any invariance of the inner product on $\mathfrak{k}_{\mathbb{C}}$. This observation opened the door for a development (in [10]) of a unitary Taylor expansion map in the setting of an arbitrary simply connected complex Lie group $G$. (That is, $G$ is no longer required to be the complexification of a compact group $K$.) It is worth noting that some vestige of infinite-dimensional analysis lurks in [10]. Driver and Gross
builds on a similar analysis in [9]) use the concept of “Taylor expansion along paths” to obtain some key estimates on the pointwise behavior of $L^2$ holomorphic functions on $G$. These estimates can be understood as follows: One can turn a holomorphic function $F$ on $G$ into an “endpoint function” on the path space over the Lie algebra $\mathfrak{g}$ of $G$. This endpoint function belongs to a Segal–Bargmann space over an infinite-dimensional linear space. The bounds on $F$ can be viewed as coming from applying the standard pointwise bounds on functions in that Segal–Bargmann space and minimizing over all paths with the same endpoint. Recently, Driver and Gross, along with L. Saloff-Coste, have extend the results of [10] to the case of a sub-Riemannian metric on $G$ (see [11]).

The results of Driver and Gross have been extended to give Taylor maps for (certain classes of) infinite-dimensional Lie groups by M. Gordina [18, 19, 20]. Gordina’s work allows for the development of a Cameron–Martin group along with a “skeleton map” or “restriction map,” generalizing results for the linear case. Specifically, Gordina considers a space of $L^2$ holomorphic functions on a certain sort of infinite-dimensional group $G$. She then constructs a Cameron–Martin subgroup $G_{CM}$ and shows that each $L^2$ holomorphic function on $G$ has a well-defined “restriction” to $G_{CM}$. This holds even though $G_{CM}$ is a set of measure zero within $G$. Further work in this direction has been done by M. Cecil and Driver [8, 7].

Meanwhile, concerning the Hermite expansion, J. Mitchell [48, 49, 50] has studied the asymptotics of the “Hermite functions” on a compact group $K$ as the time parameter tends to zero. The Hermite functions are defined as logarithmic-type derivatives of the heat kernel, and the limit as the time parameter tends to zero is equivalent to the limit as the metric on $K$ is scaled by a large constant, so that the curvature tends to zero. (Compare [35].) Mitchell shows that in this limit, the Hermite functions tend to the ordinary Hermite polynomials on $\mathbb{R}^n$. More generally, Mitchell develops an asymptotic expansion (in powers of $t$) of the Hermite functions on $K$, the coefficients of which are polynomials on $\mathbb{R}^n$.

The results of [25, 26] concerning the Segal–Bargmann transform for a compact group $K$ have been extended to the setting of compact symmetric spaces by M. Stenzel [54]. More recently, work has begun on understanding the Segal–Bargmann transform in the setting of noncompact symmetric spaces. Except in the Euclidean case, noncompact symmetric spaces present a major conceptual difficulty, because of singularities that do not occur in the compact or Euclidean cases. Any theorem in the noncompact setting has to provide a way of dealing with these singularities, typically by means of a “cancellation of singularities” result. Complementary approaches to this problem have been given B. Krötz, G. Ólafsson, and B. Stanton [44] on the one hand, and by Mitchell and me on the other hand [36, 37]. Comparing the two approaches, we may say that the approach of [44] is more general (it applies to arbitrary symmetric spaces of the noncompact type), whereas the approach of [36, 37] is more parallel to the compact and Euclidean cases, but applies (so far) only to noncompact symmetric spaces of the complex type. Both approaches draw on the work of Krötz and Stanton [45, 46] on analytic continuation of matrix entries of principal-series representations, as well as the Gutzmer–type formula of J. Faraut [14, 15]. See also [51, 52, 38, 55] for ongoing research in this direction.
Generalizing in a different way, Krötz, S. Thangavelu, and Y. Xu have developed a Segal–Bargmann transform (or “heat kernel transform”) for the Heisenberg group. This case turns out to be quite different from the $\mathbb{R}^n$ case, with the image being identified as the direct sum of two spaces, each of which is an $L^2$ space of holomorphic functions with respect to a signed measure.

In another direction, the Segal–Bargmann transform for compact groups, along with the associated “coherent states” [35] have been applied in several ways in the setting of loop quantum gravity, which is an alternative to string theory in the attempt to unify gravity with quantum mechanics. In [4], A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, and T. Thiemann use the generalized Segal–Bargmann transform to deal with certain “reality conditions” in the theory. Meanwhile, a series of papers by T. Thiemann and O. Winkler [56, 57, 58] have used the coherent states on $K$ that come from the Segal–Bargmann transform to investigate the classical limit of Thiemann’s “quantum spin dynamics” approach to loop quantum gravity.

Meanwhile, the paper [31] shows that the Segal–Bargmann transform for a compact group can be viewed as a unitary pairing map in the setting of geometric quantization. In light of [12], this can be seen as a “quantization commutes with reduction” result, as explained in [29]. Further understanding of the situation has been given by C. Florentino, P. Matias, J. Mourão, and J. Nunes [16, 17] and by J. Huebschmann [43]. Florentino and co-authors understand the heat equation on $K$ as coming from parallel transport in a Hilbert bundle associated to rescalings of the complex structure on $K^C$. Huebschmann, in turn, sees the heat equation as connected to a Peter–Weyl-type decomposition on $K^C$ and to the Kirillov character formula.

In the paper [34], W. Lewkeeratiyutkul and I develop a theory of “holomorphic Sobolev spaces,” that allows us to characterize the image of $C^\infty_0(K)$ under the generalized Segal–Bargmann transform. Thangavelu [59] has given a similar characterization of the image of distributions on $K$, and has extend both results to the setting of compact symmetric spaces. Thangavelu’s work draws on the very sharp estimates for heat kernels on noncompact symmetric spaces developed by J.-P. Anker and P. Ostellari [2, 3]; see also [40].

In conclusion, heat kernel analysis on Lie groups has become an active subject of its own. Although it has by now evolved far from its origins, the subject owes its existence to the work of Leonard Gross.

References


Brian C. Hall: Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556-4618, USA

E-mail address: bhall@nd.edu

URL: http://www.nd.edu/~bhall
QUADRATIC WIENER FUNCTIONALS OF SQUARE NORMS ON MEASURE SPACES

SETSUO TANIGUCHI

Abstract. Stochastic oscillatory integrals associated with quadratic Wiener functionals obtained as square norms on measure spaces of first order Wiener chaos is investigated. As applications, the square norm of the Brownian sheet and quadratic Wiener functional related to the KdV equation will be studied.

1. Introduction

Abstract Wiener spaces were introduced by L. Gross in 1965 [3], and since then, have been playing a key role in infinite dimensional stochastic analysis. In this paper, we investigate the quadratic Wiener functional on an abstract Wiener space \((X, H, \nu)\) of the form

\[
F = \frac{1}{2} \int_E (\nabla^* f_e)^2 \sigma(de),
\]

where \((E, \mathcal{E}, \sigma)\) is a \(\sigma\)-finite measure space, \(f_e \in H\) for every \(e \in E\), and \(\nabla^*\) stands for the adjoint operator of the Malliavin gradient \(\nabla\) on \(X\). For the precise definitions, see Sections 2 and 3. We shall give an exact infinite product expression of stochastic oscillatory integral

\[
\int_X e^{\zeta F + \nabla^* h} d\nu
\]

for sufficiently small \(\zeta \in C\) and any \(h \in H\). Moreover, it will be applied to the study of two quadratic Wiener functionals; the first one is

\[
\mathfrak{h} = \int_{[0,T]^2} W(s, t)^2 ds dt,
\]

the square norm of the Brownian sheet \(\{W(s, t)\}_{(s, t) \in [0, T]^2}\) on \([0, T]^2\), and the second is the quadratic Wiener functionals representing reflecting potentials, which plays a key role in the study of soliton solutions to the KdV equation (cf. [9, 10]). We also apply our result to show the corresponding Lévy-Khinchin formulas.

The investigation of such functionals as \(F\) goes back to the work of Cameron and Martin ([1]), who studied the functional

\[
\int_0^T W(s)^2 ds
\]

2000 Mathematics Subject Classification. Primary 60H30; Secondary 60H20.

Key words and phrases. quadratic Wiener functional, infinite product, first order Wiener chaos, Brownian sheet, KdV equation.

* This research is supported in part by Grant-in-Aid for Scientific Research (B) 18340038.
of the 1-dimensional Brownian motion \( \{W(s)\}_{s \geq 0} \) starting from the origin at time 0. The functional has a deep connection with the Schrödinger operator corresponding to harmonic oscillator; \((d/dx)^2 + x^2\). Recently, Deheuvels, Peccati, and Yor ([2]) studied the functional \( h \) with the help of Karhunen-Loève expansions and the result due to Cameron-Martin. In their paper, a reason why one is interested in \( h \) and hence why \( h \) is not a useless generalization of the functional studied by Cameron-Martin, can be found. Our result, which is based on the complex change of variable formula on the abstract Wiener space obtained via the Malliavin calculus ([5]), covers all their exact expressions. See Section 4.

A concrete bijective correspondence between reflectionless potentials and stochastic oscillatory integrals with Ornstein-Uhlenbeck processes in phase function was established in [4, 10]. In particular, in [10], quadratic Wiener functionals obtained as square norms of Wiener integrals with respect to the Brownian sheet played a fundamental role. In Section 5, such quadratic Wiener functionals will be studied as an example of quadratic Wiener functional of the form (1.1).

2. Preliminaries

In this section, we review several results on abstract Wiener spaces.

Let \((X, H, \nu)\) be a real abstract Wiener space, i.e., \(X\) is a real separable Banach space, \(H\) is a real Hilbert space with inner product \(\langle \cdot, \cdot \rangle_H\) and norm \(\| \cdot \|_H\), which is imbedded in \(X\) densely and continuously, and \(\nu\) is a Gaussian measure on \((X, \mathcal{B}(X))\), \(\mathcal{B}(X)\) being the Borel \(\sigma\)-field of \(X\), such that

\[
\int_X e^{\sqrt{-1} \langle \ell, \nu \rangle} d\nu = e^{-\|\ell\|^2_H/2}
\]

for any \(\ell \in X^*\), where \(X^*\) is the dual space of \(X\) and we have used the standard identification of \(H^*\) and \(H\) to have the inclusion that \(X^* \subset H^* = H \subset X\). For a separable Hilbert space \(K\), we say that a \(K\)-valued Wiener functional \(F\) belongs to \(D_\infty(K)\) if \(F\) is infinitely differentiable in the sense of the Malliavin calculus and it and its Malliavin derivatives of all orders are \(p\)-integrable with respect to \(\nu\) for any \(p > 1\) (cf. [8]). Denoting by \(H \otimes K\) the Hilbert space of Hilbert-Schmidt operators of \(H\) to \(K\), we define the adjoint operator \(\nabla^* : D_\infty(H \otimes K) \to D_\infty(K)\) of the Malliavin gradient \(\nabla\) by

\[
\int_X \langle \nabla^* F, G \rangle_K d\nu = \int_X \langle F, \nabla G \rangle_{H \otimes K} d\nu \quad \text{for any } F \in D_\infty(H \otimes K), G \in D_\infty(K).
\]

For a symmetric Hilbert-Schmidt operator \(U : H \to H\), we set

\[
Q_U = (\nabla^*)^2 U,
\]

where we have thought of \(U\) as a constant function defined on \(X\) with values in the Hilbert space \(H^{\otimes 2}\) of Hilbert-Schmidt operators on \(H\). If \(U\) is of trace class, we can define

\[
q_U = Q_U + \text{tr} U.
\]

It is easily checked that the third Malliavin derivative of functional \(F \in D_\infty(R)\) vanishes, i.e., \(\nabla^3 F = 0\), if and only if \(F\) admits an expression as

\[
F = \frac{1}{2} Q_U + \nabla^* h + c,
\]

(2.1)
where
\[ U = \nabla^2 F, \quad h = \int_X \nabla F \, d\nu, \quad \text{and} \quad c = \int_X F \, d\nu. \]

Applying the complex change of variables formula shown in [5], we obtain that

**Proposition 2.1.** Let \( U : H \to H \) be a symmetric Hilbert-Schmidt operator and \( h \in H \). For \( \zeta \in C \) with \( |\zeta| < 1/\|U\|_{\text{op}} \), where \( \| \cdot \|_{\text{op}} \) stands for the operator norm, it holds that

\[
\int_X e^{\zeta/2} Q_U + \eta \nabla^* h \, d\nu = \left\{ \det_2 (I - \zeta U) \right\}^{-1/2} e^{\eta^2/2 (I - \zeta U)^{-1} h, h} / 2
\]

for every \( \eta \in C \), where \( \det_2 \) is the Carleman-Fredholm determinant, and \( (\cdot, \cdot)_H \) was extended complex bi-linearly to the complexified Hilbert space \( H \otimes C \) of \( H \). If, in addition, \( U \) is of trace class, then

\[
\int_W e^{\zeta/2} q_U + \eta \nabla^* h \, d\mu = \left\{ \det (I - \zeta U) \right\}^{-1/2} e^{\eta^2/2 (I - \zeta U)^{-1} h, h} / 2
\]

where \( \det \) is the Fredholm determinant.

It is routine to extend the above identity to \( \zeta \)'s in much wider domain in \( C \) by holomorphic continuation.

We now recall the Lévy-Khinchin formulas of the distributions of \( Q_U \) and \( q_U \) given in [6]. Let \( \{a_n\}_{n=1}^\infty \) be a sequence of eigenvalues of \( U \) repeated according to multiplicity. Define

\[
f_U(x) = \begin{cases} 
\frac{1}{2} \sum_{n: x_n > 0} \frac{1}{|x|} \exp(-x/a_n), & \text{if } x \neq 0, \\
0, & \text{if } x = 0.
\end{cases}
\]

**Proposition 2.2.** For any \( \lambda \in R \), it holds that

\[
\int_X e^{\sqrt{-1} \lambda Q_U} / 2 \, d\nu = \exp \left( \int_R \{e^{\sqrt{-1} \lambda x} - 1 - \sqrt{-1} \lambda x \} f_U(x) \, dx \right).
\]  \( (2.2) \)

If, in addition, \( U \) is of trace class, then

\[
\int_X e^{\sqrt{-1} \lambda q_U} / 2 \, d\nu = \exp \left( \int_R \{e^{\sqrt{-1} \lambda x} - 1 \} f_U(x) \, dx \right).
\]  \( (2.3) \)

**Proof.** The identity (2.2) was shown in [6, Theorem 2]. To see (2.3), it suffices to note that \( \int_R x f_U(x) \, dx = \text{tr} U / 2 \). \( \square \)

### 3. Square Norm on Measure Space

Let \( (X, H, \nu) \) be a real abstract Wiener space, \( E \) a topological space, \( \mathcal{E} \) its Borel \( \sigma \)-field, and \( \sigma \) a \( \sigma \)-finite measure on \((E, \mathcal{E})\). Consider a continuous mapping \( E \ni e \mapsto f_e \in H \), where the topology of \( H \) is the strong one, i.e., comes from the norm. Assume that

\[
(A) \quad \int_E \| f_e \|_H^2 \sigma(de) < \infty.
\]
Every $h \in H$ is an element of $D^\infty(H)$, being thought of as a constant Wiener functional. The functional $\nabla^* h$ satisfies that
\[
\|\nabla^* h\|_{L^p(\nu)} = C_p \|h\|_H, \quad p > 0, \quad \text{where} \quad C_p = \left(\int_R \frac{|x|^p}{\sqrt{2\pi}} e^{-x^2/2} dx\right)^{1/p}, \quad (3.1)
\]
and $\|\cdot\|_{L^p(\nu)}$ denotes the norm of the $L^p$-space $L^p(\nu)$ associated with $\nu$. It follows from this identity that
\[
\|(\nabla^* f_e)^2 - (\nabla^* f_{e'})^2\|_{L^2(\nu)} \leq C_4^2 \|f_e - f_{e'}\|_H \|f_e + f_{e'}\|_H, \quad e, e' \in E.
\]
Thus the mapping $E \ni e \mapsto (\nabla^* f_e)^2 \in L^2(\nu)$ is strongly continuous, and hence strongly measurable. By virtue of Assumption (A) and (3.1),
\[
\int_E \|(\nabla^* f_e)^2\|_{L^2(\nu)} \sigma(de) \leq C_4^2 \int_E \|f_e\|^2_H \sigma(de) < \infty.
\]
Hence the function $E \ni e \mapsto (\nabla^* f_e)^2 \in L^2(\nu)$ is Bochner integrable. Define $F \in L^2(\nu)$ by the Bochner integral
\[
F = \frac{1}{2} \int_E (\nabla^* f_e)^2 \sigma(de).
\]
Due to Assumption (A), we can define $A : H \to H$ by
\[
\langle Ah, g \rangle_H = \int_E \langle f_e, h \rangle_H \langle f_e, g \rangle_H \sigma(de), \quad h, g \in H.
\]
It also follows from Assumption (A) that $A$ is a non-negative definite, symmetric Hilbert-Schmidt operator of trace class and
\[
\text{tr} A = \int_E \|f_e\|^2_H \sigma(de).
\]

**Proposition 3.1.** $F$ belongs to $D^\infty(R)$ and coincides with $q_A/2$.

**Proof.** Let $p > 2$. Every bounded and measurable $G : X \to R$ is in the dual space of $L^2(\nu)$, and due to (3.1), satisfies that
\[
\left| \int_X FGd\nu \right| \leq \frac{1}{2} \int_E \left| \int_X (\nabla^* f_e)^2 Gd\nu \right| \sigma(de) \leq C_4^2 \left( \int_E \|f_e\|^2_H \sigma(de) \right) \|G\|_{L^q(\nu)},
\]
where $q = p/(p - 1)$. By Assumption (A), this implies that $F \in L^p(\nu)$.

For the $H$-valued functional $(\nabla^* f_e)f_e$, by virtue of Assumption (A) and (3.1), we have that
\[
\|(\nabla^* f_e)f_e - (\nabla^* f_{e'})f_{e'}\|_{L^2(\nu; H)} \leq C_2 \|f_e\|_H + \|f_{e'}\|_H \|f_e - f_{e'}\|_H, \quad e, e' \in E,
\]
\[
\int_E \|(\nabla^* f_e)f_e\|_{L^2(\nu; H)} \sigma(de) \leq C_2 \int_E \|f_e\|^2_H \sigma(de) < \infty,
\]
where $\|\cdot\|_{L^p(\nu; H)}$ stands for the norm of $L^p(\nu; H)$, the space of $p$th integrable $H$-valued functionals with respect to $\nu$. Hence we can define $L^2(\nu; H)$-valued $F'$ by the Bochner integral
\[
F' = \int_E (\nabla^* f_e)f_e \sigma(de).
\]
The same argument as given in the previous paragraph implies that $F' \in L^p(\nu; H)$ for any $p > 1$. 
Let $G \in D^\infty(H)$ and $K \in D^\infty(H^\otimes 2)$, where $H^\otimes 2$ stands for the Hilbert space of Hilbert-Schmidt operators on $H$. Since $\nabla(\nabla^* f_e) = f_e$, it holds that

$$\int_X F \nabla^* G d\nu = \int_X (F', G)_{H^2} d\nu, \quad \int_X F (\nabla^*)^2 K d\nu = \int_X (A, K)_{H^\otimes 2} d\nu.$$ 

Thus $F \in D^\infty(R)$ and

$$\nabla F = \int_E (\nabla^* f_e) f_e \sigma(de), \quad \nabla^2 F = A, \quad \nabla^k F = 0, \ k \geq 3.$$  \hspace{1cm} (3.2)

By (3.1), we see that

$$\int_X F d\nu = \frac{1}{2} \int_E \left( \int_X (\nabla^* f_e)^2 d\nu \right) \sigma(de) = \frac{1}{2} \int_E \|f_e\|^2_{H^\sigma} \sigma(de) = \frac{1}{2} \text{tr} A.$$

Moreover, the expression of $\nabla F$ in (3.2) yields that

$$\int_X \nabla F d\nu = 0.$$ \hspace{1cm} \text{(2.1)}

Proposition 3.2. Let $\{\phi_n\}_{n=1}^\infty$ be an orthonormal basis of $H$ consisting of eigenvectors of $A$ and put

$$a_n = \int_E \langle f_e, \phi_n \rangle_H^2 \sigma(de), \quad n = 1, 2, \ldots$$

For sufficiently small $\zeta \in C$, it holds that

$$\int_X e^{\zeta F + \nabla^* h} d\nu = \left( \prod_{n=1}^\infty (1 - \zeta a_n) \right)^{-1/2} \exp \left( \frac{1}{2} \sum_{n=1}^\infty \frac{\langle h, \phi_n \rangle_H^2}{1 - \zeta a_n} \right)$$

for every $h \in H$.

Proof. By Proposition 2.1, for sufficiently small $\zeta \in C$, it holds that

$$\int_X e^{\zeta F + \nabla^* h} d\nu = \text{det}(I - \zeta A)^{-1/2} \exp((I - \zeta A)^{-1} h, h)_{H^2}/2).$$

Due to the eigenfunction expansion $A = \sum_{n=1}^\infty a_n \phi_n \otimes \phi_n$, where $\phi_n \otimes \phi_n$ denotes the Hilbert-Schmidt operator such that $(\phi_n \otimes \phi_n)(h) = \langle \phi_n, h \rangle_H \phi_n$, $h \in H$, we obtain that

$$\det(I - \zeta A) = \prod_{n=1}^\infty (1 - \zeta a_n), \quad \langle (I - \zeta A)^{-1} h, h \rangle_H = \sum_{n=1}^\infty \frac{\langle h, \phi_n \rangle_H^2}{1 - \zeta a_n},$$

which implies the the desired expression. \hspace{1cm} \Box

4. Brownian Sheet

In this section we investigate the stochastic oscillatory integrals associated with the square norm of Brownian sheet.
4.1. Abstract Wiener spaces. We introduce the abstract Wiener spaces associated with the Brownian sheet.

Let $T > 0$ and set $W = \{ w : [0, T]^2 \to \mathbf{R} | w \text{ is continuous and } w(t, 0) = w(0, t) = 0, t \in [0, T] \}$. Denote by $\mathcal{H}$ the set of all $h \in W$ of the form

$$h(s, t) = \int_{[0,s] \times [0,t]} h'(u, v)du dv, \quad (s, t) \in [0, T]^2$$

for some $h' \in L^2([0,T]^2)$ (resp. the space of all real square integrable functions on $[0,T]^2$ with respect to the Lebesgue measure). $W$ is a real separable Banach space with the norm

$$\|w\| = \sup_{(s,t)\in[0,T]^2} |w(s,t)|, \quad w \in W,$$

and $\mathcal{H}$ is a real separable Hilbert space with the inner product

$$\langle h, g \rangle_{\mathcal{H}} = \int_{[0,T]^2} h'(s, t)g'(s, t)ds dt, \quad h, g \in \mathcal{H}.$$ 

We denote by $\| \cdot \|_{\mathcal{H}}$ the associated norm of $\mathcal{H}$. For $(s, t) \in [0, T]^2$, define the coordinate function $W(s, t) : W \to \mathbf{R}$ by

$$W(s, t)(w) = W(s, t; w) = w(s, t), \quad w \in W.$$ 

There exists a unique probability measure $\mu$ on $W$ such that $\{ W(s, t) \}_{(s, t) \in [0, T]^2}$ is a continuous Gaussian field with mean 0 and covariance function

$$\int_W W(s, t)W(u, v)d\mu = (s \wedge u)(t \wedge v) \quad \text{for any } (s, t), (u, v) \in [0, T]^2. \quad (4.1)$$

It is easily seen that $(W, \mathcal{H}, \mu)$ is a real abstract Wiener space.

Put

$$W_0 = \{ w \in W | w(T, t) = w(t, T) = 0, t \in [0, T] \}, \quad \mathcal{H}_0 = \mathcal{H} \cap W_0.$$ 

$W_0$ is a real separable Banach space with the same norm as $W$, and $\mathcal{H}_0$ is a real separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_0}$ inherited from $\mathcal{H}$. Define $\pi : W \to W_0$ by

$$(\pi w)(s, t) = w(s, t) - s T w(T, t) - t T w(s, T) + \frac{s t T}{2} w(T, T), \quad w \in W.$$ 

Note that $\pi(\mathcal{H}) = \mathcal{H}_0$. Let $\mu_0$ be the induced measure of $\mu$ on $W_0$ through $\pi$;

$$\mu_0 = \mu \circ \pi^{-1}.$$ 

Write $B(s, t)$ for the restriction of $W(s, t)$ onto $W_0$ to emphasize working on $W_0$. Under $\mu_0$, $\{ B(s, t) \}_{(s, t) \in [0, T]^2}$ is a continuous Gaussian field with mean 0 and covariance function

$$\int_{W_0} B(s, t)B(u, v)d\mu_0 = \left\{ s \wedge u - \frac{s u}{T} \right\} \left\{ t \wedge v - \frac{t v}{T} \right\}, \quad (s, t), (u, v) \in [0, T]^2, \quad (4.2)$$

and $(W_0, \mathcal{H}_0, \mu_0)$ is an abstract Wiener space.

We introduce the third abstract Wiener space associated with Brownian sheet. Let $W' = \{ w \in W | w(s, T) = 0, s \in [0, T] \}$, $\mathcal{H}' = \mathcal{H} \cap W'$.
$W$ is a separable Banach space equipped with the norm of uniform convergence, and $H'$ is a separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{H'}$ inherited from $H$. Define $\pi' : W \to W'$ by

$$(\pi'w)(s, t) = w(s, t) - \frac{t}{T}w(s, T), \quad w \in W.$$  

Then $\pi'(H) = H'$. Set $\mu' = \mu \circ (\pi')^{-1}$.

Denote the restriction of $W(s, t)$ to $W'$ by $B'(s, t)$ to emphasize dealing with $W'$. Under $\mu'$, \{B'(s, t)\}_{(s, t) \in [0, T]^2}$ is a continuous Gaussian field with mean 0 and covariance function

$$\int_{W'} B'(s, t)B'(u, v)d\mu' = (s \wedge u) \left\{ t \wedge v - \frac{tv}{T} \right\}, \quad (s, t), (u, v) \in [0, T]^2, \quad (4.3)$$

and $(W', H', \mu')$ is an abstract Wiener space.

For the sake of simplicity of notation, in what follows, all Malliavin gradients and their adjoints on $W$, $W_0$, and $W'$ will be denoted by the same symbol $\nabla$ and $\nabla^*$, respectively.

4.2. Square norm of Brownian sheet. Define $h : W \to R$ by

$$h = \frac{1}{2} \int_{[0, T]^2} W(s, t)^2dsdt,$$

and denote by $h_0$ and $h'$ its restriction to $W_0$ and $W'$, respectively:

$$h_0 = \frac{1}{2} \int_{[0, T]^2} B(s, t)^2dsdt \quad \text{and} \quad h' = \frac{1}{2} \int_{[0, T]^2} B'(s, t)^2dsdt.$$  

Define $f_{s,t} \in W^*$ by $f_{s,t}(w) = w(s, t), \ w \in W$. Recall that on an abstract Wiener space $(X, H, \nu)$, it holds that $\nabla^*\ell = \ell$ for $\ell \in X^*$. Hence we have that

$$\mathfrak{h} = \frac{1}{2} \int_{[0, T]^2} (\nabla^*f_{s,t})^2dsdt.$$

Following the convention that we use the same symbol $\nabla$ and $\nabla^*$ for the Malliavin gradients and its adjoint operators on $W$, $W_0$, and $W'$, we denote the restrictions $f_{s,t}$ to $W_0$ and $W'$ by the same letter $f_{s,t}$. Then $f_{s,t}$ belong to $W_0^*$ and $(W')^*$, and we have the similar expression as above;

$$h_0 = \frac{1}{2} \int_{[0, T]^2} (\nabla^* f_{s,t})^2dsdt \quad \text{and} \quad h' = \frac{1}{2} \int_{[0, T]^2} (\nabla^* f_{s,t})^2dsdt.$$  

Define $A : H \to H$, $A_0 : H_0 \to H_0$, and $A' : H' \to H'$ by

$$(Ah)(s, t) = \int_{[0, a] \times [0, t]} \int_{[a, s] \times [v, T]} h(a, b)dbdv, \quad h \in H, \ (s, t) \in [0, T]^2$$

and

$$A_0 = \pi_0 \circ A \quad \text{and} \quad A' = \pi' \circ A.$$
Lemma 4.1. (i) \{\phi_{m,n}\}_{m,n\in \mathbb{Z}_+}, \{\psi_{m,n}\}_{m,n\in \mathbb{N}}, and \{\psi'_{m,n}\}_{m\in \mathbb{Z}_+,n\in \mathbb{N}} are orthonormal bases of \mathcal{H}, \mathcal{H}_0, and \mathcal{H}', respectively.

(ii) \(A, A_0,\) and \(A'\) admit the following eigenfunction expansions:

\[
A = \sum_{m,n\in \mathbb{Z}_+} \left( \frac{T}{(m+\frac{1}{2})\pi} \right)^2 \left( \frac{T}{(n+\frac{1}{2})\pi} \right)^2 \phi_{m,n} \otimes \phi_{m,n},
\]

\[
A_0 = \sum_{m,n\in \mathbb{N}} \left( \frac{T}{m\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \psi_{m,n} \otimes \psi_{m,n},
\]

\[
A' = \sum_{m\in \mathbb{Z}_+,n\in \mathbb{N}} \left( \frac{T}{(m+\frac{1}{2})\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \psi'_{m,n} \otimes \psi'_{m,n}.
\]

(iii) It holds that \(\mathfrak{h} = q_A/2, \mathfrak{h}_0 = q_{A_0}/2,\) and \(\mathfrak{h}' = q_{A'}/2.\)

Proof. (i) Obviously \(\phi_{m,n} \in \mathcal{H}, \psi_{m,n} \in \mathcal{H}_0,\) and \(\psi'_{m,n} \in \mathcal{H}_0.\) It is an elementary exercise of Fourier series to show that \{\phi_{m,n}\}_{m,n\in \mathbb{Z}_+}, \{\psi_{m,n}\}_{m,n\in \mathbb{N}}, and \{\psi'_{m,n}\}_{m\in \mathbb{Z}_+,n\in \mathbb{N}} are orthonormal bases of \(\mathcal{H}, \mathcal{H}_0,\) and \(\mathcal{H}',\) respectively, once one notices that \(h \in \mathcal{H}_0\) and \(g \in \mathcal{H}'\) satisfy that

\[
\int_0^T h'(s,t)ds = 0 \quad \text{a.e. } t, \quad \int_0^T h'(s,t)dt = \int_0^T g'(s,t)dt = 0 \quad \text{a.e. } s.
\]

(ii) It is easily checked that

\[
A\phi_{m,n} = \left( \frac{T}{(m+\frac{1}{2})\pi} \right)^2 \left( \frac{T}{(n+\frac{1}{2})\pi} \right)^2 \phi_{m,n}, \quad A_0\psi_{m,n} = \left( \frac{T}{m\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \psi_{m,n},
\]

\[
A'\psi'_{m,n} = \left( \frac{T}{m\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \psi'_{m,n}.
\]

In conjunction with (i), these identities imply the desired expansions.

(iii) This is an immediate consequence of Proposition 3.1 and the integration by parts formula on \([0,T].\) \(\square\)

The first goal of this section is

Theorem 4.2. For sufficiently small \(\zeta \in \mathcal{C},\) the following identities hold.

\[
\int_{\mathcal{W}} e^{C\mathfrak{h}} d\mu = C(T^2\zeta)^{-1/2}, \quad \int_{\mathcal{W}_0} e^{C\mathfrak{h}_0} d\mu_0 = S(T^2\zeta)^{-1/2},
\]

\[
\int_{\mathcal{W}'} e^{C\mathfrak{h}'} d\mu' = \tilde{C}(T^2\zeta)^{-1/2}, \quad (4.4)
\]
where
\[ C(x) = \prod_{m=0}^{\infty} \cos \left( \frac{2x}{(2m+1)\pi} \right), \quad S(x) = \prod_{n=1}^{\infty} \frac{n\pi \sin(x/(n\pi))}{x}, \]
\[ \tilde{C}(x) = \prod_{n=1}^{\infty} \cos \left( \frac{x}{n\pi} \right). \]

Proof. By Lemma 4.1 (ii) and the well known identities that
\[ \prod_{n=0}^{\infty} \left( 1 - \frac{x^2}{(2n+1)^2} \right) = \cos \left( \frac{\pi x}{2} \right), \quad \prod_{m=1}^{\infty} \left( 1 - \frac{x^2}{m^2} \right) = \frac{\sin(\pi x)}{\pi x} \]
we see that
\[ \det(I - \zeta^2 A) = \prod_{m=0}^{\infty} \prod_{n=0}^{\infty} \left\{ 1 - \zeta^2 \left( \frac{T}{(m + \frac{1}{2})\pi} \right)^2 \left( \frac{T}{(n + \frac{1}{2})\pi} \right)^2 \right\} = C(T^2 \zeta), \]
\[ \det(I - \zeta^2 A_0) = \prod_{m,n=1}^{\infty} \left\{ 1 - \zeta^2 \left( \frac{T}{m\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \right\} = S(T^2 \zeta), \]
\[ \det(I - \zeta^2 A') = \prod_{m=0}^{\infty} \prod_{n=1}^{\infty} \left\{ 1 - \zeta^2 \left( \frac{T}{(m + \frac{1}{2})\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \right\} = \tilde{C}(T^2 \zeta). \]
In conjunction with Lemma 4.1 (iii) and Proposition 3.2, these implies the identities described in (4.4). □

As another application of Lemma 4.1, we obtain the following Lévy-Khinchin formulas.

**Proposition 4.3.** For \( \lambda \in \mathbb{R} \), it holds that
\[
\int_{W} e^{\sqrt{-1}\lambda h} d\mu = \exp \left( \int_{(0,\infty)} \{ e^{\sqrt{-1}\lambda x} - 1 \} \times \right.
\left. \frac{1}{4x} \sum_{m=0}^{\infty} \left\{ \Theta \left( \frac{(2m+1)^2\pi^4 x}{16T^4} \right) - \Theta \left( \frac{(2m+1)^2\pi^4 x}{4T^4} \right) \right\} dx \right),
\]
\[
\int_{W_0} e^{\sqrt{-1}\lambda_0 h} d\mu_0 = \exp \left( \int_{(0,\infty)} \{ e^{\sqrt{-1}\lambda x} - 1 \} \frac{1}{4x} \sum_{n=1}^{\infty} \left\{ \Theta \left( \frac{n^2\pi^4 x}{T^4} \right) - 1 \right\} dx \right),
\]
\[
\int_{W'} e^{\sqrt{-1}\lambda h'} d\mu' = \exp \left( \int_{(0,\infty)} \{ e^{\sqrt{-1}\lambda x} - 1 \} \frac{1}{4x} \sum_{m=0}^{\infty} \left\{ \Theta \left( \frac{(2m+1)^2\pi^4 x}{4T^4} \right) - 1 \right\} dx \right),
\]
where \( \Theta(u) \) is Jacobi’s Theta function;
\[ \Theta(u) = \sum_{n=-\infty}^{\infty} e^{-n^2 u}, \quad u \in \mathbb{R}. \]
Proof. By Lemma 4.1 (ii) and the very definition, \( f_A(x) = f_{A_0}(x) = f_{A'}(x) = 0 \) if \( x \leq 0 \). For \( x > 0 \), by Lemma 4.1 again, we have that
\[
\begin{align*}
 f_A(x) &= \frac{1}{2} \sum_{m,n=0}^{\infty} \frac{1}{x} \exp \left( -x \left( \frac{T}{m + \frac{1}{2}} \pi \right)^2 \left( \frac{T}{n + \frac{1}{2}} \pi \right)^2 \right) \\
 &= \frac{1}{4x} \sum_{m=0}^{\infty} \left\{ \Theta \left( \frac{(2m+1)^2\pi^4x}{16T^4} \right) - \Theta \left( \frac{(2m+1)^2\pi^4x}{4T^4} \right) \right\}, \\
 f_{A_0}(x) &= \frac{1}{2} \sum_{m,n=1}^{\infty} \frac{1}{x} \exp \left( -x \left( \frac{T}{m\pi} \right)^2 \left( \frac{T}{n\pi} \right)^2 \right) = \frac{1}{4x} \sum_{m=1}^{\infty} \left\{ \Theta \left( \frac{m^2\pi^4x}{4T^4} \right) - 1 \right\}, \\
 f_{A'}(x) &= \frac{1}{2} \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{1}{x} \exp \left[ -x \left( \frac{T}{m + \frac{1}{2}} \pi \right)^2 \left( \frac{T}{n\pi} \right)^2 \right] \\
 &= \frac{1}{4x} \sum_{m=1}^{\infty} \left\{ \Theta \left( \frac{(2m+1)^2\pi^4x}{4T^4} \right) - 1 \right\}.
\end{align*}
\]
Combined with Proposition 2.2, these imply the desired identities. \( \square \)

We next consider an application to conditional expectations of \( e^{\zeta h} \).

**Theorem 4.4.** (i) Let \( x \in \mathbb{R} \). For sufficiently small \( \zeta \in C \), it holds that
\[
E[e^{\zeta h} | W(T, T) = x] = \left\{ C(T^2 \zeta) T(T^2 \zeta) \right\}^{-1/2} \exp \left( -\frac{x^2}{2T^2} \left\{ \frac{1}{T(T^2 \zeta)} - 1 \right\} \right),
\]
where
\[
T(a) = \frac{4}{a} \sum_{m=0}^{\infty} \frac{1}{(2m+1)\pi} \tan \left( \frac{2a}{(2m+1)\pi} \right).
\]
(ii) Let \( \alpha, \beta \in C([0, T]) \) satisfy \( \alpha(0) = \beta(0) = 0 \) and \( \alpha(T) = \beta(T) \). For sufficiently small \( \zeta \in C \), it holds that
\[
E_{\alpha}[\exp(\zeta h)|W(\cdot, T) = \alpha, W(T, \cdot) = \beta]
= S(T^2 \zeta)^{-1/2} \exp \left( \frac{\zeta^2}{2} I(\alpha, \beta) + 2\zeta^4 T^4 \sum_{m,n=1}^{\infty} \frac{K^2_{m,n}}{m^2n^2\pi^4 - \zeta^2 T^4} \right),
\]
where
\[
I(\alpha, \beta) = \frac{T}{3} \int_0^T \{ \alpha(u)^2 + \beta(u)^2 \} du + \frac{T^2}{9} \alpha(T)^2 + \frac{2}{T^2} \int_0^T s\alpha(s)ds \int_0^T t\beta(t)dt - \frac{2\alpha(T)}{3} \int_0^T \{ u\alpha(u) + u\beta(u) \} du,
\]
\[
\alpha_m = \int_0^T \alpha(s) \sin \left( \frac{m\pi s}{T} \right) ds, \quad \beta_n = \int_0^T \beta(t) \sin \left( \frac{n\pi t}{T} \right) dt,
\]
\[
K_{m,n} = \frac{(-1)^{m+1}\beta_n}{m\pi} + \frac{(-1)^{n+1}\alpha_m}{n\pi} = \frac{(-1)^{m+n}T\alpha(T)}{m\pi},
\]
\[
\sum_{m,n=1}^{\infty} \frac{K^2_{m,n}}{m^2n^2\pi^4 - \zeta^2 T^4}.
\]
(iii) Suppose that $\gamma \in C([0,T])$ satisfies that $\gamma(0) = 0$. For sufficiently small $\zeta \in C$, it holds that

$$E_{\mu}[\exp(\zeta^2 h)|W(\cdot, T) = \gamma] = \mathcal{C}(T^2\zeta)^{-1/2} \exp\left(\frac{\zeta^2 T I(\gamma)}{6} + 2\zeta^4 T^4 \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{\gamma_m^2}{(m+\frac{1}{2})^2 n^2 \pi^4 - \zeta^2 T^4 \frac{1}{n^2 \pi^2}}\right),$$

where

$$I(\gamma) = \int_0^T \gamma(s)^2 ds, \quad \gamma_m = \int_0^T \gamma(s) \sin\left(\frac{(m+\frac{1}{2})\pi s}{T}\right) ds.$$ 

Remark 4.5. The expression in the second assertion is different from that given in [2]. Their formula seems incomplete because the right hand side of the identity given at the bottom of [2, p.526], on which their formula is based, has no boundary values $y(\cdot, 1)$ and $y(1, \cdot)$.

Proof. (i) Since $W(T, T)$ obeys the normal distribution with mean 0 and variance $T^2$, it holds that

$$E[e^{\xi^2 h} | W(T, T) = w] = \sqrt{2\pi T^2} e^{\xi^2 / (2T^2)} \int_W e^{\xi h \delta_x(W(T, T))} d\mu,$$

where $\delta_x(W(T, T))$ is Watanabe’s pull-back of the Dirac measure $\delta_x$ concentrated at $x \in \mathbb{R}$ through the smooth and non-degenerate Wiener functional $W(T, T) : W \to \mathbb{R}$. Thus it suffices to show that

$$\int_W e^{\xi^2 h \delta_x(W(T, T))} d\mu = \{2\pi T^2 C(T^2 \zeta) T(T^2 \zeta)\}^{-1/2} \exp\left(-\frac{x^2}{2T^2(T^2 \zeta)}\right). \quad (4.5)$$

To see (4.5), define $\ell \in W^*$ by $\ell = \pi T; \ell(w) = w(T, T), w \in W$. Then $\nabla^* \ell = W(T, T)$. By Proposition 3.2, Lemma 4.1, and Theorem 4.2, we have that

$$\int_W e^{\xi^2 h \delta_x(W(T, T))} d\mu = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\sqrt{-1} \eta x} \left(\int_W e^{\xi^2 h + \sqrt{-1} \eta \nabla^* \ell} d\mu\right) d\eta$$

$$= C(T^2 \zeta)^{-1/2} \{2\pi \{(I - \zeta^2 A)^{-1} \ell, \ell\}_\mathcal{H}\}^{-1/2} \exp\left(-\frac{x^2}{2\pi \{(I - \zeta^2 A)^{-1} \ell, \ell\}_\mathcal{H}}\right).$$

By Lemma 4.1 (ii), we see that

$$\{(I - \zeta^2 A)^{-1} \ell, \ell\}_\mathcal{H} = \sum_{m, n \in \mathbb{Z}_+} \left\{1 - \zeta^2 \left(\frac{T}{(m+\frac{1}{2})\pi}\right)^2 \left(\frac{T}{(n+\frac{1}{2})\pi}\right)^2\right\}^{-1} \phi_{m,n}(T, T)^2$$

$$= \sum_{m, n \in \mathbb{Z}_+} \frac{4T^2}{(m+\frac{1}{2})^2(n+\frac{1}{2})^2 \pi^4 - \zeta^2 T^4}$$

$$= \frac{4}{\zeta} \sum_{m \in \mathbb{Z}_+} \frac{1}{(2m+1)\pi} \tan\left(\frac{2\zeta T^2}{(2m+1)\pi}\right) = T^2 T(T^2 \zeta).$$

Thus (4.5) holds.

(ii) Since $\pi W$ and $(I - \pi)W$ are independent, for $y \in (I - \pi)(W)$, the conditional
distribution of $W$ under $\mu$ given the condition that $(I - \pi)W = y$ coincides with the distribution of $B + y$ under $\mu_0$;

$$\mu(W \in \bullet | (I - \pi)W = y) = \mu_0(B + y \in \bullet).$$

This implies that

$$\mu(W \in \bullet | W(\cdot, T) = y(\cdot, T), W(T, \cdot) = y(T, \cdot)) = \mu_0(B + y \in \bullet).$$

Define $y \in (I - \pi)W$ and $\ell \in W_0^*$ by

$$y(s, t) = \frac{s}{T} \beta(t) + \frac{t}{T} \alpha(s) - \frac{st}{T^2} \alpha(T), \quad (s, t) \in [0, T]^2,$$

$$\ell(w) = \int_{[0,T]^2} w(s, t)y(s, t)dsdt, \quad w \in W_0.$$

By Proposition 2.1 and Theorem 4.2, we have that

$$E_\mu[\exp(\zeta^2 h)]W(\cdot, T) = \alpha, W(T, \cdot) = \beta$$

$$= \exp \left( \frac{\zeta^2}{2} \int_{[0,T]^2} y(s, t)^2dsdt \right) \int_{W_0} \exp(\zeta^2 h_0 + \zeta^2 \ell) d\mu_0$$

$$= \exp \left( \frac{\zeta^2}{2} \int_{[0,T]^2} y(s, t)^2dsdt \right) S(T^2\zeta)^{-1/2} \exp(\zeta^4((I - \zeta^2 A_0)^{-1}\ell, \ell)_{H_0}/2).$$

It holds that $\int_{[0,T]^2} y(s, t)^2dsdt = I(\alpha, \beta)$. By virtue of Lemma 4.1 (ii), we see that

$$\langle (I - \zeta^2 A_0)^{-1}\ell, \ell \rangle_{H_0} = \sum_{m,n=1}^{\infty} \frac{m^2 n^2 \pi^4}{m^2 n^2 \pi^4 - \zeta^2 T^4} \left\{ \int_{[0,T]^2} \psi_{m,n}(s, t)y(s, t)dsdt \right\}^2.$$

A direct computation yields that

$$\int_{[0,T]^2} \psi_{m,n}(s, t)y(s, t)dsdt = \frac{2T^2}{mn\pi^2} K_{m,n}.$$

Thus the desired identity follows.

(iii) Taking the advantage of the independence of $\pi'W$ and $(I - \pi')W$, we see that

$$\mu(W \in \bullet | W(\cdot, T) = \gamma) = \mu'(B' + z \in \bullet),$$

where $z \in W$ is given by

$$z(s, t) = \frac{t}{T} \gamma(s), \quad (s, t) \in [0, T]^2.$$

Define $\ell \in (W')^*$ by

$$\ell(w) = \int_{[0,T]^2} w(s, t)z(s, t)dsdt, \quad w \in W'.$$

Then, due to Proposition 2.1 and Theorem 4.2, we have that

$$E_\mu[\exp(\zeta^2 h)]W(\cdot, T) = \gamma$$

$$= \exp \left( \frac{\zeta^2}{2} \int_{[0,T]^2} z(s, t)^2dsdt \right) \tilde{C}(T^2\zeta)^{-1/2} \exp(\zeta^4((I - \zeta^2 A')^{-1}\ell, \ell)_{W'/2}).$$
It follows from Lemma 4.1 that
\[
\int_{[0,T]^2} z(s,t)^2 ds dt = \frac{T}{3} \int_0^T \gamma(s)^2 ds,
\]
and
\[
\langle (I - \zeta^2 A')^{-1} \ell, \ell \rangle_{\mathcal{H}} = \sum_{n=0}^{\infty} \sum_{n=1}^{\infty} \frac{(m + \frac{1}{2})^2 n^2 \pi^4}{(m + \frac{1}{2})^2 n^2 \pi^4 - \zeta^2 T^4} \langle \psi'_{m,n}, \ell \rangle_{\mathcal{H}}^2,
\]
and
\[
\langle \psi'_{m,n}, \ell \rangle_{\mathcal{H}} = \frac{2T^2}{(m + \frac{1}{2})n\pi^2} \frac{\gamma_m(-1)^{n+1}}{n\pi}.
\]
Hence the desired identity follows. \( \square \)

5. Reflectionless Potential

In this section, we continue to work on the abstract Wiener space \((\mathcal{W}, \mathcal{H}, \mu)\) associated with the Brownian sheet.

A reflectionless potential \( u : \mathbb{R} \to \mathbb{R} \) with scattering data \( \{\eta_i, m_i\}_{1 \leq i \leq n} \), where \( \eta_i, m_i > 0 \), \( 1 \leq i \leq n \), and \( \eta_i \neq \eta_j \) if \( i \neq j \), is the function defined by
\[
u(x) = -2 \frac{d^2}{dx^2} \log \det(I + G(x)),
\]
where
\[
G(x) = \left(\sqrt[m_i]{m_j} e^{-(\eta_i + \eta_j)x} / \eta_i + \eta_j\right)_{1 \leq i,j \leq n}.
\]
It is widely known that \( u(x,t) = -q(x,t) \), \( q \) being the reflectionless potential with scattering data \( \{\eta_i, m_i \exp[-2\eta_i^2 t]\}_{1 \leq i \leq n} \) gives rise of an \( n \)-soliton solution of the KdV equation
\[
\frac{\partial v}{\partial t} = 3 \frac{\partial v}{\partial x} + \frac{1}{4} \frac{\partial^3 v}{\partial x^3}, \quad (5.1)
\]
For examples, see [7].

In [10], using the Brownian sheet, quadratic Wiener functionals related to reflectionless potentials are studied. More precisely, let \( \{(p_j, c_j)\}_{j=1}^n \subset (\mathbb{R}^2)^n \) satisfy that \( p_i \neq p_j \) if \( i \neq j \) and \( c_j > 0 \), \( 1 \leq j \leq n \). Define \( q_j \), \( 0 \leq j \leq n \) by
\[
q_0 = 0, \quad q_j = \sum_{k=1}^{j} |p_k - p_{k-1}|,
\]
where \( p_0 = -|p_1| - 1 \). Assume \( q_n \leq T \). Define \( e_1, \ldots, e_n \in L^2[0,T] \), the space of square integrable functions on \([0,T]\) with respect to the Lebesgue measure, by
\[
e_j(s) = \frac{1}{\sqrt{q_j - q_{j-1}}} \chi_{[q_{j-1}, q_j)}(s), \quad s \in [0,T],
\]
where \( \chi_{[a,b)} \) stands for the indicator function of \([a,b)\). Let \( x \leq T \). For \( y \in [0,x] \), define \( f_y \in \mathcal{H} \) by
\[
f'_y(s,t) = \sum_{j=1}^{n} c_j e_j(s) e^{(y-t)p_j} \chi_{[0,y)}(t), \quad (s,t) \in [0,T]^2.
\]
The function
\[ u(x) = 4 \frac{d^2}{dx^2} \log \left( \int_W \exp \left( -\frac{1}{2} \int_0^x (\nabla^* f_y)^2 dy \right) d\mu \right) \]
determines a reflectionless potential, and conversely, every reflectionless potential
admits such an expression ([10, 4]).

Define the Hilbert-Schmidt operator \( A : \mathcal{H} \to \mathcal{H} \) by
\[ A = \nabla^2 F = \int_0^x f_y \otimes f_y dy, \text{ where } F = \frac{1}{2} \int_0^x (\nabla^* f_y)^2 dy \]
(cf. Proposition 3.1). The aim of this section is to specify eigenvalues and eigenvectors of \( A \).

5.1. \( \ker A \). Let \( \mathcal{K} \) be the subspace of \( \mathcal{H} \) consisting of all \( h \in \mathcal{H} \) of the form
\[ h'(s, t) = \sum_{i=1}^n c_i(s) h_i(t), \quad (s, t) \in [0, T]^2 \quad (5.2) \]
for some \( h_1, \ldots, h_n \in L^2[0, T] \). Denote by \( \mathcal{P} \) the set of all \( h \in \mathcal{K} \) satisfying that
\[ h_i = g_i - p_i \int_0^x g_i(s) ds \quad \text{a.e. on } [0, x], \quad 1 \leq i \leq n, \quad (5.3) \]
for some \( g_1, \ldots, g_n \in L^2[0, T] \) with \( \sum_{i=1}^n c_i g_i = 0 \) a.e. on \([0, x]\). Let \( \mathcal{G} \) be the space of all \( h \in \mathcal{H} \) such that \( h'(s, t) = u(s)v(t) \) for some \( u, v \in L^2[0, T] \) with \( \int_0^T u(s)e_i(s) ds = 0, \quad 1 \leq i \leq n \). We shall show that
\[ \ker A = \overline{\mathcal{G}} \oplus \mathcal{P}, \quad (5.4) \]
where \( \overline{\mathcal{G}} \) is the closure of \( \mathcal{G} \).

Notice that
\[ \mathcal{K} = \mathcal{G}^\perp \quad \text{and} \quad (\ker A)^\perp \subset \mathcal{K}, \quad (5.5) \]
where \( A^\perp \) stands for the orthogonal complement of \( A \) in \( \mathcal{H} \). Hence, the proof of (5.4) completes once one has shown that
\[ \mathcal{P} = \ker A \cap \mathcal{K}. \quad (5.6) \]

To show (5.6), we first give an expression of \( Ah \) for \( h \in \mathcal{K} \). Let \( h \in \mathcal{K} \) and represent it as in (5.2). Putting
\[ \xi^i(y; h) = e^{yp} \int_0^y e^{-zp} h_i(z) dz, \quad i = 1, \ldots, n, \]
we see that \( (f_y, h)_\mathcal{H} = \sum_{i=1}^n c_i \xi^i(y; h) \), and hence that, for \( (s, t) \in [0, T]^2 \),
\[ (Ah)'(s, t) = \sum_{j=1}^n \left( e^{-tp_j} \int_t^x e^{qp_j} c_j \left( \sum_{i=1}^n c_i \xi^i(y; h) \right) dy \right) e_j(s) \chi_{[0, x]}(t). \quad (5.7) \]

Let \( h \in \mathcal{K} \) and represent it as in (5.2). By the very definition, it holds that
\[ (\xi^i)'(y; h) = h_i(y) + p_i \xi^i(y; h). \quad (5.8) \]
Suppose first that \( h \in \mathcal{P} \). Then, substituting the expression (5.3) into (5.8), we obtain that
\[
\xi^i(y; h) = \int_0^y g_i(z)dz, \quad y \in [0, x].
\]
Since \( \sum_{i=1}^n c_i g_i = 0 \) a.e. on \([0, x]\),
\[
\sum_{i=1}^n c_i \xi^i(y; h) = 0, \quad y \in [0, x].
\] (5.9)

Plugging this into (5.7), we see that \( h \in \ker A \). Next suppose that \( h \in \ker A \).
Since \( e_j(s) \neq 0 \) if and only if \( s \in [q_j-1, q_j) \), it follows from (5.7) that (5.9) holds.
Moreover, due to (5.8), we have that
\[
h_i(y) = (\xi^i)'(y; h) - p_i \int_0^y (\xi^i)'(z; h)dz, \quad a.e. \ y \leq x,
\]
which means that \( h \in \mathcal{P} \). Thus (5.6) has been verified.

5.2. Non-zero eigenvalues. Let \( \lambda \neq 0 \). By (5.5) and (5.7), \( \lambda \) is an eigenvalue of \( A \) if and only if there exists \( h \in \mathcal{K} \setminus \{0\} \) such that, under the representation as in (5.2), it holds that
\[
\lambda \begin{pmatrix} h_1(t) \\ \vdots \\ h_n(t) \end{pmatrix} = e^{-tD} \int_t^x e^{uD} (c \otimes c) \xi(y; h)dy \chi_{[0,x]}(t), \quad t \in [0, T],
\] (5.10)

where \( D \) is the diagonal matrix with elements \( p_1, \ldots, p_n \), \( c \otimes c \) denotes the \( n \times n \) matrix \((c_i c_j)_{1 \leq i,j \leq n}\), and
\[
\xi(y; h) = \begin{pmatrix} \xi^1(y; h) \\ \vdots \\ \xi^n(y; h) \end{pmatrix}.
\]

The identity (5.10) implies immediately that
\[
h_i(t) = 0, \quad t > x, \ i = 1, \ldots, n.
\]
By virtue of (5.8), the identity (5.10) reads as
\[
\lambda e^{tD} [\xi'(t; h) - D\xi(t; h)] = \int_t^x e^{uD} (c \otimes c) \xi(y; h)dy, \quad t \leq x.
\]
This is equivalent to the ordinary differential equation that
\[
\xi'' - B(1/\lambda) \xi = 0 \quad \text{with } \xi(0) = 0 \text{ and } \xi'(x) - D\xi(x) = 0,
\]
where \( B(a) = D^2 - a(c \otimes c) \).
Let \( s(t; \lambda) = \sinh(tB(1/\lambda)^{1/2}) \) and \( c(t; \lambda) = \cosh(tB(1/\lambda)^{1/2}) \). The above ordinary differential equation has a solution \( \xi(t; h) \) if and only if there exists a \( u \in \mathbb{R}^n \) such that
\[
\{c(x; \lambda) - Ds(x; \lambda)\}u = 0,
\]
and then \( \xi(t; h) = s(t; \lambda)u, \ t \leq x \). Moreover, \( h \neq 0 \) if and only if \( u \neq 0 \).
5.3. Eigenvalues and eigenvectors. Summing up the above observations, we have that

**Proposition 5.1.** (i) \( \ker A = \mathcal{G} \oplus \mathcal{P} \).

(ii) \( \lambda \in \mathbb{R} \setminus \{0\} \) is an eigenvalue of \( A \) if and only if

\[
\det \left[ c(x; \lambda) - Ds(x; \lambda) \right] = 0.
\]

Moreover, in this case, the corresponding eigenvector \( h \in \mathcal{H} \) is of the form

\[
h'(s, t) = \sum_{i=1}^{n} c_i(s) h_j(t),
\]

with

\[
\begin{pmatrix}
h_1(t) \\
h_2(t) \\
\vdots \\
h_n(t)
\end{pmatrix} = \begin{cases} 
\{c(t; \lambda) - Ds(t; \lambda)\} u, & t \leq x, \\
0, & t > x,
\end{cases}
\]

for some \( u \in \ker [c(x; \lambda) - Ds(x; \lambda)] \setminus \{0\} \).

**References**

COMPLEX HERMITE POLYNOMIALS: FROM THE SEMI-CIRCULAR LAW TO THE CIRCULAR LAW

MICHEL LEDOUX

Abstract. We study asymptotics of orthogonal polynomial measures of the form $|H_N|^2 d\gamma$ where $H_N$ are real or complex Hermite polynomials with respect to the Gaussian measure $\gamma$. By means of differential equations on Laplace transforms, interpolation between the (real) arcsine law and the (complex) uniform distribution on the circle is emphasized. Suitable averages by an independent uniform law give rise to the limiting semi-circular and circular laws of Hermitian and non-Hermitian Gaussian random matrix models. The intermediate regime between strong and weak non-Hermiticity is clearly identified on the limiting differential equation by means of an additional normal variable in the vertical direction.

1. Introduction

Let $A^N$ be a $N \times N$ random Hermitian matrix from the Gaussian Unitary Ensemble (GUE)

$$\mathbb{P}(dX) = \frac{1}{Z} \exp \left( - \text{Tr}(X^2)/2 \right) dX$$  \hspace{1cm} (1.1)

where $dX$ is Lebesgue measure on the space of $N \times N$ Hermitian matrices $X$ and $Z$ the normalization constant. Equivalently, the entries $A^N_{kl}, 1 \leq k \leq l \leq N,$ of $A^N$ are independent complex Gaussian variables with mean zero and variance one. It is a classical result going back to E. Wigner [10] that the empirical measure $\frac{1}{N} \sum_{k=1}^{N} \delta_{\lambda^N_k}$ on the (real) eigenvalues $\lambda^N_1, \ldots, \lambda^N_N$ of $A^N$ converges as $N \to \infty$ to the semi-circular law $\frac{2}{\pi} \frac{1}{(1-x^2)^{1/2}} \mathbf{1}_{\{|x|<1\}} dx$.

Consider now an independent copy $B^N$ of $A^N$, and form the random matrix $A^N + iB^N$ whose entries are independent complex Gaussian variables with mean zero and variance two. This is a canonical non-Hermitian ensemble of random matrix theory widely referred to as the (complex) Ginibre Ensemble. Girko’s classical theorem [3] indicates that the empirical measure on the (complex) eigenvalues of $(A^N+iB^N)/\sqrt{2N}$ converges as $N \to \infty$ towards the circular law $\frac{1}{\pi} \mathbf{1}_{\{a^2+x^2+y^2 \leq 1\}} dxdy$ on the plane.

Interpolation from the Ginibre Ensemble to the GUE is provided by the family $A^N + i\rho B^N$ for some real parameter $\rho, |\rho| \leq 1$, yielding as limiting spectral measure Girko’s elliptic law $\frac{1}{\pi a b} \mathbf{1}_{\{(x^2/a^2) + (y^2/b^2) \leq 1\}} dxdy$ where $a^2 = \frac{2}{1+\rho^2}, b^2 = \frac{2\rho^2}{1+\rho^2}$.

2000 Mathematics Subject Classification. Primary 15A52, 33A65; Secondary 60.

Key words and phrases. random matrices, Hermite polynomials, semi-circular law, circular law.
These results extend to families of random matrices with non-Gaussian entries (cf. [10, 1, 3, ...]). In the Gaussian case, the determinantal structure of the joint law of the eigenvalues (cf. [8]) allows, by the so-called orthogonal polynomial method, for the analysis of the mean spectral measure through the Hermite polynomials. Following [9], let \( H_\ell \), \( \ell \in \mathbb{N} \), be the Hermite polynomials defined by the generating series
\[
e^{\lambda x - \lambda^2/2} = \sum_{\ell=0}^{\infty} \frac{\lambda^\ell}{\sqrt{\ell!}} H_\ell(x), \quad \lambda \in \mathbb{R}, \; x \in \mathbb{R}.
\]
The family \( (H_\ell)_{\ell \in \mathbb{N}} \) forms an orthonormal basis of the Hilbert space of real-valued square integrable functions with respect to the standard normal distribution \( d\gamma(x) = e^{-x^2/2} \frac{dx}{\sqrt{2\pi}} \) on \( \mathbb{R} \). Extend now the polynomials \( H_\ell \) to the complex plane. Then, for each \( \tau > 1 \), the family
\[
\mathcal{H}_\tau(z) = (2\tau^2 - 1)^{-\ell/2} H_\ell(\tau x + i\sqrt{\tau^2 - 1} y), \quad z = x + iy \in \mathbb{C}, \; \ell \in \mathbb{N},
\]
defines an orthonormal sequence with respect to the standard Gaussian measure \( d\gamma(z) = d\gamma(x) d\gamma(y) \) on \( \mathbb{C} \). Here and below we identify \( z = x + iy \in \mathbb{C} \) and \( (x,y) \in \mathbb{R}^2 \). The value \( \tau = 1 \) corresponds to the real case, while when \( \tau \to \infty \), \( \mathcal{H}_\tau(z) \sim \frac{z^\ell}{\sqrt{2\pi \ell^2}} \), which form an orthonormal basis of the space of square integrable analytic functions on \( \mathbb{C} \).

As presented in [8], the correlation functions of the random matrix ensemble \( A^N + i\rho B^N \) are completely described by the Hermite kernel \( \frac{1}{N} \sum_{\ell=0}^{N-1} \mathcal{H}_\ell(z) \mathcal{H}_\ell(z') \) with \( \tau = (1 - \rho^2)^{-1/2} \). In particular, the mean eigenvalue density \( \mu^N \) on the eigenvalues \( \lambda_1^N, \ldots, \lambda_N^N \) of \( A^N + i\rho B^N \) is given by
\[
\langle f, \mu^N \rangle = \mathbb{E} \left( \frac{1}{N} \sum_{k=1}^{N} f(\lambda_k^N) \right) = \int_{\mathbb{C}} f(z) \frac{1}{N} \sum_{\ell=0}^{N-1} |\mathcal{H}_\ell(z)|^2 \, d\gamma(z)
\]
for every bounded measurable function \( f \) on \( \mathbb{C} \). Asymptotics of \( \mu^N \) may thus be read off from the behavior of the probability density \( \frac{1}{N} \sum_{\ell=0}^{N-1} |\mathcal{H}_\ell(z)|^2 \, d\gamma(z) \) on \( \mathbb{C} \) suitably rescaled.

In the contribution [5], inspired by the work [4] by U. Haagerup and S. Thorbjørnsen, simple Markov integration by parts and differential equations on Laplace transforms have been emphasized in the context of real orthogonal polynomial ensembles, demonstrating in particular the underlying universal character of the arcsine law. It was shown namely that, properly rescaled, measures \( \int \mathcal{H}_\tau^2 \, d\gamma \) on the line converge weakly to the arcsine law \( \frac{1}{2} (1 - x^2)^{-1/2} \mathbb{1}_{[|x|<1]} \, dx \) (\( \xi \) will denote below a random variable with this distribution). By a simple averaging procedure, rescaled measures \( \frac{1}{N} \sum_{\ell=0}^{N-1} \mathcal{H}_\ell^2 \, d\gamma \) converge to the product \( \sqrt{U} \xi \) where \( U \) is uniform on \([0,1]\) and independent from \( \xi \), giving thus rise to the semi-circular law for the limiting spectral measure of the GUE by the representation (1.2) (cf. [5]).

The purpose of this note is to show that a similar structure arises in the context of non-Hermitian random matrices with complex eigenvalues, where the central role is now played by the uniform distribution on the unit circle. For any fixed \( \rho \) such that \( 0 < |\rho| \leq 1 \), equivalently \( \tau > 1 \), we show namely that rescaled measures \( |\mathcal{H}_\tau^\ell|^2 \, d\gamma \) converge to \((a \cos \Theta, b \sin \Theta)\) where \( \Theta \) is uniform on the unit circle. Note
that projections of $\Theta$ on diameters have the arcsine distribution. The analysis again relies on second order differential equations for Laplace transforms from which the limiting distribution may easily be identified.

These results and methods extend to non-compactly supported models aspects of the classical theory developed in [7] in which recurrence equations for orthogonal polynomials $P_\ell, \ell \in \mathbb{N}$, of measures $\mu$ on the torus or with compact support on the line are used to show that measures $P_\ell^2d\mu$ converge to the uniform distribution on the circle or the arcsine law. Following [5, 6], the same line of investigation should cover the complex extension of the other classical orthogonal polynomial ensembles.

After averaging, $\frac{1}{N} \sum_{\ell=0}^{N-1} |H_\ell^2|d\gamma$ converges weakly, for every fixed $\tau > 1$, to $\sqrt{U}(a \cos \Theta, b \sin \Theta)$ where $U$ is uniform on $[0, 1]$ and independent from $\Theta$, giving thus rise to the elliptic law for the non-Hermitian random matrix models. Interestingly enough, the approach allows us to investigate as easily the intermediate regime studied by Y. Fyodorov, B. Khoruzhenko and H.-J. Sommers [2] concerned with weak non-Hermiticity as $\rho \to 0 (b \to 0)$ with $N$. The differential equation approach indeed clearly identifies a Gaussian perturbation in the, properly rescaled, vertical direction, and it provides in particular a simple description of the limiting distribution put forward in [2].

In the first section of this note, we present the classical results in the strong non-Hermitian regime $A_N + i\rho B_N$ with $0 < |\rho| \leq 1$ fixed. In the second part, we analyze the weak non-Hermitian regime in which $\rho \to 0$ with $N$.

2. Strong Non-Hermitian Random Matrices

In this section, we deal with strong non-Hermitian matrices given by the Ginibre Ensemble $A_N + i\rho B_N$ with $0 < |\rho| \leq 1$ fixed. In the orthogonal polynomial description, $\tau > 1$ (possibly infinite) is therefore fixed independently of $N$. Following the strategy of [5], the first proposition describes the limiting distribution of the measures $|H_N^2|d\gamma$ properly renormalized by means of differential equations on Laplace transforms.

**Proposition 2.1.** Let $\tau > 0$ be fixed, and let $Z_N = (X_N, Y_N)$ be a random variable with distribution $|H_N^2|d\gamma$ on $\mathbb{C}$. Then, as $N \to \infty$,

$$Z_N \sqrt{\frac{2N}{2\tau^2 - 1}} \to (a \cos \Theta, b \sin \Theta)$$

in distribution, where $\Theta$ is uniform on the unit circle and where $a, b > 0$ are such that $a^2 = \frac{2\tau^2}{2\tau^2 - 1}, b^2 = \frac{2(\tau^2 - 1)}{2\tau^2 - 1}$.

**Proof.** As announced, we follow the general strategy of [5]. Along these lines, it is easy (although somewhat tedious) to show similarly that if we let

$$\varphi(t) = \int_C e^{(\alpha x + \beta y/b)} |H_N(z)|^2d\gamma(z), \quad t \in \mathbb{R},$$

where $\alpha, \beta \in \mathbb{R}, \alpha^2 + \beta^2 = 1$, then $\varphi$ solves the second order differential equation

$$t\varphi'' + [1 + t^2(1 - 2\alpha^2)]\varphi' - t[\alpha^2 t^2(1 - \alpha^2) + 2\alpha^2 + 2N]\varphi = 0$$

(2.2)
where \( c^2 = \frac{a^2}{\tau^2} + \frac{\beta^2}{\tau^2} \). Considering \( q(t/\sqrt{2N}) \) and letting \( N \to \infty \), the limiting differential equation is given, whatsoever the choice of \( \alpha \) and \( \beta \), by \( t\Phi'' + \Phi' - t\Phi = 0 \) which is the characterizing differential equation of the Laplace transform of the arcsine law on \((-1,+1)\). Arguing as in [5] thus shows that, for every \( \alpha, \beta \in \mathbb{R} \) with \( \alpha^2 + \beta^2 = 1 \),

\[
\frac{1}{\sqrt{2N}} \left( \alpha \frac{X_N}{a} + \beta \frac{Y_N}{b} \right) \to \xi
\]

in distribution where \( \xi \) is distributed according to the arcsine law on \((-1,+1)\). Since \( \alpha \cos \Theta + \beta \sin \Theta \), where \( \Theta \) is uniform on the unit circle, is distributed as \( \xi \), the conclusion follows.

Proposition 2.1 interpolates between the limiting cases \( \tau \to 1 \) for which \( a = \sqrt{2}, \ b = 0 \), which gives rise to the arcsine law (the distribution of \( \cos \Theta \)), and \( \tau \to \infty \) for which \( a = b = 1 \), which gives rise to the uniform distribution on the unit circle.

To apply the preceding conclusions to the spectral measure of non-Hermitian random matrix models, we have to deal, according to the representation (1.2), with averages \( \frac{1}{N} \sum_{\ell=0}^{N-1} |H_{\ell}|^2 d\gamma \). This is accomplished by a suitable mixture with an independent uniform random variable. Namely, let \( f: \mathbb{C} \to \mathbb{R} \) be bounded and continuous. Then

\[
\int_{\mathbb{C}} f\left( \frac{z}{\sqrt{2N}} \right) \frac{1}{N} \sum_{\ell=0}^{N-1} |H_{\ell}'(z)|^2 d\gamma(z) = \frac{1}{N} \sum_{\ell=0}^{N-1} \int_{\mathbb{C}} f\left( \frac{\sqrt{\ell}}{N \cdot \sqrt{2N}} \right) |H_{\ell}'(z)|^2 d\gamma(z).
\]

Together with Proposition 2.1 and Lebesgue’s theorem, the limiting distribution as \( N \to \infty \) is thus given by \( \sqrt{U}(a \cos \Theta, b \sin \Theta) \) where \( U \) is uniform on \([0,1]\) and independent from \( \Theta \). Now, the law of \( \sqrt{U}(a \cos \Theta, b \sin \Theta) \) is uniform on the ellipse \( x^2 a^2 + y^2 b^2 \leq 1 \). By (1.2), we thus easily recover Girko’s elliptic law interpolating between the Ginibre Ensemble and the GUE.

**Corollary 2.2.** Let \( A_N \) and \( B_N \) be independent copies from the GUE. For every fixed \( \rho, 0 < |\rho| \leq 1 \), the mean spectral distribution \( \mu_N \) of \( (A_N + \rho B_N)/\sqrt{2N} \) converges as \( N \to \infty \) towards the uniform law on the ellipse \( \frac{x^2}{\tau} + \frac{y^2}{\tau^2} \leq 1 \) where \( a^2 = \frac{2}{1+\rho^2}, \ b^2 = \frac{2\rho^2}{1+\rho^2} \).

As announced, for \( \rho = \pm 1 \), we recover the circular law, and when (formally) \( \rho = 0 \), the semi-circular law (the distribution of \( \sqrt{U} \cos \Theta \)).

### 3. Weak Non-Hermitian Random Matrices

In the weak non-Hermitian regime, the parameter \( \rho \) interpolating between the Ginibre Ensemble and the GUE tends to 0 with \( N \). As investigated in the work [2] by Y. Fyodorov, B. Khoruzhenko and H.-J. Sommers, a new limiting distribution of the spectral measure develops in this regime, provided an appropriate zoom is put on the imaginary part of the eigenvalues. This deformation may be easily identified on the limiting differential equation on Laplace transforms on which it is actually seen that the elliptic law is deformed by a Gaussian variable in the vertical coordinate.
We start again with equation (2.2). Set, for a parameter \( \kappa > 0 \), \( \varphi_\kappa(t) = \varphi(\kappa t) \), \( t \in \mathbb{R} \), where \( \varphi \) is defined in (2.1), and
\[
\psi(t) = e^{-\beta^2 \sigma^2 t^2 / 2} \varphi_\kappa(t), \quad t \in \mathbb{R},
\]
where \( \sigma^2 \geq 0 \). From (2.2), it is easily seen that \( \psi \) solves the second order differential equation
\[
t' \psi'' + \left[ 1 + t^2 (\kappa^2 (1 - 2c^2) + 2\beta^2 \sigma^2) \right] \psi' = 0,
\]
\[
- \left( \kappa^2 c^2 t^2 (1 - c^2) - \beta^2 \sigma^4 - \beta^2 \sigma^2 (1 - 2c^2) \right) \psi' + t \left[ \kappa^2 (2c^2 + 2\beta^2 \sigma^2) \right] \psi = 0.
\]
(3.1)

Therefore, provided that \( \kappa \sim \frac{1}{\sqrt{2N}} \) and \( \kappa^2 c^2 \sim \beta^2 \sigma^2 \), the limiting differential equation is given as above by \( t \Psi'' + \Psi' - t \Phi = 0 \). Thus, with \( \sigma \sim \sigma_N \), \( 2N \sigma_N^2 \to 0 \) and \( 2N \sigma_N^2 b_N^2 \sim 1 \),
\[
\frac{1}{\sqrt{2N}} \left( \frac{X_N}{\sigma_N} + \beta \frac{Y_N}{b_N} \right) \to \xi + \beta \sigma G
\]
in distribution, where \( \xi \) is distributed according to the arcsine law on \((-1, +1)\) and \( G \) is an independent standard normal variable. As in the preceding section, we then conclude in particular to the following result.

**Proposition 3.1.** Let \( Z_N = (X_N, Y_N) \) be a random variable with distribution \( |H_N|^2 d\gamma \) on \( \mathbb{C} \). Then, as \( \sigma_N \to \sigma \geq 0 \) and \( \tau = \tau_N \to 1 \) \( \langle b_N \to 0 \rangle \) such that \( 2N \sigma_N^2 b_N^2 \to \eta^2 > 0 \), \( 0 < \eta < \infty \), \( N \to \infty \),
\[
\left( \frac{X_N}{2\sqrt{N}}, \sigma_N Y_N \right) \to (\cos \Theta, \eta \sin \Theta + \sigma G)
\]
in distribution, where \( \Theta \) is uniform on the unit circle and \( G \) is an independent standard normal variable.

As in the preceding section, Proposition 3.1 may be translated at the level of the averages \( \frac{1}{N} \sum_{\ell=0}^{N-1} |H_{\ell N}^r|^2 d\gamma \), and thus, by (1.2), for the mean spectral measure of the random matrix ensemble. Let \( Z_N = (X_N, Y_N) \) be a random variable with law \( |H_N^r|^2 d\gamma \). For every bounded continuous function \( f : \mathbb{C} \to \mathbb{R} \),
\[
\int_{\mathbb{C}} f \left( \frac{x}{\sqrt{N}}, \tau_N \right) \frac{1}{N} \sum_{\ell=0}^{N-1} |H_{\ell N}^r|^2 d\gamma = \int_1^\infty \int_{\mathbb{C}} f \left( \sqrt{U_N(r)} \frac{x}{NU_N(r)}, \sigma_N y \right) |H_{\ell N}^r|^2 d\gamma dr
\]
where \( U_N(r) = \ell / N \) for \( \ell / N < r \leq (\ell + 1) / N \), \( \ell = 0, 1, \ldots, N - 1 \) \( \langle U_N(0) = 0 \rangle \). Since \( U_N(r) \to r \), \( r \in (0, 1) \), and since \( 2NU_N \sigma_N^2 b_N^2 \to \eta^2 r \), it follows from Proposition 3.1 that
\[
\int_{\mathbb{C}} f \left( \frac{x}{\sqrt{N}}, \sigma_N y \right) \frac{1}{N} \sum_{\ell=0}^{N-1} |H_{\ell N}^r|^2 d\gamma \to \mathbb{E} \left( f(\sqrt{U} \cos \Theta, \eta \sqrt{U} \sin \Theta + \sigma G) \right)
\]
where \( U \) is uniform on \([0,1]\) and independent from \( \Theta \) and \( G \). In the language of
the spectral measure, we recover in this way the main result of [2].

**Corollary 3.2.** Let \( A_N \) and \( B_N \) be independent copies from the GUE, and denote
by \( \lambda_1^N, \ldots, \lambda_N^N \) the eigenvalues of \((A_N + i\rho_N B_N)/2\sqrt{N}\). Set, for each \( N \geq 1 \)
and \( k = 1, \ldots, N \), \( \hat{\lambda}_k^N = \Re(\lambda_k^N) + 2i\sigma_N\sqrt{N} \Im(\lambda_k^N) \), and denote by \( \hat{\mu}_N \) the mean
empirical measure on \( \hat{\lambda}_1^N, \ldots, \hat{\lambda}_N^N \). Then, as \( \sigma_N \to \sigma \geq 0 \) and \( 4N\sigma_N^2\rho_N^2/(1+\rho_N^2) \to \eta^2 > 0 \), \( 0 < \eta < \infty \), \( N \to \infty \), \( \hat{\mu}_N \) converges towards the distribution of
\[
\left( \sqrt{U} \cos \Theta, \eta \sqrt{U} \sin \Theta + \sigma G \right)
\]
where \( U \) is uniform on \([0,1]\), \( \Theta \) is uniform on \([0,2\pi]\), \( G \) is a standard normal
variable, \( U, \Theta \) and \( G \) being independent.

Typically, \( \sigma_N = \sigma \) and \( \rho_N \sim \frac{\eta}{2\sigma_N^2} \) as investigated in [2]. It is not difficult to
represent the density of the limiting distribution in Corollary 3.2 as
\[
\frac{1}{2\pi \eta} \int_{\eta - \sqrt{1-x^2}}^{\eta + \sqrt{1-x^2}} e^{-t^2/2\sigma^2} \frac{dt}{\sqrt{2\pi \sigma^2}}, \quad (x, y) \in (-1, +1) \times \mathbb{R}
\]
(cf. [F-K-S]).

**Acknowledgment.** Thanks are due to B. Khoruzhenko for suggesting that the

**References**

1. Bai, Z. D.: Methodologies in the spectral analysis of large dimensional random matrices: a
5. Ledoux, M.: Differential operators and spectral distributions of invariant ensembles from the
6. Ledoux, M.: Differential operators and spectral distributions of invariant ensembles from the
7. Máté, A., Nevai, P., Totik, V.: Strong and weak convergence of orthogonal polynomials,

MICHEL LEDOUX: INSTITUT DE MATHEMATIQUES, UNIVERSITE DE TOULOUSE, 31062 TOULOUSE,
FRANCE

E-mail address: ledoux@math.univ-toulouse.fr
LOG-SOBOLEV INEQUALITIES WITH POTENTIAL FUNCTIONS ON PINNED PATH GROUPS

SHIGEKI AIDA

Abstract. We establish a refined version of Gross’s log-Sobolev inequalities on pinned path groups. We explain the reason why it is useful in a lower bound estimate of Schrödinger operators on path spaces.

1. Introduction

It is well-known that the hypercontractivity of the diffusion semi-group and the equivalent notion of the validity of the log-Sobolev inequality are used to give a lower bound on the bottom of spectrum of a Schrödinger operator which is given by the sum of the generator of the semi-group and a potential function. We applied this lower bound estimate to study the semi-classical behavior of the bottom of spectrum of Schrödinger operators on path spaces over compact Riemannian manifolds ([3, 4]) partly motivated by an application to $P(\phi)$-type Hamiltonian and an extension of [19] to infinite dimensional curved spaces.

In this paper, we establish a refined version of Gross’s log-Sobolev inequalities on a pinned path group. Pinned path group $P_{e,a}(G)$ is a space of continuous paths with values in a compact Lie group $G$ over the time interval $[0,1]$ with a fixed starting point $e$ (unit element) and the fixed end point $a$. We will apply the log-Sobolev inequalities with potential functions to study the semi-classical behavior of the low lying spectrum of Schrödinger operators over pinned path groups in a separate paper.

The structure of the paper is as follows. In Section 2, we consider smooth pinned path spaces over a general compact Riemannian manifold $M$. We introduce a Riemannian structure on the pinned path space using a metric connection on $M$. Next we calculate the gradient of the energy function $E(\gamma) = \frac{1}{2} \int_0^1 |\dot{\gamma}(t)|^2 dt$. This calculation and a formal argument show that an LSI with a special potential function may be useful for the study of semi-classical behavior of low lying spectrum of Schrödinger operators over a pinned path space. This kind of log-Sobolev inequality with special potential function already appeared in [13, 10]. In Section 3, we consider a pinned path group and introduce an $H$-derivative on $P_{e,a}(G)$ and the Dirichlet form in $L^2$-space with respect to the (scaled) pinned Brownian motion measure with scaling (semi-classical) parameter $\lambda$. The $H$-derivative is considered as the gradient operator on the path space which is defined by the right invariant
connection on $G$. We prove a special kind of LSI which is introduced in Section 2 in the case of the pinned path group with respect to the Dirichlet form. The generator of the Dirichlet form is the Ornstein-Uhlenbeck type operator. However we cannot apply the argument in [3, 4] to the Ornstein-Uhlenbeck type operator on $P_{e,a}(G)$ itself, since the function $\frac{1}{2} |b(1)|^2$ does not satisfy the exponential integrability condition. We refer the reader to Section 2 and Section 3 for the definition of $b(1)$. For that reason, we consider a sequence of bounded domains of $\{\Omega_{L,\varepsilon}\}_{L>0}$ of $P_{e,a}(G)$ which exhaust the path space in Section 4. $\Omega_{L,\varepsilon}$ is the $\varepsilon$-neighborhood in the uniform convergence topology of the level set of the energy function, $\Omega_L = \{ \gamma \mid \sqrt{E(\gamma)} \leq L \}$. The similar subset appeared in the study of [9]. We prove that $\frac{1}{2} |b(1)|^2$ satisfies good integrability properties on these subsets. This integrability properties will be applied to determine the semi-classical behavior of low lying spectrum of the Ornstein-Uhlenbeck operator with Dirichlet boundary condition on $\Omega_{L,\varepsilon}$ in the forthcoming paper.

2. Path Integral and Logarithmic Sobolev Inequalities

Let $(M,g)$ be a $d$-dimensional complete connected Riemannian manifold. Let $\Gamma$ be a metric connection whose torsion $T$ satisfies that

\[ g(T(X,Y),Z) = -g(Y,T(X,Z)) \]

for any vector fields $X, Y, Z$. We refer the reader to [8] for the notion, “torsion skew symmetric connection”. Let $x \in M$ and consider a smooth path $\gamma(t)$ $(0 \leq t \leq 1)$ on $M$ starting at $x$. Along $\gamma$, the parallel translation operator $\tau(\gamma)_t : T_x(M) \rightarrow T_{\gamma(t)}M$ is defined by the connection $\Gamma$. Let $P_{x,y,H^1}(M)$ be the space of $H^1$ maps from $[0,1]$ to $M$ with $\gamma(0) = x, \gamma(1) = y \in M$. Let $T_xP_{x,y,H^1}(M)$ be the tangent space at $\gamma$ which consists of mapping $h$ from $[0,1]$ to $TM$ such that $h(t) \in T_{\gamma(t)}M$, $h(0)$ and $h(1)$ is 0, and its $H^1$-norm

\[ ||h||_{T_xP_{x,y,H^1}} = \left\{ \int_0^1 \left| \frac{d}{dt} (\tau(\gamma)^{-1}h(t)) \right|^2 dt \right\}^{1/2} \]

is finite. This Hilbert norm defines a Riemannian metric on $P_{x,y,H^1}(M)$. The gradient operator $\nabla$ which is naturally defined by the metric is given explicitly for a smooth cylindrical function $f(\gamma) = F(\gamma(t_1), \ldots, \gamma(t_n)) \in \mathcal{S}_b^\infty(P_{x,y,H^1}(M))$ by

\[ (\nabla f)(\gamma)_t = \sum_{i=1}^n \tau(\gamma)^{-1}_{t_1} \nabla(i)F(\gamma)(t \wedge t_i - tt_i). \tag{2.1} \]

Here $t \wedge t_i = \min(t, t_i)$ and $\nabla(i)F(\gamma) \in T_{\gamma(t)}M$ denotes the covariant derivative with respect to the $i$-th variable. Now we consider measures on path spaces. Formally, we consider the Riemannian measure (Feynman measure) $d\gamma$ on $P_{x,y,H^1}(M)$ which is defined by the Riemannian metric. On the other hand, the Brownian bridge measure which is denoted by $\nu_{x,y}$ is rigorously defined on $P_{x,y}(M)$ which is a space of continuous map from $[0,1]$ to $M$ such that $\gamma(0) = x, \gamma(1) = y$. The formal expression of $\nu_{x,y}$ is given by

\[ d\nu_{x,y}(\gamma) = Z^{-1}_\lambda \exp \left( -\lambda \gamma(\gamma) \right) d\gamma, \]
Thus we assume that the following LSI holds:

\[ \mathcal{E}_\lambda(f, f) = \int_{P_{x,y}(M)} \frac{|(\nabla f)(\gamma)|^2}{\lambda_2} \, d\mu_{x,y}(\gamma), \]

where \( f \in \mathcal{C}_b^\infty(P_{x,y}(M)) \). The gradient operator \( \nabla \) is also defined by the formula (2.1) using the stochastic parallel translation. We denote the generator by \(-L_\lambda\). Let \( V \) be a real-valued measurable function on \( P_{x,y}(M) \) and consider a Schrödinger operator \(-L_{\lambda,V} = -L_\lambda + \lambda^2 V\) on \( L^2(P_{x,y}(M), d\mu_{x,y}) \). Set \( E(\lambda, V) = \inf \sigma(-L_{\lambda,V}) \). We are interested in determining the asymptotics of \( E(\lambda, V) \) when \( \lambda \to \infty \). Let \( \Delta \) be the Laplace-Bertlami operator on \( P_{x,y}(M) \) which is formally defined by the Riemannian metric. Then \(-L_{\lambda,V} \) on \( L^2(P_{x,y}(M), d\mu_{x,y}) \) is formally unitary equivalent to

\[ -H_{\lambda,V} = -\Delta + \lambda^2 \left( \frac{1}{4} |(\nabla E)(\gamma)|^2_{T_1(P_{x,y}(M))} + V(\gamma) \right) - \frac{\lambda}{2} \Delta E(\gamma) \]

on \( L^2(P_{x,y}(M), d\gamma) \) by the mapping

\[ f \in L^2(P_{x,y}(M), d\mu_{x,y}) \to f : \lambda_2^{-1/2} e^{-\lambda (\gamma)} \in L^2(P_{x,y}(M), d\gamma). \]

For \( \gamma \in P_{x,y,H^1}(M) \), let \( b(t) = \int_0^t \tau(\gamma)^{-1} \gamma(s) ds \). Then \( E(\gamma) = \frac{1}{4} \int_0^1 |\dot{b}(t)|^2 dt \). Note that

\[ \nabla_h b(t) = h(t) + \int_0^t \left( \int_0^s \overline{R(\gamma)(u)}(h(u), \dot{b}(u)) du \right) \dot{b}(s) ds - \int_0^t \overline{T(\gamma)(h(u), \dot{b}(s))} ds, \]

where \( R(X, Y)Z \) denotes the curvature tensor and for any \( \xi_i \in T_x M \),

\[ \overline{R(\gamma)}(\xi_1, \xi_2)\xi_3 = \tau(\gamma)^{-1} R(\gamma(t))(\tau(\gamma)\xi_1, \tau(\gamma)\xi_2)\tau(\gamma)\xi_3, \]

\[ \overline{T(\gamma)}(\xi_1, \xi_2) = \tau(\gamma)^{-1} T(\gamma(t))(\tau(\gamma)\xi_1, \tau(\gamma)\xi_2). \]

By the skew symmetric property of the curvature tensor and the torsion, we have \( (\nabla E)(\gamma) = b(t) - tb(1) \) and

\[ \frac{1}{4} |(\nabla E)(\gamma)|^2_{T_1(P_{x,y}(M))} = \frac{1}{4} \int_0^1 |\dot{b}(t)|^2_{T_2(P_{x,y}(M))} dt - \frac{1}{4} |b(1)|^2_{T_2(P_{x,y}(M))}. \]

Thus

\[ -H_{\lambda,V} = -\Delta + \lambda^2 \left( \frac{1}{2} E(\gamma) - \frac{1}{4} |b(1)|^2_{T_2(P_{x,y}(M))} + V(\gamma) \right) - \frac{\lambda}{2} \Delta E(\gamma). \]

We assume that the following LSI holds:

\[ \int_{P_{x,y}(M)} f(\gamma)^2 \log \left( \frac{f(\gamma)^2}{\|f\|^2_{L^2(\mu_{x,y})}} \right) d\mu_{x,y} \leq \frac{2}{\lambda} \left( 1 + \frac{C}{\lambda} \right) \mathcal{E}_{\lambda,V_{\lambda,x,y}}(f, f), \]

(2.2)
where
\[ E_{\lambda,V}(f,f) = \int_{P_{x,y}(M)} |(\nabla f)(\gamma)|^2 \, d\nu_{\lambda,x,y} \]
\[ + \int_{P_{x,y}(M)} \lambda^2 V_{\lambda,x,y}(\gamma) f(\gamma)^2 \, d\nu_{\lambda,x,y} \]
\[ V_{\lambda,x,y}(\gamma) = \frac{1}{4} \left\{ \frac{|b(1)|^2}{T_x(M)} + \frac{2}{\lambda} \log \left( \lambda^{-d/2} p \left( \frac{1}{\lambda}, x, y \right) \right) + \frac{1}{\lambda} W(\gamma) \right\} \]
and \( C \) is a positive constant and \( W \) is a real-valued measurable function. Getzler [10] and Gross [13] proved LSI's in the above forms in the case where \( M \) is a compact Lie group and \( x = y \) are unit element \( e \). They point out that \( \frac{1}{4} |b(1)|^2 \) is main term in the inequalities. LSI on a loop space \( P_{x,x}(M) \) over general compact Riemannian manifolds were studied in [2, 11]. However the potential functions in their inequalities are very different from the above. Let \( -L_{\lambda,V_{\lambda,x,y}} \) be the generator of the form \( E_{\lambda,V_{\lambda,x,y}} \). Then by the results in [12],
\[ E(\lambda,V) = \inf \sigma \left( -L_{\lambda,V_{\lambda,x,y}} - \lambda^2 V_{\lambda,x,y}(\gamma) + \lambda^2 V \right) \geq -\frac{\lambda}{2} \left( 1 + \frac{C}{\lambda} \right)^{-1} \log I(\lambda). \] (2.3)
Here
\[ I(\lambda) = \int_{P_{x,y}(M)} \exp \left\{ \lambda(1 + \frac{C}{\lambda}) \left( \frac{1}{2} |b(1)|^2_{T_x(M)} \right) - 2V(\gamma) \right\} \]
\[ \exp \left\{ 2(1 + \frac{C}{\lambda}) W(\gamma) \right\} \lambda^{-d/2} d\nu_{\lambda,x,y}(\gamma) \]
and \( \nu_{\lambda,x,y} = p(1/\lambda, x, y)\nu_{\lambda,x,y} \). Let
\[ U(\gamma) = \frac{1}{2} E(\gamma) - \frac{1}{4} |b(1)|^2_{T_x(M)} + V(\gamma). \]
We assume that \( U(\gamma) \geq 0 \) for all \( \gamma \) and \( \min U = 0, N = \{ \gamma \mid U(\gamma) = 0 \} \) are finite set and the hessian of \( U \) at them are non-degenerate. These assumptions are standard in semi-classical analysis in finite dimensions. We note that \( \lim_{\lambda \to \infty} I(\lambda) \) converges under these and certain additional integrability assumptions. See [16]. This implies that \( \liminf_{\lambda \to \infty} \frac{E(\lambda,V)}{\lambda} < -\infty \). In the case where \( M = \mathbb{R}^d \), \( b(1) = y - x \) holds and the asymptotics of \( I(\lambda) \) is a classical problem. We determined the value of \( \lim_{\lambda \to \infty} \frac{E(\lambda,V)}{\lambda} \) under certain assumptions on \( V \) in the case where \( M = \mathbb{R}^d \) in [3]. Now we consider the case where \( M \) is a compact connected Lie group \( G \) and the metric is bi-invariant under the action of \( G \) and \( x \) is the unit element \( e \). Let \( \Gamma \) be the right invariant connection. That is, for a smooth curve \( \gamma(t) \ (\gamma(0) = e) \), the parallel translation is given by \( \tau(\gamma)^{-1}(h(t)) = (R_{\gamma(t)^{-1}})_x h(t) \), where \( R_g h = hg \ (g,h \in G) \).
Since this is a torsion skew symmetric connection, if the inequality (2.2) holds, then the above formal argument could be applied to this case. In the next section, we prove an LSI as in (2.2).
3. Refined Version of Gross’s Logarithmic Sobolev Inequalities on Pinned Path Groups

Let $G$ be a $d$-dimensional compact connected and simply connected Lie group. Let $e$ be the unit element of $G$. We denote the Lie algebra of $G$ by $\mathfrak{g}$ which is identified with $T_e G$. Actually, $G$ is isomorphic to a Lie subgroup of $n$-dimensional unitary group $SU(n)$ and the Lie algebra is isomorphic to a Lie subalgebra of $\mathfrak{su}(n)$. In this case, $e$ is the identity matrix and the Lie bracket is given by $[A, B] = AB - BA$. By this result, we may assume that $G$ is a matrix group. That is $G \subset SU(n) \subset M(n, \mathbb{C}) \cong \mathbb{C}^{n^2}$. $M(n, \mathbb{C})$ denotes the all $n \times n$-matrices whose elements are complex numbers. For $A, B \in M(n, \mathbb{C})$, let $(A, B) = \text{real part of } \text{tr}AB^*$. This defines an inner product on $M(n, \mathbb{C})$ and a bi-invariant Riemannian metric on $G$. We denote by $dx$ the Riemannian measure (Haar measure). Then by the bi-invariance of the metric, we have

$$
\int_G f(x)dx = \int_G f(gx)dx = \int_G f(xy)dx = \int_G f(x^{-1})dx.
$$

(3.1)

For $A \in \mathfrak{g}$, $e^A, \exp A$ denotes the matrix exponential element as well as the exponential map in Riemannian geometry sense. Let $i(G)$ be the injectivity radius of $G$. That is,

$$
i(G) = \sup \{ r \mid \exp : B_r(0) \to G \text{ gives a local chart at } e \},$$

where $B_r(0) = \{ v \in T_e G \mid |v| < r \}$. Let $P_\gamma(G)$ be a set of continuous paths $\gamma(t) \ (0 \leq t \leq 1)$ with values in $G$ with $\gamma(0) = e$. Let $\lambda > 0$. There exists a probability measure which is called the Brownian motion measure $\nu_\lambda$ on $P_\gamma(G)$ such that for $0 = t_0 < t_1 < \cdots < t_n \leq 1$,

$$
\nu_\lambda (\gamma_{t_1} \in A_1, \ldots, \gamma_{t_n} \in A_n)
= \int_{A_1} dx_1 \cdots \int_{A_n} dx_n \prod_{i=1}^{n} p \left( \lambda^{-1}(t_i - t_{i-1}), x_{i-1}, x_i \right),
$$

where $x_0 = e$. Here $p(t, x, y)$ denotes the heat kernel of $e^{t\Delta/2}$. Let $P_{\gamma, a}(G)$ be a subset of $P_\gamma(G)$ such that $\gamma(1) = a$. Also we denote $P_{\gamma, a, H}(G) = \{ \gamma \in P_{\gamma, a}(G) \mid E(\gamma) := \frac{1}{2} \int_0^1 |\dot{\gamma}(t)|^2 dt < \infty \}$. There exists a probability measure $\nu_{\lambda, a}$ which is called a Brownian bridge measure on $P_{\gamma, a}(G)$ such that for $0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = 1$

$$
\nu_{\lambda, a} (\gamma_{t_1} \in A_1, \ldots, \gamma_{t_n} \in A_n)
= p(\lambda^{-1}, e, a)^{-1} \int_{A_1} dx_1 \cdots \int_{A_n} dx_n \prod_{i=1}^{n+1} p \left( \lambda^{-1}(t_i - t_{i-1}), x_{i-1}, x_i \right),
$$

where $x_0 = e, x_{n+1} = a$. Let

$$
H(\mathfrak{g}) = \left\{ h : [0, 1] \to \mathfrak{g} \mid h(0) = 0, \| h \|_{H}^2 := \int_0^1 |\dot{h}(t)|^2 dt < \infty \right\}
$$

$$
H_0(\mathfrak{g}) = \left\{ h \in H(\mathfrak{g}) \mid h(1) = 0 \right\}.
$$
Below, we use the notation $X$ to denote $P_{e,a}(G)$, $P_{e}(G)$. Let $\mathfrak{A}C^\infty_b(X)$ be the set of smooth cylindrical functions on $X$. We define the $H$-derivative $(\nabla F)(\gamma)$ of $F(\gamma) \in \mathfrak{A}C^\infty_b(X)$ by the unique element of $H(\mathfrak{g})$ (or $H_0(\mathfrak{g})$) satisfying that
\[
((\nabla F)(\gamma), h)_H := \lim_{\varepsilon \to 0} \frac{F(e^{\varepsilon h} \gamma(\cdot)) - F(\gamma)}{\varepsilon},
\]
for all $h \in H(\mathfrak{g})$ (or $h \in H_0(\mathfrak{g})$).

We introduce necessary tools for our analysis. We define $g$-valued process $b(t, \gamma) = \int_0^t d\gamma(s) \circ \gamma(s)^{-1}$ for all $t \leq \sup \gamma$. The above integral is Stratonovich integral when $\gamma(t)$ is the Brownian motion. However $b(t)$ could be still defined under $\nu_{\lambda,a}$ by the quasi-sure analysis [17]. We collect basic results for $b(t)$. Below $\{\varepsilon_i\}_{i=1}^d$ denotes an orthonormal basis on $g$.

**Lemma 3.1.** Below, we consider the stochastic processes under the law of $\nu_{\lambda}$. (1) The distribution of $b(\cdot)$ are the Brownian motion measure which satisfies
\[
E[(b(t, \varepsilon_i)(b(t, \varepsilon_j))] = \frac{1}{\lambda} \delta_{i,j} s \wedge t.
\]
(2) Let $C = \sum_{i=1}^d \varepsilon_i^2 \in M(n, \mathbb{C})$. Then
\[
d(\gamma(t)^{-1}) = -\gamma(t)^{-1} db(t) + \frac{1}{2\lambda} \gamma(t)^{-1} C dt.
\]
The proof of this lemma is standard and we refer the proof to [13, 8]. Also we note that $C$ is called the Casimir element which commutes any matrices in $G$.

In the lemma below, we use the notation in the Watanabe’s distribution theory [18]. The reader may find the proof in [13].

**Lemma 3.2.** (1)
\[
\nabla_h b(t) = h(t) + \int_0^t [h(s), db(s)]
\]
(2) $b(t, e^{\xi} \gamma(\cdot)) = \int_0^t Ad(e^{s\xi}) db(s) + t\xi$
(3)
\[
\int_{P_e(G)} F(e^{\xi} \gamma(\cdot)) G(b) \exp \left(-\lambda(b(1), \xi) - \frac{\lambda}{2} |\xi|^2 \right) \delta_{e}(\gamma(1)) d
nu_{\lambda}(\gamma)
\]
\[
= \int_{P_e(G)} F(\gamma) G \left(\int_0^t Ad(e^{-s\xi}) db(s) - \xi \right) \delta_{e}(\gamma(1)) d
nu_{\lambda}(\gamma).
\]
(4) It holds that
\[
\int_{P_{e,a}(G)} \nabla_h F(\gamma) G(\gamma) d
\nu_{\lambda,a}(\gamma)
\]
\[
= \int_{P_{e,a}(G)} F(\gamma) \{-\nabla_h G(\gamma) + \lambda(b, h) H G(\gamma)\} d
\nu_{\lambda,a}(\gamma).
\]

(3.3)
There exist constants $C_1, C_2 > 0$ such that for any sufficiently large $\lambda > 0$ and $f \in \mathfrak{g}C_0^\infty(P_{\gamma,a}(G))$, it holds that
\[
\int_{P_{\gamma,a}(\mathbb{G})} f^2(\gamma) \log \left( \frac{f^2(\gamma)}{\|f\|^2_{L^2(\nu_{\lambda,a})}} \right) d\nu_{\lambda,a}(\gamma) \leq \frac{2}{\lambda} \left( 1 + \frac{C_1}{\lambda} \right) \mathcal{E}_{\lambda,V_{\gamma,a}}(f, f),
\]
where
\[
\mathcal{E}_{\lambda,V_{\gamma,a}}(f, f) = \int_{P_{\gamma,a}(\mathbb{G})} |(\nabla f)(\gamma)|^2 d\nu_{\lambda,a}
+ \int_{P_{\gamma,a}(\mathbb{G})} \lambda^2 V_{\lambda,a}(\gamma) f(\gamma)^2 d\nu_{\lambda,a},
\]
\[
V_{\lambda,a}(\gamma) = \frac{1}{4} \left\{ |b(1)|^2 + \frac{2}{\lambda} \log \left( \lambda^{-d/2} p(1/\lambda, e, a) \right) \right\}
+ \frac{C_2}{\lambda} \left\{ 1 + |b(1)|^2 + \left( \int_0^1 |b(s)| ds \right)^2 \right\}.
\]

Remark 3.4. If $e$ and $a$ are sufficiently close, then there exist $C_1, C_2 > 0$ such that for all sufficiently small $t$
\[
C_1 t^{-d/2} e^{-d(e,a)^2/(2t)} \leq p(t, e, a) \leq C_2 t^{-d/2} e^{-d(e,a)^2/(2t)}.
\]
Thus (3.4) is equivalent to the inequality which is obtained by replacing $V_{\lambda,a}(\gamma)$ by $\tilde{V}_{\lambda,a}(\gamma)$:
\[
\tilde{V}_{\lambda,a}(\gamma) = \frac{1}{4} (|b(1)|^2 - d(e,a)^2) + \frac{C_2}{\lambda} \left\{ 1 + |b(1)|^2 + \left( \int_0^1 |b(s)| ds \right)^2 \right\}.
\]

We use the following lemmas to prove Theorem 3.3.

**Lemma 3.5.** Let $\delta > 0$ and set $U_\delta(x) = \{ y \in G \mid d(x, y) < \delta \}$. For a sufficiently small $\delta$, there exists a smooth map $\log : U_\delta(e) \to \mathfrak{g}$ such that $\exp(\log y) = y$ for any $y \in U_\delta(e)$ and $\log e = 0$.

**Lemma 3.6.** (1) Assume that $d(x, e)$ is sufficiently small. Let
\[
\Psi(\xi) = \log(x \exp(\xi)).
\]
Then
\[
\Psi'(0) = \xi + \frac{1}{2} \log x, \xi + A_{\log x} \xi,
\]
where $A_v$ is a linear map on $\mathfrak{g}$ such that $\|A_v\|_{op} \leq C|v|^2$.

(2) For $\gamma$ with $d(a, \gamma(1))$ to be sufficiently small,
\[
\nabla_h \left( \log(a\gamma(1)^{-1}) \right) = -h(1) - \frac{1}{2} \log(a\gamma(1)^{-1}) + h(1) + A_{\log(a\gamma(1)^{-1})} h(1).
\]
Moreover, in the calculation below, $I$ denotes the identity matrix. We have
\[
\Psi'(0) = x\xi + \sum_{n=2}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n-1} (x - I)^k x\xi (x - I)^{n-k-1}
\]
\[
= \xi + \frac{1}{2}[x - I,\xi] + \sum_{n=2}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n-1} (x - I)^k x\xi (x - I)^{n-k-1}
+ \sum_{n=3}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n-1} (x - I)^k x\xi (x - I)^{n-k-1}.
\]
Note that $x\xi$ etc are matrix product in $GL(n, \mathbb{C})$. This follows from the Taylor expansion:
\[
\Psi(\varepsilon) = \log \left( (I + (x - I) + \varepsilon x\xi + O(\varepsilon^2)) \right)
= \log x + \varepsilon \left( x\xi + \sum_{n=2}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{k=0}^{n-1} (x - I)^k x\xi (x - I)^{n-k-1} \right) + O(\varepsilon^2).
\]
Noting $\log x = x - I + O(|x - I|^2)$, we get (3.5). □

**Lemma 3.7.** Let $\delta$ be a sufficiently small positive number. Let $S : P_{e, U_0}(G) \to P_{e, a}(G)$ be the map which is defined by
\[
S(\gamma)(t) = \exp \left( t \log(a\gamma(1)^{-1}) \right) \gamma(t) \quad (0 \leq t \leq 1).
\]
Here $P_{e, U_0}(G) = \{ \gamma \in P_e(G) \mid \gamma(1) \in U_0(a) \}$. Then, the following statements hold.
(1) Let $T(\gamma)$ be the bounded linear operator from $H(\mathfrak{g})$ to $H_0(\mathfrak{g})$ such that
\[
(T(\gamma)h)(t) = \left( (R_{S(\gamma)(t)})_t \right)^{-1} \frac{d}{d\varepsilon} \left( S(t\varepsilon \gamma)(t) \right)|_{\varepsilon=0}.
\]
$R_\varepsilon$ denotes the derivative of the right translation. Then for any $f \in \mathcal{F}C^\infty_{\mathfrak{g}}(P_{e, a}(G))$ and $h \in H(\mathfrak{g})$, it holds that
\[
(\nabla(f \circ S)(\gamma), h)_H = ((\nabla f)(S_\gamma), T(\gamma)h)_{H_0}.
\]
Moreover $T(\gamma)$ can be written in the following form:
\[
(T(\gamma)h)(t) = h(t) - th(1) + \frac{t(t-1)}{2} \left[ \log(a\gamma(1)^{-1}), h(1) \right]
+ t[t[\log(a\gamma(1)^{-1}), h(1)] - th(1)] + A_{\log(a\gamma(1)^{-1})}(t)h(t). \tag{3.7}
\]
$A_v(t)$ is a $GL(n, \mathbb{R})$-valued smooth function of $t \in [0, 1]$ and $v \in \mathfrak{g}$ and satisfies that
\[
\max_{0 \leq t \leq 1} \left\{ \| A_v(t) \|_{op} + \left\| \frac{\partial}{\partial t} A_v(t) \right\|_{op} \right\} \leq C\| v \|_{\mathfrak{g}}^2
\]
and $A_v(1) = 0$ for all $v$.
(2) For any $f \in \mathcal{F}C^\infty_{\mathfrak{g}}(P_{e, a}(G))$, it holds that
\[
\| \nabla(f \circ S)(\gamma) \|_H^2 \leq (1 + C\| \log(a\gamma(1)^{-1}) \|_{\mathfrak{g}}^2)(\| \nabla f(S_\gamma) \|_{H_0}^2
+ 2(B_{\log(a\gamma(1)^{-1})}(\nabla f)(S_\gamma), (\nabla f)(S_\gamma))_{H_0}).
\]
Here $B_{v}$ is a bounded linear operator on $H_{0}(g)$ such that $B_{v}h(t) = M^{*}([v, h(t)])$ and $M^{*}$ is the adjoint operator of $M$ on $H_{0}(g)$ which is given by $(Mh)(t) = th(t)$.

**Proof.** (1) (3.6) is the chain rule. We prove (3.7). For $x$ which is sufficiently close to $e$, we denote $K_{x}E = \xi = \frac{1}{2}\log x, E + A_{0}x_{x}$, where this operator appeared in Lemma 3.6 (1). Let $k_{e}(t) = S(e^{h}(\gamma)(t))$. Then by (3.5),

$$k_{e}(t) - k_{0}(t)$$

$$= \exp\left\{ t \left( (a\gamma(1)^{-1} - \epsilon K_{a\gamma(1)}^{-1} h(1) + O(\epsilon^{2})) e^{h(t)} \right) \right\} \gamma(t)$$

$$- \exp\left\{ t \left( (a\gamma(1)^{-1}) \right) \gamma(t) \right\}$$

$$= \left\{ \exp\left[ t \left( (a\gamma(1)^{-1} - \epsilon K_{a\gamma(1)}^{-1} h(1) + O(\epsilon^{2})) \right) \right] \right\} e^{h(t)} \gamma(t)$$

$$+ \exp\left[ t \left( (a\gamma(1)^{-1}) \right) \right] (e^{h(t)} - I) \gamma(t) := I_{1}(\epsilon) + I_{2}(\epsilon).$$

Below, we denote $[A_{v}, B_{+}] = AB + BA$ and $A_{v}(t)$ denotes an $GL(n, \mathbb{R})$-valued smooth function of $0 \leq t \leq 1$ and $v \in g$ and satisfies that $\|A(t)\|_{op} \leq C\|v\|_{g}^{2}$ for all $v \in g$. We have

$$I_{1}(\epsilon) = -\left[ \epsilon tK_{a\gamma(1)}^{-1}h(1) + \frac{\epsilon^{2}}{2} t^{2} \left\{ \log(a\gamma(1)^{-1}), K_{a\gamma(1)}^{-1}h(1) \right\} \right.$$ \n
$$+ \frac{\epsilon}{2} t^{n} \left\{ \sum_{k=0}^{n-1} (\log(a\gamma(1)^{-1}))^{k} K_{a\gamma(1)}^{-1}h(1) \left( \log(a\gamma(1)^{-1}) \right)^{n-k-1} \right\}$$

$$+ O(\epsilon^{2}) \right\} e^{h(t)} \gamma(t),$$

$$\lim_{\epsilon \to 0} \frac{I_{1}(\epsilon)}{\epsilon}$$

$$= -\left[ tK_{a\gamma(1)}^{-1}h(1) + \frac{t^{2}}{2} \left\{ \log(a\gamma(1)^{-1}), K_{a\gamma(1)}^{-1}h(1) \right\} \right.$$ \n
$$\left. + \sum_{n=3}^{\infty} \frac{t^{n}}{n!} \left\{ \sum_{k=0}^{n-1} (\log(a\gamma(1)^{-1}))^{k} K_{a\gamma(1)}^{-1}h(1) \left( \log(a\gamma(1)^{-1}) \right)^{n-k-1} \right\} \right\} \gamma(t)$$

$$= -\left\{ tK_{a\gamma(1)}^{-1}h(1) + \frac{t^{2}}{2} \left\{ \log(a\gamma(1)^{-1}), K_{a\gamma(1)}^{-1}h(1) \right\} \right.$$ \n
$$+ A_{\log(a\gamma(1)^{-1})}^{(1)}(h(1)) \gamma(t) \right\} =: J_{1}(t).$$

Therefore

$$J_{1}(t)S(\gamma)(t)^{-1}$$

$$= -\left\{ th(1) - t^{2}h(1) \log(a\gamma(1)^{-1}) + \frac{1}{2} t \left\{ \log(a\gamma(1)^{-1}), h(1) \right\} \right.$$ \n
$$+ \frac{t^{2}}{2} \left\{ \log(a\gamma(1)^{-1}), K_{a\gamma(1)}^{-1}h(1) \right\} + A_{\log(a\gamma(1)^{-1})}^{(2)}(h(1)).$$
We consider the term $I_2(\varepsilon)$.

\[
\lim_{\varepsilon \to 0} \frac{I_2(\varepsilon)}{\varepsilon} = \exp \left[ t \log(a\gamma(1)^{-1}) \right] h(t) =: J_2(t).
\]

Hence

\[
J_2(t)S(\gamma)^{-1} = Ad \left( \exp \left[ t \log \left( a\gamma(1)^{-1} \right) \right] \right) h(t)
= h(t) + t \left[ \log(a\gamma(1)^{-1}), h(t) \right] + A_{\log(a\gamma(1)^{-1})}^{(3)}(t)h(t).
\]

Therefore, we have

\[
(J_1(t) + J_2(t)) (S(\gamma)(t))^{-1}
= h(t) - th(1) + \frac{t(t - 1)}{2} \left[ \log(a\gamma(1)^{-1}), h(1) \right]
+ t\left[ \log(a\gamma(1)^{-1}), h(t) - th(1) \right] + A_{\log(a\gamma(1)^{-1})}^{(4)}(t)h(t) + A_{\log(a\gamma(1)^{-1})}^{(5)}(t)h(1).
\]

Since $J_1(1) + J_2(1) = 0$ for all $h$, $A_{\nu}^{(i)}(1) = 0$ ($i = 4, 5$) holds for all $v$ and this implies (3.7) holds. We prove (2). Let $\{e_n\}$ be the complete orthonormal system on $H(\mathfrak{g})$ as follows: $e_n(t) = e_n t$ ($1 \leq n \leq d$) and $e_n(1) = 0$ for all $n \geq d + 1$. Then it holds that for all $1 \leq n \leq d$

\[
|((\nabla f)(S\gamma), T(\gamma)e_n)|^2 \leq C\| \log(a\gamma(1)^{-1}) \|_\theta^2 (\| \nabla f \|_{H_0})^2.
\]

Also for $n \geq d + 1$,

\[
|((\nabla f)(S\gamma), T(\gamma)e_n)|^2 \leq ((\nabla f)(S\gamma), e_n)^2 + ((\nabla f)(S\gamma), [\tilde{e}_n, \log(a\gamma(1)^{-1})])^2
+ 2((\nabla f)(S\gamma), e_n)([(\nabla f)(S\gamma), [\tilde{e}_n, \log(a\gamma(1)^{-1})])]
+ C\| (\nabla f)(S\gamma) \|_{H_0}^2 \| \log(a\gamma(1)^{-1}) \|_{\theta},
\]

where $\tilde{e}_n(t) = te_n(t)$. Since $\{e_n\}_{n=d+1}^\infty$ is a complete orthonormal system of $H_0(\mathfrak{g})$,

\[
\sum_{n=d+1}^\infty (((\nabla f)(S\gamma), e_n)(((\nabla f)(S\gamma), [\tilde{e}_n, \log(a\gamma(1)^{-1})]))
= (B_{\log(a\gamma(1)^{-1})}((\nabla f)(S\gamma), (\nabla f)(S\gamma)))_{H_0}.
\]

This completes the proof. 

\[\square\]

**Lemma 3.8.** For $f \in H^\infty_{\delta}(P_{c,a}(G))$, let

\[
\hat{f}\phi(\gamma) = f(S(\gamma))\Phi(\gamma)\phi(\gamma),
\]

\[
\Phi(\gamma) = \exp \left( -\lambda \frac{1}{2} (\log(a\gamma(1)^{-1}), b(1)) - \frac{\lambda}{4} \| \log(a\gamma(1)^{-1}) \|^2 \right),
\]

\[
\phi(\gamma) = \phi \left( \sqrt{\lambda} \log(a\gamma(1)^{-1}) \right) \psi \left( \frac{d(a\gamma(1)^{-1})}{\delta} \right) D_{\lambda}^{-1}.
\]

Here $\phi$ is a non-negative smooth function on $\mathfrak{g}$ such that $\phi(v) = \phi(-v)$ for all $v$ and $\phi(v) = 0$ for $v$ with $\|v\| \geq 1$. $\psi$ is a smooth function with compact support on $\mathbb{R}$.
and takes a value 1 near 0. $\delta$ is a sufficiently small number such that $\log(a_\gamma(1)^{-1})$ is well-defined. Also we set

$$D_\lambda = \{p(1/\lambda, e, a)E_\lambda\}^{1/2}, \quad E_\lambda = \int_G \phi \left(\sqrt{\lambda} \log x\right)^2 \psi \left(\frac{d(e, x)}{\delta}\right)^2 dx.$$  

Then

(1) $\lim_{\lambda \to \infty} \lambda^{d/2} E_\lambda$ converges,

$$\int_{P_e(G)} f_\phi(\gamma) d\nu_\lambda(\gamma) = \int_{\mathbb{P}_a(G)} f(\gamma)^2 d\nu_{a, \lambda}(\gamma), \quad (3.8)$$

(2) For sufficiently large $\lambda > 0$, it holds that $\nabla(f_\phi(\gamma)) = I_1 + I_2 + I_3$, where

$$I_1 = T(\gamma)^* ((\nabla f)(S(\gamma))) \Phi(\gamma) \bar{\phi}(\gamma), \quad I_2 = \frac{\lambda}{2} f(S(\gamma)) \left\{ t \left( I + \frac{1}{2} \delta \right) \log(a_\gamma(1)^{-1}) + Q^{(1)}_{\log(a_\gamma(1)^{-1})} \right\} b(1) \left\{ \nabla \phi \right\}, \quad (3.9)$$

$$I_3 = f(S(\gamma)) \Phi(\gamma) D_{\lambda}^{-1} \left\{ \sqrt{\lambda} \psi \left(\frac{d(a, \gamma(1))}{\delta}\right) \right\} \left\{ -t(I + \frac{1}{2} \delta \log(a_\gamma(1)^{-1})) + tQ^{(1)}_{\log(a_\gamma(1)^{-1})} \nabla \phi \right\}, \quad (3.10)$$

where $Q_v^{(i)}$ are linear maps satisfying that $|Q_v^{(i)}| \leq C|v|^2$.

Proof. (1) This is easily proved by that $\log: \mathbb{U}_k(e) \to \mathfrak{g}$ is a smooth one to one invertible mapping and the image measure of log is equal to the Lebesgue measure on $\mathfrak{g}$ multiplied by a smooth positive function.

(2) By Lemma 3.2 (3),

$$\int_{P_e(G)} f_\phi(\gamma)^2 d\nu_\lambda(\gamma) \leq \frac{\lambda}{2} \left[ \log(ax)^{-1} \right] b(1) - \frac{\lambda}{2} |\log(ax) + 1| \right\} d\nu_\lambda(\gamma) D_{\lambda}^{-2} \quad (3.11)$$

$$= \int_{P_e(G)} \delta a(1) f(\gamma(1)) \left(1 + \frac{1}{2} \delta \right) \right\} d\nu_\lambda(\gamma) D_{\lambda}^{-2} \quad (3.12)$$

$$= \int_{P_e(G)} \delta a(1) f(\gamma(1)) \left(1 + \frac{1}{2} \delta \right) \right\} d\nu_\lambda(\gamma) D_{\lambda}^{-2} \quad (3.13)$$

$$= E_\lambda \int_{P_e(G)} \phi \left(\sqrt{\lambda} \log(x)\right)^2 \psi \left(\frac{d(e, x)}{\delta}\right)^2 \left(1 + \frac{1}{2} \delta \right) \right\} f(\gamma)^2 d\nu_{a, \lambda}(\gamma). \quad (3.9)$$
In the last equality we have used the invariance of the Riemannian volume (3.1).
(3) This follows from (3.2) and Lemma 3.6. □

We prove Theorem 3.3.

Proof. The proof is similar to Gross’s inheritance method. Note that
\[
\int_{P_x(G)} f_\phi(\gamma)^2 \log \left( f_\phi(\gamma)^2 \right) d\nu_\lambda(\gamma)
= \int_{P_x(a)} f(\gamma)^2 \log f(\gamma)^2 d\nu_{\lambda,a}(\gamma) + \int_{P_x(G)} f_\phi(\gamma)^2 \log \left( f_\phi(\gamma)^2 \right) d\nu_\lambda(\gamma)
+ \int_{P_x(G)} f_\phi(\gamma)^2 \log \left( \phi \left( \sqrt{\lambda} \log(a_\gamma(1)^{1-1}) \right)^2 \psi \left( \frac{d(a_\gamma(1))}{\delta} \right)^2 \right) d\nu_\lambda(\gamma)
- 2 \log D_\lambda \int_{P_x(a)} f(\gamma)^2 d\nu_{\lambda,a}(\gamma)
= J_1(\lambda) + J_2(\lambda) + J_3(\lambda) + J_4(\lambda).
\]

We estimate each terms. By the calculation similar to (3.8),
\[
J_2(\lambda) = \int_G D_\lambda^{-2} \phi \left( \sqrt{\lambda} \log x \right)^2 \psi \left( \frac{d(e,x)}{\delta} \right)^2 J_2(\lambda,x) dx
\]
\[
J_2(\lambda,x) = \int_{P_x(G)} f(\gamma)^2 \left\{ -\lambda(\log x,b(1)) - \frac{\lambda}{2} |\log x|^2 \right\} \delta_{\lambda}(\gamma) d\nu_\lambda(\gamma).
\]

By \( \log(x^{-1}) = -\log x \), \( \phi(v) = \phi(-v) \) and the bi-invariance of \( dx \), we see that the integral of the term containing \( \log(x,b(1)) \) is zero. Thus, we have \( J_2(\lambda) \geq -\frac{\lambda}{2} C \|f\|_{L^2(P_x(a,G))}^2 \). By the similar calculation to (3.9) and the fact in (1), we have
\[
J_3(\lambda) = E_x^{-1} \int_G \phi \left( \sqrt{\lambda} \log x \right)^2 \psi \left( \frac{d(e,x)}{\delta} \right)^2 \log \left\{ \phi \left( \sqrt{\lambda} \log x \right)^2 \psi \left( \frac{d(e,x)}{\delta} \right)^2 \right\} dx
\]
\[
\cdot \int_{P_x(a)(G)} f(\gamma)^2 d\nu_{\lambda,a}(\gamma) \geq -C \int_{P_x(a)(G)} f(\gamma)^2 d\nu_{\lambda,a}(\gamma),
\]

where \( C \) is a positive constant. Again by the result in (1), there exists a positive constant \( C \) such that
\[
J_4(\lambda) \geq -\log \left( C \lambda^{-1/2} p(1/\lambda,e,a) \right) \|f\|_{L^2(P_x(a,G))}^2.
\]

Therefore, by Gross’s LSI on \( P_x(G) \) [13],
\[
\int_{P_x(a)(G)} f^2 \log \left( \frac{f^2}{\|f\|_{L^2(P_x(a,G))}^2} \right) d\nu_{\lambda,a}
\leq \frac{2}{\lambda} \int_{P_x(G)} |(\nabla f_\phi)(\gamma)|^2 d\nu_\lambda(\gamma) + \left\{ \log \left( C_1 \lambda^{-1/2} p(1/\lambda,e,a) \right) + C_2 \right\} \|f\|_{L^2(P_x(a,G))}^2.
\]
Next we estimate the integral of \(|(\nabla f(x))(\gamma)|^2\). By Lemma 3.8 (3), we need to estimate \(|I_1|^2\) and the cross terms \((I_1, I_j)\). In the calculation below, we denote \(\phi_\lambda(x) = \phi\left(\sqrt{\lambda} \log x\right)^2 \psi\left(\frac{s(x)}{\lambda}\right)^2\).

1. **\(I_1\):** By the similar calculation to (3.9), we have

\[
\int_{P_e(P_c(G))} |I_1|^2 d\nu_\lambda(\gamma) \leq \left(1 + \frac{C}{\lambda}\right) \int_{P_e(P_c(G))} |(\nabla f(x))(\gamma)|^2 d\nu_\lambda(\gamma) + \int_G dx \int_{P_e(P_c(G))} 2\phi_\lambda(x) (B_{\log x} (\nabla f) (\gamma), (\nabla f) (\gamma)) d\nu_\lambda(\gamma).
\]

Since \(B_{\log(x^{-1})} = -B_{\log(x)}\) and \(\phi_\lambda(x) = \phi_\lambda(x^{-1})\), the second integral is 0.

2. **\(I_2\):** By Lemma 3.2 (1) and (2),

\[
\int_{P_e(P_c(G))} |I_2|^2 d\nu_\lambda(\gamma) = \frac{\lambda^2}{4E_A} \int_{P_e(P_c(G))} \int_G I_2(x, \gamma) \phi_\lambda(x) f(\gamma)^2 dx d\nu_\lambda(\gamma).
\]

Then

\[
I_2(x, \gamma) = \int_0^1 b(1) + \int_0^1 (Ad (e^{-s \log x}) - I) db(s) - 2 \log x + \frac{1}{2} \text{ad}(\log x) \left(\int_0^1 Ad (e^{-u \log x}) db(u)\right)

+ \frac{1}{2} Q_{\log x}^2 \left(\int_0^1 Ad (e^{-s \log x}) db(s)\right)

- \text{ad}(\log x) \left(\int_0^1 Ad (e^{-u \log x}) db(u) - t \log x\right) \right|^2 dt.
\]

Then

\[
I_2(x, \gamma) \leq |b(1)|^2 + C|\log x|^2 \left\{1 + |b(1)|^2 + \left(\int_0^1 |b(s)| ds\right)^2\right\} + \lambda \log x(\gamma),
\]

where \(v \rightarrow \lambda(\gamma)\) is a linear function. By the similar reason to \(I_1\),

\[
\frac{\lambda^2}{4E_A} \int_G I_2(x, \gamma) \phi_\lambda(x) dx \leq \frac{\lambda^2}{4} |b(1)|^2 + C\lambda \left\{1 + |b(1)|^2 + \left(\int_0^1 |b(s)| ds\right)^2\right\}.
\]

Therefore, we have

\[
\int_{P_e(P_c(G))} |I_2|^2 d\nu_\lambda(\gamma) \leq \int_{P_e(P_c(G))} \frac{\lambda^2}{4} |b(1)|^2 f(\gamma)^2 d\nu_\lambda(\gamma)

+ \lambda C \int_{P_e(P_c(G))} \left\{1 + |b(1)|^2 + \left(\int_0^1 |b(s)| ds\right)^2\right\} f(\gamma)^2 d\nu_\lambda(\gamma).
\]

3. **\(I_3\):** We have \(|I_3|^2 \leq C \lambda \|f\|_{L^2(P_e(P_c(G)), \nu_\lambda(\gamma))}^2\).

4. The cross term \((I_2, I_3)\).
Since $\phi^2(\xi) = \phi^2(-\xi)$, we have $(\nabla \phi^2)(\xi) = -(\nabla \phi^2)(-\xi)$. By the invariance of the Haar measure,

$$
\int_G (\nabla \phi^2) \left( \sqrt{\lambda} \log x \right) \psi \left( \frac{d(e, x)}{\delta} \right)^2 dx
= \int_G (\nabla \phi^2) \left( -\sqrt{\lambda} \log x \right) \psi \left( \frac{d(e, x)}{\delta} \right)^2 dx.
$$

Therefore, this integral value is 0. Thus, the integral of the term containing $\left( b(1), \nabla \phi \left( \sqrt{\lambda} \log(\alpha \gamma(1)-1) \right) \right)$ is 0. Consequently,

$$
\int_{P_{e,a}(G)} (I_1, I_3) d\nu_{\lambda,a}(\gamma)
\leq C \lambda \int_{P_{e,a}(G)} \left\{ 1 + |b(1)|^2 + \left( \int_0^1 |b(s)| ds \right)^2 \right\} f(\gamma)^2 d\nu_{\lambda,a}(\gamma).
$$

(5) The cross terms: $(I_1, I_2), (I_1, I_3)$:

To estimate these terms, we use the integration by parts formula (3.3) and the fact that $T(\gamma) \eta = A_{\log(\alpha \gamma(1)-1)}(t) \eta(t)$ for $\eta(t) = t \xi$. Here $A_{\cdot}$ is the operator in (3.7). The calculations are similar to the previous terms. So we omit the details. Thus, we have

$$
\int_{P_{e,a}(G)} (I_1, I_1) d\nu_{\lambda,a}(\gamma)
\leq C \lambda \int_{P_{e,a}(G)} \left\{ 1 + |b(1)|^2 + \left( \int_0^1 |b(s)| ds \right)^2 \right\} f(\gamma)^2 d\nu_{\lambda,a}(\gamma).
$$

Combining all the above estimates, we have

$$
\int_{P_{e,a}(G)} f^2(\gamma) \log \left( \frac{f^2(\gamma)}{\|f\|^2_{L^2(P_{e,a}(G))}} \right) d\nu_{\lambda,a}(\gamma)
\leq \frac{2}{\lambda} \left( 1 + \frac{C_1}{\lambda} \right)
\times \left\{ \int_{P_{e,a}(G)} |(\nabla f)(\gamma)|^2 d\nu_{\lambda,a}(\gamma) + \int_{P_{e,a}(G)} \frac{\lambda^2}{4} |b(1)|^2 f(\gamma)^2 d\nu_{\lambda,a}(\gamma) \right\}
+ \left( \log(C_2 \lambda^{-d/2} p(1/\lambda, e, a)) \right) \|f\|^2_{L^2(P_{e,a}(G))}
+ \int_{P_{e,a}(G)} 2C_3 \left\{ 1 + |b(1)|^2 + \left( \int_0^1 |b(s)| ds \right)^2 \right\} f(\gamma)^2 d\nu_{\lambda,a}(\gamma)
$$

which completes the proof. 

4. Exponential Integrability

It is crucial to check on which domains $e^{\frac{d}{2} |b(1)|^2}$ is integrable to use the lower bound estimate (2.3). To this end, we introduce the following:
Definition 4.1. For $\gamma, \eta \in P_e,\alpha(G)$, define $d(\gamma, \eta) = \max_{0 \leq t \leq 1} d(\gamma(t), \eta(t))$. We denote

$$\Omega_L = \{ \gamma \in P_e,\alpha(G) \mid \sqrt{E(\gamma)} \leq L \},$$

$$B_\varepsilon(\eta) = \{ \gamma \in P_e,\alpha(G) \mid d(\gamma, \eta) < \varepsilon \},$$

$$\Omega_{L,\varepsilon} = \{ \gamma \in \Omega_L \mid \text{there exists } \eta \in \Omega_L \text{ such that } \gamma \in B_\varepsilon(\eta) \}.$$ Also we define for $0 < \alpha < 1$,

$$\|\gamma\|_\alpha = \sup_{0 \leq s,t \leq 1} \frac{d(\gamma(t), \gamma(s))}{|t-s|^\alpha}.$$

**Lemma 4.2.** (1) For any $\varepsilon > 0$, it holds that $\lim_{L \to \infty} \nu_{\lambda,\alpha}(\Omega_{L,\varepsilon}) = 1$.

(2) We denote $l(\gamma) = \int_0^1 |\dot{\gamma}(t)| dt$. Take a finite energy path $\gamma_0 \in P_{e,\alpha,H^1}(G)$. Let $p > 0$. There exist $C_i > 0$ $(1 \leq i \leq 4)$ and $m \in \mathbb{N}$ such that, for $\varepsilon < (18pC_1)^{-1/2} =: \varepsilon_p$, it holds that

$$\int_{B_\varepsilon(\gamma_0)} \exp \left( \frac{p\lambda}{2} |b(1)|^2 \right) d\nu_{\lambda,\alpha}(\gamma) \leq \frac{C_2 \lambda^{m+(d/2)}}{\sqrt{1 - 18p^2C_3}} \exp \left\{ \frac{3p\lambda}{2} (|b(1,\gamma_0)|^2 + C_3 + C_4 \varepsilon^2 l(\gamma_0)^2) \right\},$$

The constants $C_i$ and $m$ are independent of $p, \varepsilon, \gamma_0$.

(3) Let $p > 0$ and set $\varepsilon < \varepsilon_p/2$. There exist positive numbers $C_i$ $(i = 5, 6)$ such that the following inequality holds. $C_5$ depends on $L$ and $\varepsilon$ but $C_6$ is independent of $L$.

$$\int_{\Omega_{L,\varepsilon}} \exp \left( \frac{p\lambda}{2} |b(1)|^2 \right) d\nu_{\lambda,\alpha}(\gamma) \leq \frac{C_6 \lambda^{m+(d/2)}}{\sqrt{1 - 72p^2C_7}} e^{p\lambda C_6(1+L^2)}.$$ (4.1)

Also for any $M$ and $p$, we have $\int_{\|\gamma\|_\alpha < M} e^{\frac{1}{\varepsilon} |b(1)|^2} d\nu_{\lambda,\alpha}(\gamma) < \infty$.

Proof. (1) Note that for $0 < \alpha < 1/2$, $\nu_{\lambda,\alpha}(\|\gamma\|_\alpha < \infty) = 1$ holds. Take a positive number $M$. Note that if $d(x,y) \leq i(G)$ $(x,y \in G)$, then $x$ and $y$ are joined by the unique minimal geodesic. Take a positive integer $N$ such that $MN^{-\alpha} < i(G)$. Suppose that $\gamma$ satisfies that $\|\gamma\|_\alpha < M$. Let $t_k = k/N$. Since $d(\gamma(t_{k+1}), \gamma(t_k)) \leq MN^{-\alpha}$, $\log (\gamma(t_{k+1})\gamma(t_k)^{-1})$ is well-defined. Let $\gamma_N$ be the piecewise geodesic path such that $\gamma_N(t) = \gamma(t_k)$ for $t = t_k$ $(0 \leq k \leq N)$ and $\gamma_N(t) = \exp \left( \frac{t-t_k}{t_{k+1}-t_k} \log (\gamma(t_{k+1})\gamma(t_k)^{-1}) \right) \gamma(t_k)$ for $t_k < t < t_{k+1}$. Then for $t_k \leq t \leq t_{k+1}$,

$$d(\gamma_N(t), \gamma(t)) \leq d(\gamma_N(t), \gamma_N(t_k)) + d(\gamma_N(t_k), \gamma(t_k)) + d(\gamma(t_k), \gamma(t)) \leq 2MN^{-\alpha}.$$ Furthermore, we take $N$ such that $2MN^{-\alpha} < \varepsilon$, that is, $N > \max \left( \frac{M}{(\frac{\varepsilon}{\alpha})^{1/\alpha}}, \frac{2M}{\varepsilon} \right)^{1/\alpha}$.

Finally note that $E(\gamma_N)^{1/2} \leq MN^{1-\alpha}$. Hence for $L > M \max \left( \frac{1}{(\frac{\varepsilon}{\alpha})^{1/\alpha}}, \frac{2\varepsilon}{\alpha} \right)^{\frac{1}{\alpha}-1}$, it holds that $\{ \gamma \mid \|\gamma\|_\alpha < M \} \subset \Omega_{L,\varepsilon}$. This proves (1).
(2) We consider $b(1)$ in $B_{\varepsilon}(\gamma_0)$. Note that the calculation below should be understood as the calculation under $\nu_\lambda$ at first. After that it can be extended to pinned measure case too by the quasi-sure analysis.

By the Itô formula,

$$b(1) = \int_0^1 d\gamma(t) \circ \gamma(t)^{-1}$$

$$= \int_0^1 \dot{\gamma}_0(t)\gamma(t)^{-1} dt + \int_0^1 (\gamma(t) - \gamma_0) \circ \gamma(t)^{-1} dt$$

$$= \int_0^1 \dot{\gamma}_0(t)\gamma(t)^{-1} dt - \int_0^1 (\gamma(t) - \gamma_0) \circ d(\gamma(t)^{-1}).$$

By Lemma 3.1,

$$b(1) = \int_0^1 \dot{\gamma}_0(t)\gamma(t)^{-1} dt + \int_0^1 (\gamma(t) - \gamma_0) \circ \gamma(t)^{-1} dt$$

$$- \frac{1}{2\lambda} \int_0^1 (\gamma(t) - \gamma_0) \gamma(t)^{-1} C dt$$

$$= b(1, \gamma_0) + \int_0^1 \dot{\gamma}_0(t) (\gamma(t)^{-1} - \gamma_0(t)^{-1}) dt$$

$$- \frac{1}{2\lambda} \int_0^1 (\gamma(t) - \gamma_0) \gamma(t)^{-1} C dt + M(1, \gamma),$$

where $M(t, \gamma) = \int_0^t (\gamma(s) - \gamma_0(s)) \gamma(s)^{-1} ds$.

Noting that $|A^{-1} - B^{-1}| \leq |A^{-1}| |A - B| |B^{-1}|$, we have

$$|b(1)| \leq |b(1, \gamma_0)| + \frac{C_1 \varepsilon}{2\lambda} + C_2 \varepsilon l(\gamma_0) + |M(1, \gamma)|.$$

Therefore,

$$|b(1)|^2 \leq 3 \left(|b(1, \gamma_0)|^2 + C^2 \varepsilon^2 (1 + l(\gamma_0))^2 + |M(1, \gamma)|^2\right).$$

Hence using the lower bound estimate on $p(t, \varepsilon, a)$,

$$\int_{B_{\varepsilon}(\gamma_0)} e^{\frac{\varepsilon}{2A}|b(1)|^2} d\lambda, a(\gamma)$$

$$\leq \exp \left[ \frac{3p\lambda}{2} \left(|b(1, \gamma_0)|^2 + C^2 \varepsilon^2 (1 + l(\gamma_0))^2\right) \right]$$

$$\times p(1/\lambda, \varepsilon, a)^{-1} J_\lambda$$

$$\leq C \lambda^{d/2} \exp \left[ \frac{3p\lambda}{2} \left(|b(1, \gamma_0)|^2 + C^2 (1 + \varepsilon l(\gamma_0))^2\right) \right] J_\lambda, \quad (4.2)$$

where

$$J_\lambda = \int_{B_{\varepsilon}(\gamma_0)} \Psi_\lambda(b) \delta_a(\gamma(1, b)) d\mu_\lambda(b),$$
we use very rough estimate: there exists $4.3$ such that many geodesics diverges. Let us prove this in the case where $\mu_k$ denotes the diagonal matrix whose $(v)$

For each integral on the right-hand side can be estimated as in (4.2) by replacing $\mu$. Here we use very rough estimate:

$$J_\lambda \leq \int_B \Psi_\lambda(b) \delta_a(\gamma(1,b)) d\mu(b)$$

$$\leq \frac{C\lambda^n}{\sqrt{1-18peC}}.$$  

This estimate follows from the integration by parts formula and an estimate on exponential martingale. The natural number $m$ depends on how many times we apply the integration by parts formula on Wiener space. 

(3) Since $\Omega_L$ is a compact subset in $P_{e,a}(G)$, there exists a finite set of smooth curves $\{\gamma_i\}_{i=1}^N \subset \Omega_L$ such that $\Omega_L \subset \cup_{i=1}^N B_\varepsilon(\gamma_i)$. This implies $\Omega_{L,\varepsilon} \subset \cup_{i=1}^N B_{2\varepsilon}(\gamma_i)$. Hence

$$\int_{\Omega_{L,\varepsilon}} e^{\frac{p\lambda}{2}|b(1)|^2} d\nu_{\lambda,a}(\gamma) \leq \sum_{i=1}^N \int_{B_{2\varepsilon}(\gamma_i)} e^{\frac{p\lambda}{2}|b(1)|^2} d\nu_{\lambda,a}(\gamma).$$

For each integral on the right-hand side can be estimated as in (4.2) by replacing $\gamma_0$ by $\gamma_i$. Note that $N$ depends on $L$ and $\varepsilon$. This and (2) proves (4.1). The last statement follows from (4.1) and the fact which we proved in the proof of (1).

Remark 4.3. In general, $\int_{P_{e,a}(G)} e^{\lambda \frac{|b(1)|^2}{2}} d\nu_{\lambda,a}(\gamma) = +\infty$. Actually, we can prove a stronger statement: The integral of $e^{\lambda \frac{|b(1)|^2}{2}}$ in a neighborhood of countably many geodesics diverges. Let us prove this in the case where $G = SU(n)$. Let $v_0$ be an element of $g$ such that $l_0(t) = e^{tv_0}$ is a geodesic between $e$ and $a$. Then there exists $g \in G$ such that $g v_0 g^{-1} = D[\sqrt{-1}\gamma_1, \ldots, \sqrt{-1}\gamma_n]$, where $D[a_1, \ldots, a_n]$ denotes the diagonal matrix whose $(i,i)$-element is $a_i$. Take $v_i = g^{-1} D[(\eta_i + 2\pi k_1^{(i)})/\sqrt{-1}, \ldots, (\eta_n + 2\pi k_n^{(i)})/\sqrt{-1}]$, where $(k^{(i)})_{1 \leq j \leq n} \in \mathbb{Z}^n$ denotes the distinct points of integer lattice satisfying $\sum k_j^{(i)} = 0$. Let $l_i(t) = e^{tv_i}$. Then $\{l_i\}_{i=1}^\infty$ are distinct geodesics joining $e$ and $a$. Let $p > 0$. Take a sufficiently small positive number $\varepsilon$. By Lemma 4.2 (2), it holds that

$$\int_{B_\varepsilon(l_i)} \exp \left( \frac{p\lambda}{2} \frac{|b(1)|^2}{2} \right) d\nu_{\lambda,a}(\gamma) < \infty.$$  

Also we have $B_\varepsilon(l_i) \cap B_\varepsilon(l_j) = \emptyset$ for sufficiently small $\varepsilon$ since

$$B_\varepsilon(l_i) = \{e^{tv_i} : \gamma(t) \mid \gamma \in B_\varepsilon(l_j)\}.$$
for any $i, j$. By Lemma 3.2 (2), we can prove that
\[
\int_{B_r(t_i)} \exp \left( \frac{p \lambda}{2} |b(1)|^2 \right) d\nu_{\lambda,a}(\gamma)
\]
\[
= \int_{B_r(t_0)} \exp \left\{ \frac{\lambda}{2} \left( \int_0^1 Ad(e^{s(v_i-v_0)})db(s)^2 + (p - 1) \left| \int_0^1 Ad(e^{s(v_i-v_0)})db(s) + v_i - v_0 \right|^2 \right) \right\} d\nu_{\lambda,a}(\gamma).
\]
This shows that
\[
\int_{B_r(t_i)} \exp \left( \frac{\lambda}{2} |b(1)|^2 \right) d\nu_{\lambda,a}(\gamma) \geq \nu_{\lambda,a}(B_r(t_0)).
\]
Thus it holds that
\[
\int_{P_{\lambda,a}(G)} \exp \left( \frac{\lambda}{2} |b(1)|^2 \right) d\nu_{\lambda,a}(\gamma) > \int_{\cup_{i=1}^\infty B_r(t_i)} \exp \left( \frac{\lambda}{2} |b(1)|^2 \right) d\nu_{\lambda,a}(\gamma) = +\infty.
\]

References

Shigeki Aida: Department of Mathematical Science, Graduate School of Engineering Science, Osaka University, Toyonaka, 560-8531, JAPAN
E-mail address: aida@sigmath.es.osaka-u.ac.jp
PRICING FUNCTIONALS AND PRICING MEASURES

ERIC HILLEBRAND AND AMBAR N. SENGUPTA

Abstract. We demonstrate how pricing functionals give rise to pricing measures, using a time-independent framework. For infinite market state spaces, the Gel’fand spectral theory for \( C^\ast \)-algebras is used to obtain the pricing measure. Pricing functionals with additional market information are shown, within this model, to be given by conditional expectations. Our approach is time-independent, and the usual martingale property of prices appears as a special case in our method.

1. Introduction

The purpose of this paper is to present a method of constructing pricing measures that uses the Gel’fand theory of commutative \( C^\ast \)-algebras. This machinery has been used in quantum and statistical physics, but appears to be novel in the present economic context. Mathematically, the main results, especially Theorems 5.1 and 6.2, are in sections 5 and 6.

In addition to the specific mathematical results we also wish to emphasize that the more geometric approach of economic equilibrium theory as in Debreu [4] provides a conceptually clearer and more general framework for mathematical modelling of financial markets than the approach that has now become standard (and is summarized, for instance, in Karatzas and Shreve [12, 13]), which is heavily influenced by stock and bond market instruments. To view the fundamental pricing measure as a ‘martingale measure’, though certainly correct, places the time coordinate in a special role in the basic framework that is not necessary. In our presentation, the martingale feature is a special case of a more general property of prices: the equilibrium market price of an asset under partial information about the market is the corresponding conditional expectation of the price. A special case of this is obtained by taking ‘partial information’ to mean all information available till a given time \( t \), and this yields the usual martingale property of prices. In brief, our framework is time-independent. (This point of view is developed more fully in [15].) While this point of view is not new (indeed it is more classical, based on economic equilibrium theory), it departs from the current orthodoxy. To make an analogy with quantum physics, if the standard stochastic framework for financial markets is compared with the Schrödinger picture wave function in coordinate

2000 Mathematics Subject Classification. Primary 91B24, Secondary 46A22.

Key words and phrases. arbitrage, Gel’fand transform, Hahn-Banach, martingale measure, pricing, Riesz representation.

* The research of A. Sengupta was supported by US National Science Foundation Grants DMS-0201683 and DMS-0601141.
space for a system of particles in space, then the point of view we are stressing
corresponds to the general Hilbert space foundations of quantum theory. Example
2.1 in section 2, and comments relating to it at several places later, explain how
the ‘standard’ framework is a special case of ours.

Thus, this paper is concerned with finding a more general approach to the
so-called first fundamental theorem of asset pricing, which essentially says that
absence of arbitrage opportunities is equivalent to existence of equivalent martingale
measures. The martingale ansatz was developed in Harrison and Kreps [10]
and Harrison and Pliska [11], spawned an extensive literature, and has found a
very general formulation in Delbaen and Schachermayer [5, 6].

A consequence of the focus on the first fundamental theorem is that we are
mostly concerned with the absence of arbitrage and not so much with completeness
of the market. Broadly speaking, completeness of a market is connected with
uniqueness of the pricing measure. Completeness can also be understood in very
simple terms as the possibility of replication of any payoff function by a suitable
combination of a set of “basic” financial instruments. Market completeness is
treated in what is often called the second fundamental theorem of asset pricing,
the statement that a market is complete if and only if the equivalent martingale
measure is unique [10, 11]. Battig and Jarrow [3] propose a decoupling of the issue
of completeness from the martingale approach. A non-technical account of the
issues addressed in the second fundamental theorem is provided in Flood [8].

The fundamental relationship between price and probability has been known
since the earliest formulations of probability theory in terms of gambling returns.
Briefly put, if \( I_B \) is an instrument or asset which yields one unit of time-\( t \) cash
(or any chosen numeraire) in case event \( B \) occurs and yields nothing if \( B \) doesn’t
occur then

\[
\text{price of asset } I_B = \text{probability } Q(B) \text{ of event } B \quad (1.1)
\]

where probability is assessed by the trader who is pricing the asset. For a general
asset whose worth is \( f(\omega) \) in market scenario \( \omega \), the corresponding equation is

\[
\text{price of asset described by } f = \text{the expected value } \int f \, dQ \text{ of } f \quad (1.2)
\]

In ideal market equilibrium all traders agree on a common price for each traded
instrument and this gives rise to a common assessment of probability, which is the
market equilibrium measure. This measure describes the market’s view of
probabilities of events. Needless to say, no individual trader may truly agree with
this ideal measure which is the result of consensus emerging out of trading rather
than “real-life” probabilities (whatever that may be). Being based on the price
at which a risky asset would be exchanged for another, this measure is the “risk-
neutral measure” for the market.

There is one problem with the above analysis relating prices to probabilities:
instruments like \( I_B \) will certainly not actually exist in the real market (though of
course such “digital option” instruments would exist for certain types of events \( B \)).
Thus the pricing measure \( Q \) will have to be imputed from prices of the instruments
that are actually being traded in a market. This raises the theoretical question as
to whether such a measure exists at all:
given prices of traded instruments, is there a probability measure $Q$ on the market scenarios such that the prices result from $Q$ as expected values as in (1.2)?

Our work is devoted to this question, and here is a summary of what we do:

- in section 2 we state the formal model we are considering, essentially that of a market with a set of traded instruments with a consistent set of prices (consistency includes a no-arbitrage condition);
- in section 3 we show that the answer to question (1.3) is yes if the market state-space is finite; this is of course a standard fact (see for instance Duffie [7, Chapter 1]), but there may be some value in the way we formulate the proof;
- in section 4 we give an example where the market state-space is infinite where the answer to (1.3) is no;
- in section 5, which is our main focus, we show how the market state-space, even if infinite, may be extended mathematically so that the answer to (1.3) is yes;
- in section 6, we show how price functionals in the presence of market information are described by conditional expectations and how, as a very special case, this leads to the martingale property of market equilibrium prices. The results here are well-known but we take an approach that conforms to our model.

Expanding a state-space so that a probability measure can live on it is a common procedure in stochastic analysis (for example in the construction of Wiener measure) and in areas of physics. The method we have used in this paper is based on the Gel’fand theory of commutative $C^*$-algebras, a technique which has also been used in constructing measures in quantum field theories and statistical mechanics/thermodynamics. Sometimes projective systems of measure spaces are used to obtain limiting measures as an alternative to the Gel’fand transform method.

In [2], Balbás et al. have used a projective system of measure spaces to obtain a measure corresponding to a no-arbitrage system of prices. As in our method, they also need to expand the market state space in order to obtain the measure. The specific details of the framework are different, and we work in a more abstract setting, but the underlying issues are closely related.

2. The Model

The concepts we are formalizing are:

(i) a market which can exist in a certain set of states,
(ii) assets or instruments which have prices, in units of any chosen numeraire, in each market state,
(iii) a trader who, without exact information about the market state, associates prices to assets.
In general different traders would associate different prices to the same asset. Trading takes place when the bid and ask prices agree and so in ideal market equilibrium all traders would be using the same price mechanism.

We denote by \( \Omega \) the set of all market states. It is important to understand that an element of \( \Omega \) may be thought of as describing one entire time evolution path for the market. However, our framework is flexible enough to also permit the more restrictive interpretation of an element of \( \Omega \) being a market state at one particular time. (Such a dichotomy occurs also in quantum theory, where one has the Heisenberg picture versus the Schrödinger picture for states of a system.)

Again, let us stress that we do not need to single out the time coordinate in setting up the framework.

**Example 2.1.** For applications to stock and bond market instruments, one could take \( \Omega \) to be the continuous path space \( C_0([0, T], \mathbb{R}^N) \) (paths beginning at \( 0 \in \mathbb{R}^N \)) to model a market with \( N \) underlying factors, all evolving over a time period \([0, T]\). An element \( \omega \in \Omega \) is then a path \( t \mapsto \omega(t) \in \mathbb{R}^N \). The state-dependent discount factor for time \( t \) is \( e^{-\int_0^t r(\omega; u) \, du} \), where \( r(\omega; u) \) is the risk-free interest rate at time \( u \) for market state \( \omega \). We will view assets as having a time-stamp; thus, a particular stock at a specific time \( t \in [0, 1] \) has worth \( X(\omega; t) \) at time \( t \) in state \( \omega \). In time-0 money, converted to time-0 money, this asset’s worth is described by the function \( \omega \mapsto e^{-\int_0^t r(\omega; u) \, du} X(\omega; t) \).

Each traded asset has a specific worth (in units of a chosen numeraire) in each market state \( \omega \). Two assets which always have exactly the same worth in every scenario will be viewed as being effectively the same asset. Thus an asset can be modelled as a mapping \( f : \Omega \to \mathbb{R} : \omega \mapsto f(\omega) \), with \( f(\omega) \) denoting the value of the asset \( f \) in market state \( \omega \).

As with the state space, there are two possible formulations here. If an element of \( \Omega \) represents a market state through all time (as in Example 1.1 above), then the same physical asset/instrument will be described by different functions at different times. On the other hand, if an element of \( \Omega \) represents a possible market state at any particular time, then it is the state which evolves in time through \( \Omega \) while the asset functions remain the same.

We will work only with a class of assets for which the prices are given by bounded functions \( f \). This is a technical assumption with no larger significance, as the price functional extends uniquely to unbounded functions, given some simple continuity assumption.

Different assets may be combined to form portfolios and we also assume that each asset can be scaled in any way. Thus the set of assets forms a vector space \( V \). The sum of \( f, g \in V \) is simply the pointwise sum \( f + g \), the function \( \Omega \to \mathbb{R} \) whose value at any state \( \omega \) is \( f(\omega) + g(\omega) \). If \( f \in V \) and \( k \) is any real number then \( kf \) is the function on \( \Omega \) whose value at any \( \omega \) is \( kf(\omega) \).

Lastly, we have the trader. The trader is not assumed to know which market state \( \omega \) actually prevails but, based perhaps on his/her understanding or estimate
### Table 1. Pricing Framework.

<table>
<thead>
<tr>
<th>Mathematical object</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A set $\Omega$</td>
<td>elements of $\Omega$ correspond to market states</td>
</tr>
<tr>
<td>A vector space $V$</td>
<td>a function $f : \Omega \to \mathbb{R}$ corresponds to an instrument, with $f(\omega)$ being the price of $f$ in market state $\omega$</td>
</tr>
<tr>
<td>A linear functional $L : V \to \mathbb{R}$</td>
<td>$L(f)$ is the trader’s a priori price for the instrument $f$</td>
</tr>
<tr>
<td>$\min f \leq Lf \leq \max f$ for all $f \in V$</td>
<td>this is the no-arbitrage condition</td>
</tr>
</tbody>
</table>

of which market state $\omega$ will be realized, associates to each asset $f$ a price $Lf$ the trader is willing to pay or receive in exchange for the asset. Clearly, $Lf$ should be linear in $f$, i.e.

$$L(f + g) = L(f) + L(g) \text{ and } L(kf) = kL(f) \text{ for any constant scaling } k.$$  

Thus the trader’s role is summarized by his/her pricing procedure which is given by a linear functional

$$L : V \to \mathbb{R}$$

which we shall call the *pricing functional*.

Linearity of the pricing functional can be viewed as a form of no-arbitrage (or as lack of friction [11]).

Now we formalize the notion of no-arbitrage. An arbitrage opportunity would allow a market participant to purchase/short some asset $h$ which would yield a sure profit in some scenario, with no risk of loss in any scenario, based on the available price $Lh$. Thus an asset $h$ provides an arbitrage for the trader if the price $Lh$ is strictly less than all possible values $h(\omega)$ or is strictly greater than all possible values $h(\omega)$. Thus non-existence of arbitrage means that $L$ has the property that

$$\min f \leq Lf \leq \max f \text{ for all } f \in V \quad (2.1)$$

A stricter form would require that $\min f < Lf < \max f$ unless $f$ is constant. There are, of course, other ways and contexts to formalize no-arbitrage. In some sense, even the linearity condition on $L$ is a no-arbitrage condition.

Our model is summarized in Table 1. Note that the condition $\min f \leq Lf \leq \max f$ for all $f \in V$ is equivalent to $Lf \leq \max f$ for all $f \in V$ (just switch $f$ to $-f$ in the latter condition to obtain $Lf \geq \min f$).
3. From Pricing Functional to Pricing Measure for Finite $\Omega$

The results in this section are standard and well-known, but we include them for ease of readability and reference in the following sections. Proofs may be obtained by consulting, for example, Rudin’s standard text [14]; complete proofs of the specific results are in Sengupta [15].

The set of all functions $\Omega \rightarrow \mathbb{R}$ is denoted $\mathbb{R}^\Omega$. This is a real vector space by pointwise operations. Indeed if $\Omega$ is finite with $N$ elements then $\mathbb{R}^\Omega$ is essentially $\mathbb{R}^N$. This a linear space $V$ of real-valued functions on $\Omega$ is a subspace of $\Omega \rightarrow \mathbb{R}$.

Our goal is here is to show that if there are finitely many market states then the pricing functional $L$ arises from a measure $Q$ on the subsets of $\Omega$. The argument is in two steps: (i) first we show how to extend $L$ from $V$ to a linear functional on all of $\mathbb{R}^\Omega$ and (ii) note that every linear functional on $\mathbb{R}^\Omega$ arises by integration with respect to a measure.

The first step is contained in:

**Theorem 3.1.** Let $V$ be a subspace of $\Omega \rightarrow \mathbb{R}$ consisting of bounded functions, where $\Omega$ is a non-empty set, and $L : V \rightarrow \mathbb{R}$ a linear functional satisfying the condition $Lf \leq \max f$ for all $f \in V$. Then there is a linear functional $L' : \mathbb{R}^\Omega \rightarrow \mathbb{R}$ which coincides with $L$ on the subspace $V$ and satisfies $\min f \leq L'f \leq \max f$ for all $f \in \mathbb{R}^\Omega$.

Compare to Theorem 1 of Harrison and Kreps [10]. This is a consequence of the following slightly more general result (We omit the proof of this standard result for brevity.):

**Theorem 3.2.** Let $Z$ be a real vector space on which there is a function $p : Z \rightarrow \mathbb{R}$ such that

$$p(a + b) \leq p(a) + p(b)$$

holds for every $a, b \in Z$, and $p(ta) = tp(a)$ for every $a \in V$ and $t \geq 0$. Let $V$ be a subspace of $Z$, and $L : V \rightarrow \mathbb{R}$ a linear functional satisfying the condition $Lf \leq p(f)$ for all $f \in V$. Then there is a linear functional $L' : Z \rightarrow \mathbb{R}$ which coincides with $L$ on the subspace $V$ and, moreover, satisfies $L'f \leq p(f)$ for all $f \in Z$.

To obtain Theorem 3.1 simply take $Z = \mathbb{R}^\Omega$ and $p(f) = \sup_{\omega \in \Omega} f(\omega)$ in Theorem 3.2.

The result above is essentially the Hahn-Banach theorem. The conditions imposed on the function $p : V \rightarrow \mathbb{R}$ imply that $p$ is convex, and the result above produces from the given functional $L$ a closed half-space $\{ f \in V : L'(f) \leq 1 \}$ containing the convex set $\{ f \in V : p(f) \leq 1 \}$.

In the application to the case where $V$ is a space of functions and $p(h) = \max h$, the proof of Theorem 3.2 shows that $L$ satisfies the min-max constraints:

$$\max_{h \in V} \min_{\omega \in \Omega} [Lh - \{ h(\omega) - g(\omega) \}] \leq \min_{f \in V} \max_{\omega \in \Omega} \{ f(\omega) + g(\omega) \} - Lf$$  \hspace{1cm} (3.1)

In a 2-period model, this result follows from the Separating Hyperplane Theorem (e.g., Duffie [7], p. 4).
Next, we have the result which provides the pricing measure (we quote the result from Sengupta [15, Theorem 6.3.1] for ease of reference):

**Theorem 3.3.** Let $V$ be a subspace of $\Omega \to \mathbb{R}$, where $\Omega$ is a non-empty finite set, and $L : V \to \mathbb{R}$ a linear functional satisfying the condition $Lf \leq \max f$ for all $f \in V$. Then there is a probability measure $Q$ on the set of all subsets of $\Omega$ such that

$$Lf = \int f \, dQ$$

holds for all $f \in V$.

**Proof.** By Theorem 3.2, $L$ extends to a linear functional $L' : \mathbb{R}^\Omega \to \mathbb{R}$ satisfying $L'f \leq \max f$ for all $f \in \mathbb{R}^\Omega$. From this, applied to $f$ and then to $-f$, we have

$$\min f \leq L'f \leq \max f$$

for all $f \in \mathbb{R}^\Omega$. (3.3)

In particular,

$$L'f \geq 0 \text{ whenever } f \geq 0.$$  

Taking $f$ to be the constant function 1, we have

$$L'1 = 1$$

The vector space $\mathbb{R}^\Omega$ has a finite basis consisting of the vectors $\delta_\omega$, with $\omega$ running over $\Omega$, with $\delta_\omega$ being the element of $\mathbb{R}^\Omega$ which has value 1 on $\omega$ and value 0 at all other points of $\Omega$. Any linear functional on $\mathbb{R}^\Omega$ is uniquely specified by its values on the basis elements $\delta_\omega$. Define

$$q_\omega = L'(\delta_\omega)$$

The properties of $L'$ noted above imply that $q_\omega \geq 0$ and

$$\sum_{\omega \in \Omega} q_\omega = \sum_{\omega \in \Omega} L'(\delta_\omega) = L'\left( \sum_{\omega \in \Omega} \delta_\omega \right) = L'1 = 1$$

So the numbers $q_\omega$ specify a probability measure $Q$ on the subsets of $\Omega$:

$$Q(E) = \sum_{\omega \in E} q_\omega \text{ for all } E \subset \Omega$$

(3.4)

Then it is clear that

$$\int f \, dQ = L'f$$

for all $f \in \mathbb{R}^\Omega$, both sides being linear functionals of $f$ which agree on the basis elements $\delta_\omega$. In particular, specializing to elements $f \in V$ we have the representation (3.2).

□

Compare Theorem 3.3 to Theorem 2 in Harrison and Kreps [10], where the no-arbitrage condition enters through a martingale assumption on the price process, which introduces the emphasis on the time dimension. A generalization to a very general class of martingales is achieved in Delbaen and Schachermayer [5, 6]. Condition (3.3) generalizes in the sense that it is not necessary to specify whether a typical element $\omega \in \Omega$ is a trajectory or a state at one particular time.
4. A Counterexample for Infinite $\Omega$

Our purpose in this section is to demonstrate that the hypothesis of finiteness of $\Omega$ cannot be dropped from Theorem 3.3. (See also Back and Pliska [1] for a similar phenomenon.)

Let $\Omega$ be any countably infinite set and let $V$ be the set of all real-valued functions on $\Omega$ which are constant outside finite sets, i.e. $V$ consists of all functions $f : \Omega \to \mathbb{R}$ such that there is a finite set $S_f$ outside which $f$ is constant. Clearly $V$ is a linear space under the pointwise operations. Let $L : V \to \mathbb{R}$ be the functional which associates to each $f \in V$ the constant value which $f$ takes outside some finite set. Again it is readily seen that $L$ is a linear functional. Our point now is:

Proposition 4.1. There is no measure $Q$ on any $\sigma$-algebra of subsets relative to which each $f \in V$ is measurable and for which $Lf = \int f \, dQ$ holds for every $f \in V$.

Proof. Let’s first check that the $\sigma$–algebra $\sigma(V)$ with respect to which all functions in $V$ are measurable is in fact the set of all subsets of $\Omega$. For any point $p \in \Omega$ the indicator function $1_{\{p\}}$ is in $V$. So the one-point set $\{p\}$, being the set on which $1_{\{p\}}$ takes the value 1, must be in the $\sigma(V)$. Since $\Omega$ is countable it follows that every subset of $\Omega$ is in $\sigma(V)$.

Next suppose there is a measure $Q$ on $\sigma(V)$ such that $Lf = \int f \, dQ$ holds for every $f \in V$. Then, since $\Omega$ is the countable union of all the one-point sets $\{p\}$, we have

$$Q(\Omega) = \sum_{p \in \Omega} Q(\{p\})$$

(4.1)

But

$$Q(\{p\}) = \int 1_{\{p\}} \, dQ = L(1_{\{p\}}) = 0$$

from the definition of $L$, since the function $1_{\{p\}}$ is, by definition, 0 at all points outside $\{p\}$. On the other hand,

$$Q(\Omega) = \int 1 \, dQ = L(1) = 1,$$

the last equality again being a direct consequence of the definition of $L$. The two preceding equations again contradict the relation (4.1). The contradiction shows that no $Q$ having the desired properties can exist. □

5. Pricing Measure on the Gel’fand Spectrum

In this section we shall show that in case $\Omega$ is infinite, it is possible to extend $\Omega$ in such a way that a probability measure $Q$ with the desired properties exists for the extended space.

Standard notions and theorems about Gel’fand theory which we shall use below are available in Rudin [14, Chapters 10, 11].

As usual, we start with a non-empty set $\Omega$, a linear space $V$ of real-valued bounded functions on $\Omega$, and a linear functional $L : V \to \mathbb{R}$ which satisfies the no-arbitrage condition $Lf \leq \max f$ for all $f \in V$.

Our objective in this section is to prove the following result:
Theorem 5.1. Let \( \Omega \) be a non-empty set, \( V \) a linear space of bounded real-valued functions on \( \Omega \), and \( L \) a linear functional on \( V \) such that \( Lf \leq \max f \) for all \( f \in V \). Then there is

(a) a set \( \hat{\Omega} \),

(b) a linear space \( \hat{V} \) of functions on \( \hat{\Omega} \),

(c) a probability measure \( Q \) on a \( \sigma \)-algebra of subsets of \( \hat{\Omega} \) with respect to which all the functions in \( \hat{V} \) are measurable,

(d) a linear isomorphism \( V \to \hat{V} : f \mapsto \hat{f} \)

(e) and a map \( i : \Omega \to \hat{\Omega} \),

such that:

(i) for every \( f \in V \), the value \( Lf \) is the average value of \( \hat{f} \) with respect to the probability measure \( Q \):

\[
L(f) = \int_{\hat{\Omega}} \hat{f} \, dQ \quad \text{for every } f \in V.
\]

(ii) the range of the function \( \hat{f} \) is the closure of the range of \( f \), for each function \( f \in V \),

(iii) the measure \( Q \) is defined on the completed \( \sigma \)-algebra generated by the functions \( \hat{f} \) with \( f \) running over \( V \)

(iv) the map \( i : \Omega \to \hat{\Omega} \) identifies any two states in \( \Omega \) which have the property that all elements \( f \in V \) take the same value in the two states, i.e. \( i(\omega) = i(\omega') \) if \( f(\omega) = f(\omega') \) for all \( f \in V \).

Moreover, there is a topology on \( \hat{\Omega} \) which makes it a compact Hausdorff space, the function \( \hat{f} \) is continuous on \( \hat{\Omega} \) for each \( f \in V \), and \( i(\Omega) \) is a dense subset of \( \hat{\Omega} \).

Suppose, in the preceding framework, that there is a non-negative function \( h \in V \) such that \( Lh = 0 \) but \( h \neq 0 \). Then we have \( \hat{h} \geq 0 \) on \( \hat{\Omega} \) and \( \int \hat{h} \, dQ = 0 \). This implies that

\[
Q[\{h > 0\}] = 0
\]

Thus the measure \( Q \) assigns zero probability to such events, i.e. to events allowing a profit from instruments which have zero cost. Suppose we start with an initially given measure \( P \) (for example, a model probability measure for real-life uncertainties) such that assets such as \( I_B \) for events with \( P(B) = 0 \) have zero price. Then events which have zero \( P \)-probability continue to have zero \( Q \)-probability. With some more structure, one would have \( Q \) equivalent as a measure to the given measure \( P \).

Before proceeding to the proof of Theorem 5.1 we shall discuss its meaning.

5.1. Interpretation of Theorem 5.1. As before, \( \Omega \) consists of the set of all possible states of a market. The elements of \( V \) are functions on \( \Omega \), and for such a function \( f \) the value \( f(\omega) \) is the worth of the asset when market state \( \omega \) is realized. The linear functional \( L : V \to \mathbb{R} \) associates to each \( f \in V \) the value \( Lf \), which is what a particular trader would pay for the instrument \( f \) without knowledge of which scenario/state will actually be realized. In market equilibrium all traders use the same functional \( L \). As we have seen before, when there are infinitely many market scenarios the price \( Lf \) may not arise as the expectation value of \( f \).
for all $f \in V$ simultaneously) relative to some probability measure $Q$ on the set of market scenarios. Theorem 5.1 solves this problem by providing a probability measure, but at a cost: this measure lives on a possibly larger space $\hat{\Omega}$. (The space $\Omega$ is not necessarily larger than $\hat{\Omega}$; indeed, for the setting of Example 2.1, the market equilibrium pricing measure is often taken to be Wiener measure on $\Omega = C_0([0, T]; \mathbb{R}^N)$, and asset worths specified through martingale processes, as explained in section 6 below.)

Part (d) of the theorem assures us that the instruments/assets $f$ continue to exist on the new market-scenario space $\hat{\Omega}$, the new versions being the functions $\hat{f}$.

Part (i) tells us that the desired fundamental relationship:

\[
\text{price of an instrument} = \text{average of possible potential values in all scenarios}
\]

now holds, with the averaging done with respect to the measure $Q$.

Part (iii) assures us that the original collection of market events are expanded only minimally in the new setting, at most by including events which are viewed as having probability zero.

Part (iv) actually fixes a problem that may have been present in the original formulation. Surely two market states in which every possible asset has the same worth should be considered as essentially the same market state, and this is precisely the case in $\hat{\Omega}$.

Lastly, there is a topology on the new market state space $\hat{\Omega}$ and relative to this the original set of states is a dense subset, thereby showing again that the new state space is only a minimal expansion of the original one. However, it must be said that the topology on $\Omega$ can be quite strange and need not be viewed as having any practical significance.

The measure $Q$ need not be uniquely defined, an issue that connects to the problem of market completeness and that we take up again in the Conclusion. Compare Theorem 5.1 to Theorem 3 in Harrison and Kreps [10], which shows the existence of a measure for one special case of an infinite state space $\Omega$, the case where the price process is a diffusion. The corresponding result in Delbaen and Schachermayer [5] is Theorem 1.1.

5.2. Proof of Theorem 5.1. The proof will use Gel’fand’s spectral theory of commutative $C^*$-algebras. We have kept the material below reasonably self-contained but a detailed account, including a proof of the fundamental theorem of Gel’fand and Naimark, are available in Rudin [14].

Let $B$ be the algebra of all complex-valued bounded functions on $\Omega$. On $B$ we have the sup-norm:

\[
|f|_{\sup} = \sup_{\omega \in \Omega} |f(\omega)|
\]

Then $B$ is a complex, commutative, Banach $*$-algebra with unit 1. The involution $* : B \to B$ is simply the conjugation $f^* = \overline{f}$. Note that

\[
\|ff^*\|_{\sup} = |f|_{\sup}^2
\]

Let $B_V$ be the closure of the complex subalgebra generated by all elements $f \in V$ along with the constant function 1. This is, of course, also a complex, commutative, Banach $*$-algebra with unit. The Gel’fand spectrum of $B_V$ is the set
\( \hat{\Omega} \) of all non-zero complex homomorphisms of the algebra \( B_V \). In more detail, the elements of \( \hat{\Omega} \) are all non-zero linear maps \( \phi : B_V \to \mathbb{C} \) for which \( \phi(fg) = \phi(f)\phi(g) \) holds for all \( f, g \in B_V \).

Thus for each \( f \in B_V \) there is a function \( \hat{f} \) on \( \hat{\Omega} \) given by

\[
\hat{f}(\phi) = \phi(f)
\]  

(5.3)

This is called the Gel'fand transform of \( f \).

The set \( \hat{\Omega} \) is equipped with the smallest topology with respect to which all the functions \( \hat{f} \), as \( f \) runs over \( B_V \), are continuous. Let \( C(\hat{\Omega}) \) be the set of all complex-valued continuous functions on \( \hat{\Omega} \). This is a complex commutative Banach *-algebra under the sup-norm, and equation (5.2) holds.

The fundamental theorem of Gel'fand and Naimark [14, Theorem 11.18] applied to this situation says that the Gel'fand transform

\[
B_V \to C(\hat{\Omega}) : f \mapsto \hat{f}
\]

is an isometric algebra isomorphism which preserves the conjugation operation.

**Lemma 5.2.** The Gel'fand transform carries the set of real-valued functions in \( B_V \) onto the set real-valued functions in \( C(\hat{\Omega}) \). Moreover, it maps the set of non-negative elements in \( B_V \) onto the set of non-negative elements in \( C(\hat{\Omega}) \).

**Proof.** An element \( f \in B_V \) is real-valued if and only if \( f = f^* \), and this translates by the Gel'fand transform to \( \hat{f} = (\hat{f})^* \) which is the condition for \( \hat{f} \) being real-valued.

Next suppose \( f \in B_V \) is a real-valued non-negative function on \( \Omega \). By the Weierstrass theorem there is a sequence of polynomials \( p_n(x) \) such that \( p_n(x) \to \sqrt{f} \) uniformly on the compact interval \([0, |f|_{\text{sup}}]\). So the sequence of elements \( p_n \circ f \), which belong to \( B_V \), converge uniformly to \( \sqrt{f} \). So \( \sqrt{f} \) is an element of \( B_V \). Thus, writing \( h = \sqrt{f} \) we have \( h \in B_V \), \( h = h^* \) and \( f = hh^* \). Applying the Gel'fand transform, which preserves the conjugation operation *\( *\), we see that \( \hat{f} = gg^* = |g|^2 \), where \( g = \hat{h} \in C(\Delta) \). So \( \hat{f} \geq 0 \). The same argument can be applied to the inverse Gel'fand transform to obtain the converse result. \( \square \)

The order preserving nature of the Gel'fand transform noted above implies that

\[
\max \hat{f} = \max f \text{ for all real-valued } f \in B_V
\]

because for any real number \( c \), we have \( f \leq c \) if and only if \( \hat{f} \leq \hat{c} \) and, since the Gel'fand transform preserves \( 1 \) we have \( \hat{c} = c \).

Recall that \( B_V \) is the sup-norm closed algebra of functions on \( \Omega \) generated by all the functions in \( V \) and the constant function \( 1 \). Now the algebra \( B_V \), viewed as a real vector space, contains \( V \) as a subspace. Let \( \hat{V} \) be the image of \( V \) in \( C(\hat{\Omega}) \) under the Gel'fand transform; it is a real subspace of \( C(\hat{\Omega})_{\text{real}} \), this being the algebra of real-valued continuous functions on \( C(\hat{\Omega}) \). The real-linear functional

\[
\hat{L} : \hat{V} \to \mathbb{R} : f \mapsto Lf
\]
The no-arbitrage condition \( Lf \leq \max f \) valid for \( f \in V \), implies that \( \hat{L}h \leq \max h \) is valid for all \( h \in \hat{V} \). The extension result given by Theorem 3.2 then provides a linear functional \( \hat{L}' \) on \( C(\hat{\Omega})_{\text{real}} \) such that

\[ \hat{L}' \leq \max f \text{ for all } f \in C(\hat{\Omega})_{\text{real}}. \]

Switching \( f \) to \(-f\) gives

\[ \hat{L}' \geq \min f \]

In particular, we have

\[ \hat{L}' \geq 0 \text{ for all non-negative } f \in C(\hat{\Omega}) \]  \( (5.4) \)

as well as

\[ \hat{L}' 1 = 1 \]

The Riesz-Markov theorem then implies that there is a measure \( Q \) on the Borel \( \sigma \)-algebra of \( \hat{\Omega} \) such that

\[ \hat{L}' f = \int f \, dQ \text{ for every } f \in C(\hat{\Omega})_{\text{real}}. \]  \( (5.5) \)

The definition of \( Q \) starts with setting

\[ Q(K) = \inf_{\{ f \in C(\hat{\Omega})_{\text{real}}, f \geq 1_K \}} \hat{L}f \]  \( (5.6) \)

for every compact \( K \subset \hat{\Omega} \) and then showing that this extends to a measure on the entire Borel \( \sigma \)-algebra. Setting \( f = 1 \) in (5.5) shows that \( Q(\hat{\Omega}) = 1 \), i.e. that \( Q \) is a probability measure.

Recall that if a measure \( \mu \) is given on a \( \sigma \)-algebra \( F \) of subsets of some set \( X \) then it is often useful to work with the larger \( \sigma \)-algebra \( F_\mu \) generated by the sets of \( \mu \)-measure 0. The \( \sigma \)-algebra \( F_\mu \) is the completion of \( F \) by \( \mu \), and below we shall use the notation of subscripting by a measure to denote completion.

**Proposition 5.3.** The \( Q \)-completed Borel \( \sigma \)-algebra of \( \hat{\Omega} \) is generated by the functions \( \hat{f} \) for \( f \) running over \( V \) along with the sets of \( Q \)-measure 0.

**Proof.** Let \( B \) be the Borel \( \sigma \)-algebra of \( \hat{\Omega} \), and \( B' \) the sub-algebra generated by the functions \( f \) with \( f \) running over \( V \). Our goal is to show that

\[ B_Q = B'_Q \]  \( (5.7) \)

By definition of \( B_V \), each function in \( B_V \) is the uniform limit of a sequence of functions of the form \( P(h_1, \ldots, h_m) \), with \( P \) being polynomial in \( m \) variables and \( h_1, \ldots, h_m \in V \), and \( m \) varying over positive integers. Applying the Gel'fand transform to this observation shows that, in particular, the \( \sigma \)-algebra of subsets of \( \hat{\Omega} \) generated by the continuous functions *coincides* with the \( \sigma \)-algebra generated by just the functions \( \hat{f} \) with \( f \) running over \( \hat{V} \).

Thus, to prove that \( B \subset B'_Q \), it will suffice to show that for any closed set \( D \subset \hat{\Omega} \) there is a sequence of real-valued continuous functions \( f_n \) on \( \hat{\Omega} \) such that \( f_n(x) \to 1_D(x) \) for \( Q \)-almost every \( x \in \hat{\Omega} \). [This follows from Lusin’s theorem but we include a proof specialized to the present context.] Since \( \hat{\Omega} \) is compact
Hausdorff, closed subsets of \( \hat{\Omega} \) are compact. As noted in (5.6) the definition of \( Q \) requires that
\[
Q(D) = \inf_{f \in C(\hat{\Omega})_{\text{real}}, 1_D \leq f} \hat{L}f
\]
So there is a sequence of real-valued continuous functions \( f_n \) on \( \hat{\Omega} \) with \( f_n \geq 1_D \) such that
\[
Q(D) = \lim_{n \to \infty} \hat{L}(f_n)
\]
Replacing \( f_n \) by \( \min\{f_1, \ldots, f_n\} \) we may assume that \( f_1 \geq f_2 \geq \ldots \). So the pointwise limit \( \lim_{n \to \infty} f_n(x) \) exists for every \( x \in \hat{\Omega} \); denote this limit by \( f(x) \). Then \( f \) is the pointwise limit of a sequence of continuous functions and \( f \geq 1_D \). Moreover,
\[
\int f \, dQ = \lim_{n \to \infty} \int f_n \, dQ
\]
by dominated convergence, and so, since \( \int f_n \, dQ = \hat{L}f_n \) it follows that
\[
\int f \, dQ = Q(D)
\]
Since \( f \geq 1_D \) we conclude that \( f \) must actually be equal to \( 1_D \) almost everywhere with respect to \( Q \).

The preceding arguments prove that for any closed subset \( D \) of \( \hat{\Omega} \), the indicator function \( 1_D \) is \( Q \)-almost-everywhere the pointwise limit of a sequence of continuous functions and each such function is itself a uniform limit of a sequence of functions expressible as polynomials in the elements of \( \hat{V} \). Consequently, \( 1_D \) is measurable with respect to the \( Q \)-completed \( \sigma \)-algebra \( B_Q' \) generated by all the functions in \( \hat{V} \). So \( D \) itself is in the latter \( \sigma \)-algebra. Since the closed sets generate the Borel \( \sigma \)-algebra it follows that every Borel set is in \( B_Q' \), i.e. \( B \subset B_Q' \). It follows that
\[
B_Q \subset B_Q'
\]
Conversely, since \( B' \) is generated by a family of continuous functions we have \( B' \subset B \) and so also \( B_Q' \subset B_Q \). This completes the proof of (5.7). \( \square \)

Finally we include a proof for the claim that the range of \( \hat{f} \) is the closure of the range of \( f \):

**Proposition 5.4.** For any \( f \in V \) the range of the Gel’fand transform \( \hat{f} \) is the closure of the range of \( f \).

**Proof.** The function \( \hat{f} \) being continuous on the compact space \( \hat{\Omega} \) has compact image. Suppose \( k \) is a real number lying in the closure of the range of \( f \). Then \( f \) takes values arbitrarily close to \( k \) or has the value \( k \) itself in the range; so \( 1/(f - k) \) is either not defined everywhere or is unbounded. In fact, \( k \) is in the closure of the range of \( f \) if and only if \( f - k \) is not invertible in \( B_V \). Passing over isomorphically to \( C(\hat{\Omega}) \), this is equivalent to \( \hat{f} - k \) having no inverse in \( C(\hat{\Omega}) \), which in turn is equivalent to \( k \) being in the closure of the range of \( \hat{f} \). Since the range of \( \hat{f} \) is closed, we conclude that \( k \) is in the closure of the range of \( f \) if and only if it is in the range of \( f \). \( \square \)
6. Pricing with Additional Information

We shall now consider pricing in the presence of information. In the preceding sections we have analyzed prices of instruments decided on an *a priori* basis. The task now is to analyze prices of instruments based on knowledge of the values of a certain given set of instruments. For example, the given instruments may be all market instruments at a particular time, and the task is to understand prices at a later time.

In the first few paragraphs below we study the situation to isolate a mathematical model and then we prove a result within this model and briefly indicate its ramifications.

As before, we have our set up consisting of the market state space $\Omega$, the space $V$ of functions on $\Omega$ corresponding to traded instruments, and the functional $L$ on $V$ for which $Lf$ is the *a priori* price the trader would pay for instrument $f$. We now consider the price the trader would be pay if the values of a certain collection $A$ of instruments were known. For simplicity of exposition, let us think of $A$ as being a finite set $A = \{X_1, \ldots, X_n\}$. The price based on knowledge of the values of the functions $X_i$, would be described by a functional $f \mapsto L_A f$, where now $L_A f$ is determined by the values of $X_1, \ldots, X_n$, i.e.

the value of $L_A f$ is a function of the values of $X_1, \ldots, X_n$. \hfill (6.1)

Let us now switch over to the setting of $\hat{\Omega}$, the functional $\hat{L}$ and the corresponding probability measure $Q$ on the Borel $\sigma$–algebra $B$ of $\hat{\Omega}$. It will be convenient to define

$$\hat{L} f = \int f \, dQ$$

for all Borel functions $f$ on $Q$ for which the integral exists (certainly this is in agreement with the case for continuous functions). It will also be convenient to work with the *real* Hilbert space $L^2_r(\hat{\Omega}, B, Q)$ of Borel-measurable functions $g : \hat{\Omega} \to \mathbb{R}$ which are square-integrable, i.e. for which $\int g^2 \, dQ < \infty$. The inner-product on $L^2_r(\hat{\Omega}, B, Q)$ is given, as usual, by

$$\langle g, h \rangle = \int gh \, dQ$$

There is a convenient way to capture the notion of a function being “a function of” a given collection of functions. To this end, let $A$ be a given collection of Borel functions on $\hat{\Omega}$ and $A$ the $\sigma$–algebra generated by these functions. A function $f$ on $\hat{\Omega}$ is a “function of” the given collection $A$ if $f$ is measurable with respect to $A$. In view of this, statement (6.1) can be rewritten as

$\hat{L} A f$ is an $A$–measurable function. \hfill (6.2)

Price consistency requires that $\hat{L} A$ be linear. Actually, now something more should be true: the equation

$$\hat{L} A (kf) = k \hat{L} A f$$
Table 2. Pricing Framework with Information.

<table>
<thead>
<tr>
<th>Mathematical object</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A set of $\mathcal{A}$ of Borel functions on $\Omega$</td>
<td>$A$ corresponds to a given set of assets whose values will be known to the trader</td>
</tr>
<tr>
<td>The $\sigma$–algebra algebra $\mathcal{A}$ generated by the functions in $A$</td>
<td>Events in $\mathcal{A}$ are those determined by the given set $A$ of asset prices</td>
</tr>
<tr>
<td>A linear operator $L_A$ on $L^2(\hat{\Omega}, \mathcal{B}, Q)$</td>
<td>$\hat{L}_A f$ is the price for asset $f$ based on knowledge of the values of the given assets</td>
</tr>
<tr>
<td>$\hat{L}_A(kf) = k\hat{L}_A f$ for all $f \in L^2(\hat{\Omega}, \mathcal{B}, Q)$ and all $k \in L^2(\hat{\Omega}, \mathcal{A}, Q)$</td>
<td>$k$ being $\mathcal{A}$–measurable means $k$ is effectively a constant, being determined by the given information</td>
</tr>
<tr>
<td>$\hat{L} \left( \hat{L}_A f \right) = \hat{L} f$ for all $f \in L^2(\hat{\Omega}, \mathcal{B}, Q)$</td>
<td>this is a compatibility condition between a priori pricing by $\hat{L}$ and pricing by $\hat{L}_A$ based on given information</td>
</tr>
</tbody>
</table>

should hold not only for all constants $k$, but all $k$ which are functions of $X_1, \ldots, X_n$; thus (focusing on bounded $k$ for technical convenience) we should require that

$$\hat{L}_A(kf) = k\hat{L}_A f \text{ for all } f \in L^2(\hat{\Omega}, \mathcal{B}, Q) \text{ and all } k \in L^\infty(\mathcal{B}, Q) \quad (6.3)$$

(here $L^\infty(\mathcal{B}, Q)$ is the space of essentially bounded Borel measurable functions). The reason for this is that such $k$ are effectively constants or “known” quantities for the trader pricing with knowledge of the values of $X_1, \ldots, X_n$.

Lastly, we need to relate $L_A$ with the a priori pricing functional $\hat{L}$. The consistency condition we impose is

$$\hat{L} \left( \hat{L}_A f \right) = \hat{L} f \text{ for all } f \in L^2(\hat{\Omega}, \mathcal{B}, Q) \quad (6.4)$$

This may be understood conceptually, but in the end it is an additional assumption on the way pricing with information relates to pricing without information and we take (6.4) as an axiom.

We shall view $L_A$ as a linear operator on $L^2(\hat{\Omega}, \mathcal{B}, Q)$ with range in the closed subspace $L^2(\hat{\Omega}, \mathcal{A}, Q)$.

Table 2 summarizes our model for pricing with information.

We then have the following geometrical description of the operator $\hat{L}_A$:

**Proposition 6.1.** The operator $\hat{L}_A$ is given by the orthogonal projection of the Hilbert space $L^2(\hat{\Omega}, \mathcal{B}, Q)$ onto the closed subspace $L^2(\hat{\Omega}, \mathcal{A}, Q)$. 
Proof. For any $f \in L^2(\hat{\Omega}, B, Q)$ and any bounded $g$ in $L^2(\hat{\Omega}, A, Q)$ we have
\[
\langle f - \hat{L}Af, g \rangle = \int fg \, dQ - \int (\hat{L}Af)g \, dQ
\]
\[
= \hat{L}(fg) - \hat{L}(g\hat{L}Af)
\]
\[
= \hat{L}(fg) - \hat{L}(gf) \quad \text{because } g \text{ is } A\text{-measurable}
\]
\[
= 0
\]
Any $h \in L^2(\hat{\Omega}, A, Q)$ is the $L^2$-limit of the bounded functions obtained by truncating $h$ off above at $N$ and below at $-N$, with $N \to \infty$. So it follows that \[
\langle f - \hat{L}Af, h \rangle = 0
\]
for every $h \in L^2(\hat{\Omega}, A, Q)$. This says exactly that the element $\hat{L}Af$ in $L^2(\hat{\Omega}, A, Q)$ has the property that $f - \hat{L}Af$ is orthogonal to the subspace $L^2(\hat{\Omega}, A, Q)$. Thus, $\hat{L}Af$ is the orthogonal projection of $f$ onto $L^2(\hat{\Omega}, A, Q)$.

There is a standard interpretation of orthogonal projections in the above setting as conditional expectations. Using the relation \[
\int fg \, dQ = \int g\hat{L}Af \, dQ
\]
with $g = \mathbb{1}_E$ for any event $E \in A$, we see that $\hat{L}Af$ is the $A$-measurable function for which \[
\int_E f \, dQ = \int_E \hat{L}Af \, dQ
\]
holds for all events $E$ in $A$. Thus we have the following probabilistic view of $\hat{L}A$:

**Theorem 6.2.** For any $f \in L^2(\hat{\Omega}, B, Q)$, the price $\hat{L}Af$ based on knowledge given by the $\sigma$–algebra $A$ is the conditional expectation $E_Q(f|A)$:
\[
\hat{L}Af = E_Q(f|A) \tag{6.5}
\]

The geometrical significance of the conditional pricing functional $\hat{L}A$, or the probabilistic interpretation given above, has the following consequence:

**Proposition 6.3.** If $A$ and $B$ are collections of instruments with $A \subset B$ then:
\[
\hat{L}A \left( \hat{L}Bf \right) = \hat{L}A(f) \tag{6.6}
\]

Let us consider a special case. Take $A$ to be all market instruments up till time $s$ and $B$ to be all market instruments up till a later time $t > s$. Then, denoting $\hat{L}Af$ by $f_s$, and similarly for $B$, we have the well-known martingale condition for prices:
\[
E_Q[f_t|F_s] = f_s \tag{6.7}
\]
where $F_s$ is the collection of all market events up till time $s$. More specifically, in the setting of Example 2.1, this says that the discounted prices of traded instruments are martingales.
7. Conclusion and Open Questions

We show in this paper that the first fundamental theorem of asset pricing, namely the equivalence of the condition of no-arbitrage and existence of a pricing measure, can be achieved in the framework of the Gel'fand transform. This makes it possible to de-emphasize the time dimension and the martingale property of the price process compared to the standard approach. We show in our framework how the martingale feature arises as a special case of the more general property that the equilibrium market price of an asset under partial information about the market is the corresponding conditional expectation of the price. Partial information can but need not be indexed by time.

The question of market completeness, sometimes referred to as the second fundamental theorem of asset pricing and connected to uniqueness of the pricing measure, is not addressed in this paper and left for future research. It is particularly challenging to disentangle this issue from the time dimension since in the standard framework, completeness is usually defined to mean that for every payoff function $f$ of a “basic” instrument $X$ at maturity $T$, there is a self-financing portfolio $p$ such that $p(0)$ is the fair market price at time 0 and that $p(T) = f(X(T))$. Self-financing means that between 0 and $T$, $p$ reaches the value $f(X(T))$ without injection or extraction of funds except for the price $p(0)$. Thus, the standard definition of completeness is essentially a time concept. The possibility and desirability to achieve a definition of completeness that does not rely on the martingale approach has been recognized before, however (Battig and Jarrow [3]). As noted earlier in the Introduction, completeness of a market is, broadly, connected with uniqueness of the pricing measure. However, exactly how this uniqueness is formulated may depend on the technical framework. We leave the task of exploring market completeness issues in the Gel’fand transform framework for future work.

Acknowledgment. A. Sengupta wishes to thank Greg Lieb and Morten Andersen for many useful and enlightening conversations. The inspiration for constructing a pricing measure using the Gel’fand transform came from a set of notes [9] on quantum field theory by Leonard Gross where he explains (following earlier ideas of Irving Segal) how to construct an infinite-dimensional Gaussian measure using the Gel’fand transform.

References


**ERIC HILLEBRAND:** DEPARTMENT OF ECONOMICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803, USA

*E-mail address:* erhil@lsu.edu

**AMBAR N. SENGUPTA:** DEPARTMENT OF MATHEMATICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803, USA

*E-mail address:* sengupta@math.lsu.edu

*URL:* http://www.math.lsu.edu/~sengupta
Diffeomorphisms of the Circle and Brownian Motions on an Infinite-Dimensional Symplectic Group

Maria Gordina* and Mang Wu**

Abstract. An embedding of the group $\text{Diff}(S^1)$ of orientation preserving diffeomorphisms of the unit circle $S^1$ into an infinite-dimensional symplectic group, $\text{Sp}(\infty)$, is studied. The authors prove that this embedding is not surjective. A Brownian motion is constructed on $\text{Sp}(\infty)$. This study is motivated by recent work of H. Airault, S. Fang and P. Malliavin.

1. Introduction

The group $\text{Diff}(S^1)$ of orientation preserving diffeomorphisms of the unit circle $S^1$ has been extensively studied for a long time. One of the goals of the research has been to construct and study the properties of a Brownian motion on this group. In [1] H. Airault and P. Malliavin considered an embedding of $\text{Diff}(S^1)$ into an infinite-dimensional symplectic group.

This group, $\text{Sp}(\infty)$, can be represented as a certain infinite-dimensional matrix group. For such matrix groups, the method of [6, 7] can be used to construct a Brownian motion living in the group. This construction relies on the fact that these groups can be embedded into a larger Hilbert space of Hilbert-Schmidt operators. We use the same method to construct a Brownian motion on $\text{Sp}(\infty)$. One of the advantages of Hilbert-Schmidt groups is that one can associate an infinite-dimensional Lie algebra to such a group, and this Lie algebra is a Hilbert space. This is not the case with $\text{Diff}(S^1)$, as an infinite-dimensional Lie algebra associated with $\text{Diff}(S^1)$ is not a Hilbert space with respect to the inner product compatible with the symplectic structure on $\text{Diff}(S^1)$.

In the current paper, we describe in detail the embedding of $\text{Diff}(S^1)$ into $\text{Sp}(\infty)$, and construct a Brownian motion on $\text{Sp}(\infty)$. Our motivation comes from an attempt to use this embedding to better understand Brownian motion in $\text{Diff}(S^1)$ as studied by H. Airault, S. Fang and P. Malliavin in a number of papers (e.g. [1, 2, 4, 5]). One of the main results of the paper is Theorem 4.6, where we describe the embedding of $\text{Diff}(S^1)$ into $\text{Sp}(\infty)$ and prove that the map is not surjective. Theorem 6.17 gives the construction of a Brownian motion on $\text{Sp}(\infty)$. In order for this Brownian motion to live in the group we are forced to choose

2000 Mathematics Subject Classification. Primary 60H07; Secondary 58J65, 60J65.

Key words and phrases. $\text{Diff}(S^1)$, infinite-dimensional symplectic group, Brownian motion.

* This research is partially supported by NSF Grant DMS-0706784.

** This research is partially supported by NSF Grant DMS-0706784.
a non-\text{Ad}-invariant inner product on the Lie algebra of \text{Sp}(\infty). This fact has a potential implication for this Brownian motion not to be quasi-invariant for the appropriate choice of the Cameron-Martin subgroup of \text{Sp}(\infty). This is in contrast to results in [2]. The latter can be explained by the fact that the Brownian motion we construct in Section 6 lives in a subgroup of \text{Sp}(\infty) whose Lie algebra is much smaller than the full Lie algebra of \text{Sp}(\infty).

2. The Spaces \( H \) and \( \mathbb{H}_\omega \)

**Definition 2.1.** Let \( H \) be the space of complex-valued \( C^\infty \) functions on the unit circle \( S^1 \) with the mean value 0. Define a bilinear form \( \omega \) on \( H \) by

\[
\omega(u, v) = \frac{1}{2\pi} \int_0^{2\pi} uv' d\theta, \quad \text{for any} \; u, v \in H.
\]

**Remark 2.2.** By using integration by parts, we see that the form \( \omega \) is antisymmetric, that is, \( \omega(u, v) = -\omega(v, u) \) for any \( u, v \in H \).

Next we define an inner product \( (\cdot, \cdot)_\omega \) on \( H \) which is compatible with the form \( \omega \). First, we introduce a complex structure on \( H \), that is, a linear map \( J \) on \( H \) such that \( J^2 = -id \). Then the inner product is defined by \( (u, v)_\omega = \pm \omega(u, Jv) \), where the sign depends on the choice of \( J \). The complex structure \( J \) in this context is called the Hilbert transform.

**Definition 2.3.** Let \( \mathbb{H}_0 \) be the Hilbert space of complex-valued \( L^2 \) functions on \( S^1 \) with the mean value 0 equipped with the inner product

\[
(u, v) = \frac{1}{2\pi} \int_0^{2\pi} uv d\theta, \quad \text{for any} \; u, v \in \mathbb{H}_0.
\]

**Notation 2.4.** Denote \( \hat{e}_n = e^{in\theta}, n \in \mathbb{Z}\backslash\{0\} \), and \( \mathcal{B}_H = \{\hat{e}_n, n \in \mathbb{Z}\backslash\{0\}\} \). Let \( \mathbb{H}^+ \) and \( \mathbb{H}^- \) be the closed subspaces of \( \mathbb{H}_0 \) spanned by \( \{\hat{e}_n : n > 0\} \) and \( \{\hat{e}_n : n < 0\} \), respectively. By \( \pi^+ \) and \( \pi^- \) we denote the projections of \( \mathbb{H}_0 \) onto subspaces \( \mathbb{H}^+ \) and \( \mathbb{H}^- \), respectively. For \( u \in \mathbb{H}_0 \), we can write \( u = u_+ + u_- \), where \( u_+ = \pi^+(u) \) and \( u_- = \pi^-(u) \).

**Definition 2.5.** Define the Hilbert transformation \( J \) on \( \mathcal{B}_H \) by

\[
J : \hat{e}_n \mapsto i \text{sgn}(n) \hat{e}_n
\]

where \( \text{sgn}(n) \) is the sign of \( n \), and then extended by linearity to \( \mathbb{H}_0 \).

**Remark 2.6.** In the above definition, \( J \) is defined on the space \( \mathbb{H}_0 \). We need to address the issue whether it is well-defined on the subspace \( H \). That is, if \( J(H) \subseteq H \). We will see that if we modify the space \( H \) a little bit, for example, if we let \( C^1_0(S^1) \) be the space of complex-valued \( C^1 \) functions on the circle with mean value zero, then \( J \) is not well-defined on \( C^1_0(S^1) \). This problem really lies in the heart of Fourier analysis. To see this, we need to characterize \( J \) by using the Fourier transform.

**Notation 2.7.** For \( u \in \mathbb{H}_0 \), let \( \mathcal{F} : u \mapsto \hat{u} \) be the Fourier transformation with \( \hat{u}(n) = (u, \hat{e}_n) \). Let \( \hat{J} \) be a transformation on \( l^2(\mathbb{Z}\backslash\{0\}) \) defined by \( (\hat{J}\hat{u})(n) = i \text{sgn}(n)\hat{u}(n) \) for any \( \hat{u} \in l^2(\mathbb{Z}\backslash\{0\}) \).
The Fourier transformation $\mathcal{F} : \mathbb{H}_0 \rightarrow l^2(\mathbb{Z}\setminus\{0\})$ is an isomorphism of Hilbert spaces, and $J = \mathcal{F}^{-1} \circ \hat{J} \circ \mathcal{F}$.

**Proposition 2.8.** The Hilbert transformation $J$ is well–defined on $H$, that is $J(H) \subseteq H$.

**Proof.** The key of the proof is the fact that functions in $H$ can be completely characterized by their Fourier coefficients. To be precise, let $u \in \mathbb{H}_0$ be continuous. Then $u$ is in $C^\infty$ if and only if $\lim_{n \to \infty} n^k \hat{u}(n) = 0$ for any $k \in \mathbb{N}$. From this fact, it follows immediately that $J$ is well–defined on $H$, because $J$ only changes the signs of the Fourier coefficients of a function $u \in H$.

For completeness of exposition, we give a proof of this characterization. Though this is probably a standard fact in Fourier analysis, we found a proof (in [8]) of only one direction.

We first assume that $u$ is $C^\infty$. Then $u(\theta) = u(0) + \int_0^\theta u'(t)dt$. So

$$
\hat{u}(n) = \frac{1}{2\pi} \left( \int_0^{2\pi} \int_0^{2\pi} u'(t) \chi_{[0,\theta]}(t) e^{-int} dt \right) \theta = \frac{1}{2\pi} \int_0^{2\pi} \left( \int_t^{2\pi} e^{-int} \theta \right) u'(t)dt
$$

$$
= -\frac{1}{2\pi\imath n} \int_0^{2\pi} u'(t) - u^{\prime}(t)e^{-int}dt = \frac{\hat{u}'(n)}{\imath n},
$$

where we have used Fubini’s theorem and the continuity of $u'$. Now, $u'$ is itself $C^\infty$, so we can apply the procedure again. By induction, we get $\hat{u}(n) = \frac{\hat{u}^{(k)}(n)}{(\imath n)^k}$. But from the general theory of Fourier analysis, $\hat{u}^{(k)}(n) \rightarrow 0$ as $n \rightarrow \infty$. Therefore $n^k \hat{u}(n) \rightarrow 0$ as $n \rightarrow \infty$.

Conversely, assume $u$ is such that for any $k$, $n^k \hat{u}(n) \rightarrow 0$ as $n \rightarrow \infty$. Then the Fourier series of $u$ converges uniformly. Also by assumption that $u$ is continuous, the Fourier series converges to $u$ for all $\theta \in S^1$ (see Corollary I.3.1 in [8]). So we can write $u(\theta) = \sum_{n \neq 0} \hat{u}(n)e^{in\theta}$. Fix a point $\theta \in S^1$, then

$$
u'(\theta) = \frac{d}{dt} \left|_{t=\theta} \sum_{n \neq 0} \hat{u}(n)e^{int} \right| = \lim_{t \rightarrow \theta} \lim_{N \rightarrow \infty} \sum_{n=-N}^{N} \frac{\hat{u}(n)}{t - \theta}e^{int} - e^{in\theta}.
$$

Note that the derivatives of $\cos nt$ and $\sin nt$ are all bounded by $|n|$. So by the mean value theorem, $|\cos nt - \cos n\theta| \leq |n||t - \theta|$, and $|\sin nt - \sin n\theta| \leq |n||t - \theta|$. So

$$
\left| \frac{e^{int} - e^{in\theta}}{t - \theta} \right| \leq 2|n|, \quad \text{for any } t, \theta \in S^1.
$$

Therefore, by the growth condition on the Fourier coefficients $\hat{u}$, we have

$$
\lim_{N \rightarrow \infty} \sum_{n=-N}^{N} \hat{u}(n) \frac{e^{int} - e^{in\theta}}{t - \theta}
$$

converges at the fixed $\theta \in S^1$ and the convergence is uniform in $t \in S^1$. Therefore we can interchange the two limits, and obtain

$$
\left( \sum_{n \neq 0} \hat{u}(n)e^{in\theta} \right)' = \sum_{n \neq 0} \hat{u}(n)ine^{in\theta}.
$$
which means we can differentiate term by term. So the Fourier coefficients of \( u' \) are
given by \( \hat{u'}(n) = in\hat{u}(n) \). Clearly, \( u' \) satisfies the same condition as \( \hat{u} \): \( n^k\hat{u'}(n) \to 0 \)
as \( n \to \infty \). By induction, \( u \) is \( j \)-times differentiable for any \( j \). Therefore, \( u \) is in \( C^{\infty} \).

**Proposition 2.9.** Let \( C^1_0(S^1) \) be the space of complex-valued \( C^1 \) functions on the circle with the mean value zero. Then the Hilbert transformation \( J \) is not well-defined on \( C^1_0(S^1) \), i.e., \( J(C^1_0(S^1)) \notin C^1_0(S^1) \).

**Proof.** Let \( C(S^1) \) be the space of continuous functions on the circle. In \([8]\), it is shown that there exists a function in \( C(S^1) \) such that the corresponding Fourier series does not converge uniformly \([8, \text{Theorem II.1.3}]\), and therefore there exists an \( f \in C(S^1) \) such that \( Jf \notin C(S^1) \) \([8, \text{Theorem II.1.4}]\). Now take \( f = f_0 \) where \( f_0 \) is the mean value of \( f \). Then \( u \) is a continuous function on the circle with the mean value zero, and \( Ju \) is not continuous.

Using Notation 2.4 let us write \( u = u_+ + u_- \). Then we can use the relation

\[
  iu + Ju = 2iu_+ \quad \text{and} \quad iu - Ju = 2iu_-.
\]

to see that \( u_+ \) and \( u_- \) are not continuous. Integrating \( u = u_+ + u_- \), we have

\[
  \int_0^t u(\theta)d\theta = \int_0^t u_+(\theta)d\theta + \int_0^t u_-(\theta)d\theta.
\]

Denote the three functions in the above equation by \( v, v_1, v_2 \). By theorem I.1.6 in \([8]\),

\[
  \hat{v}(n) = \frac{\hat{u}(n)}{in}, \quad \text{and} \quad \hat{v}_1(n) = \frac{\hat{u}_+(n)}{in}, \hat{v}_2(n) = \frac{\hat{u}_-(n)}{in} \quad \text{for} \quad n \neq 0.
\]

Let \( g = v - v_0 \) where \( v_0 \) is the mean value of \( v \). Then \( g \in C^1_0(S^1) \). Write \( g = g_+ + g_- 2.4 \). Then \( g_+ = v_1 - (v_1)_0 \) and \( g_- = v_2 - (v_2)_0 \) where \( (v_1)_0 \) and \( (v_2)_0 \) are the mean values of \( v_1 \) and \( v_2 \) respectively. Then \( g_+ \), \( g_- \) \( \notin C^1_0(S^1) \) since \( v_1' = u_+ \), \( v_2' = u_- \) are not continuous.

By the relation

\[
  ig + Jg = 2ig_+ \quad \text{and} \quad ig - Jg = 2ig_-,
\]

we see that \( Jg \notin C^1_0(S^1) \). \( \square \)

**Notation 2.10.** Define an \( \mathbb{R} \)-bilinear form \( \langle \cdot, \cdot \rangle_\omega \) on \( H \) by

\[
  \langle u, v \rangle_\omega = -\omega(u, J\hat{v}) \quad \text{for any} \quad u, v \in H.
\]

**Proposition 2.11.** \( \langle \cdot, \cdot \rangle_\omega \) is an inner product on \( H \).

**Proof.** We need to check that \( \langle \cdot, \cdot \rangle_\omega \) satisfies the following properties (1) \( \langle \lambda u, v \rangle_\omega = \lambda \langle u, v \rangle_\omega \) for \( \lambda \in \mathbb{C} \); (2) \( \langle v, u \rangle_\omega = \langle u, v \rangle_\omega \); (3) \( \langle u, u \rangle_\omega > 0 \) unless \( u = 0 \).

(1) for \( \lambda \in \mathbb{C} , \)

\[
  \langle \lambda u, v \rangle_\omega = -\omega(\lambda u, J\hat{v}) = -\lambda \cdot \omega(u, J\hat{v}) = \lambda \cdot \langle u, v \rangle_\omega.
\]

To prove (2) and (3), we need some simple facts: \( H^+ = \pi^+(H) \subseteq H \) and \( H^- = \pi^-(H) \subseteq H \), and \( H = H^+ \oplus H^- \). If \( u \in H^+ \), \( v \in H^- \), then \( \langle u, v \rangle = 0 \). If \( u \in H^+ \), \( v \in H^- \), then \( \hat{u} \in H^- \), \( Ju = iu \), \( J\hat{u} = H^+ \). If \( u \in H^- \), then \( \hat{u} \in H^+ \), \( Ju = -iu \), \( J\hat{u} \in H^-. \)
\( J\dot{u} = \overline{J u} \) and \( \dot{u}'(n) = in\dot{u}(n) \). In particular, if \( u \in H^+ \), then \( u' \in H^+ \); if \( u \in H^- \), then \( u' \in H^- \).

(2) By definition,

\[
(v, u)_\omega = -\omega(v, J\dot{u}) = \omega(J\dot{u}, v) = \frac{1}{2\pi} \int (J\dot{u})v'd\theta
\]

\[
\overline{(u, v)}_\omega = -\overline{\omega(u, J\dot{v})} = \overline{\omega(J\dot{v}, u)} = \frac{1}{2\pi} \int \overline{J\dot{v}}u'd\theta = \frac{1}{2\pi} \int (J\dot{v})\dot{u}'d\theta.
\]

Write \( u = u_+ + u_- \) and \( v = v_+ + v_- \) as in Notation 2.4. Using the above fact, we can show that the above two quantities are equal to each other.

(3) Write \( u = u_+ + u_- \), then

\[
(u, u)_\omega = \frac{1}{2\pi} \int (-in\pi u_+ + i\pi u_-)d\theta = \sum_{n \neq 0} |n||\dot{u}(n)|^2.
\]

Therefore, \( (u, u)_\omega > 0 \) unless \( u = 0 \).

**Definition 2.12.** Let \( \mathbb{H}_\omega \) be the completion of \( H \) under the norm \( \| \cdot \|_\omega \) induced by the inner product \((\cdot, \cdot)_\omega\). Define

\[
\mathcal{B}_\omega = \left\{ \hat{e}_n = \frac{1}{\sqrt{n}}e^{in\theta}, n > 0 \right\} \cup \left\{ \tilde{e}_n = \frac{1}{i\sqrt{|n|}}e^{in\theta}, n < 0 \right\}.
\]

**Remark 2.13.** \( \mathbb{H}_\omega \) is a Hilbert space. Also the norm \( \| \cdot \|_\omega \) induced by the inner product \((\cdot, \cdot)_\omega\) is strictly stronger than the norm \( \| \cdot \| \) induced by the inner product \((\cdot, \cdot)\). So \( \mathbb{H}_\omega \) can be identified as a proper subspace of \( \mathbb{H}_0 \). The inner product \((\cdot, \cdot)_\omega\) or the norm induced by it is sometimes called the \( H^{1/2} \) metric or the \( H^{1/2} \) norm on the space \( H \).

One can verify that \( \mathcal{B}_\omega \) is an orthonormal basis of \( \mathbb{H}_\omega \). From the definition of the inner product \((\cdot, \cdot)_\omega\), we have the relation \( \omega(u, v) = (u, J\overline{v})_\omega \) for any \( u, v \in H \).

This can be used to extend the form \( \omega \) to \( \mathbb{H}_\omega \).

Finally, from the non–degeneracy of the inner product \((\cdot, \cdot)_\omega\), we see that the form \( \omega(\cdot, \cdot) \) on \( \mathbb{H}_\omega \) is also non–degenerate.

### 3. An Infinite-dimensional Symplectic Group

**Definition 3.1.** Let \( B(\mathbb{H}_\omega) \) be the space of bounded operators on \( \mathbb{H}_\omega \) equipped with the operator norm. For an operator \( A \in B(\mathbb{H}_\omega) \)

1. suppose \( \bar{A} \) is an operator on \( \mathbb{H}_\omega \) satisfying \( \bar{A}u = \overline{Au} \) for any \( u \in \mathbb{H}_\omega \), then \( \bar{A} \) is the **conjugate** of \( A \);
2. suppose \( A^\dagger \) is an operator on \( \mathbb{H}_\omega \) satisfying \( (Au, v)_\omega = (u, A^\dagger v)_\omega \) for any \( u, v \in \mathbb{H}_\omega \), then \( A^\dagger \) is the **adjoint** of \( A \);
3. then \( A^{\dagger\dagger} = A^\dagger \) is the **transpose** of \( A \);
4. suppose \( A^\# \) is an operator on \( \mathbb{H}_\omega \) satisfying \( \omega(Au, v) = \omega(u, A^\# v) \) for any \( u, v \in \mathbb{H}_\omega \), then \( A^\# \) is the **symplectic adjoint** of \( A \).
5. \( A \) is said to **preserve the form** \( \omega \) if \( \omega(Au, Av) = \omega(u, v) \) for any \( u, v \in \mathbb{H}_\omega \).
In the orthonormal basis $\mathcal{B}_\omega$, an operator $A \in B(H_{\omega})$ can be represented by an infinite-dimensional matrix, still denoted by $A$, with $(m,n)$th entry equal to $A_{m,n} = (A\tilde{e}_n, \tilde{e}_m)_{\omega}$.

**Remark 3.2.** If we represent an operator $A \in B(H_{\omega})$ by a matrix $\{A_{m,n}\}_{m,n \in \mathbb{Z}\setminus\{0\}}$, the indices $m$ and $n$ are allowed to be both positive and negative following Definition 2.12 of $\mathcal{B}_\omega$.

The next proposition collects some simple facts about operations on $B(H_{\omega})$ introduced in Definition 3.1.

**Proposition 3.3.** Let $A, B \in B(H_{\omega})$. Then

1. $\bar{e}_n = i\tilde{e}_n$, $J\tilde{e}_n = i\text{sgn}(n)\tilde{e}_n$, $(\tilde{e}_n)' = i\tilde{e}_n$;
2. $(A\bar{)}_{m,n} = A_{-m,-n}$;
3. $(A\bar{)}_{m,n} = A_{n,m}$;
4. $\bar{A}^T = \overline{A^T}$, and $(A^T)_{m,n} = A_{-n,-m}$;
5. if $A = \bar{A}$, then $(A\#)_{m,n} = \text{sgn}(mn)A_{n,m}$;
6. $\overline{AB} = \bar{A}\bar{B}$, $(AB)^\dagger = B^\dagger \bar{A}^\dagger$, $(AB)^T = B^T A^T$, $(AB)\# = B\# A\#$;
7. if $A$ is invertible, then $\bar{A}$, $A^T, A^\dagger, A\#$ are all invertible, and $(A^-1) = \overline{A^-1}$, $(A^T)^{-1} = (A^{-1})^T$, $(A)^{-1} = (A^{-1})^\dagger$, $(A\#)^{-1} = (A^{-1}\#)^\dagger$;
8. $(\pi^+)_{m,n} = \frac{1}{2}(\delta_{mn} + \text{sgn}(m)\delta_{mn}), (\pi^-)_{m,n} = \frac{1}{2}(\delta_{mn} - \text{sgn}(m)\delta_{mn})$, $\pi^+ = \pi^+$, $\pi^- = \pi^-$;
9. $J_{m,n} = i\text{sgn}(m)\delta_{mn}$, $J = \bar{J}$, $J = i(\pi^+ - \pi^-)$, $J^T = -J$, $J^\dagger = -J$;
10. $(A\#)_{m,n} = \text{sgn}(mn)A_{-n,-m}$.

**Proof.** All of these properties can be checked by straightforward calculations. We only prove (10).

$$(A\#)_{m,n} = (A\# \tilde{e}_n, \tilde{e}_m)_{\omega} = -\omega(A\# \tilde{e}_n, J\bar{e}_m) = \omega(J\bar{e}_m, A\# \tilde{e}_n)$$
$$= \omega(A\bar{e}_m, \tilde{e}_n) = -\omega(\tilde{e}_n, A\bar{e}_m) = -\omega(\tilde{e}_n, J(-J)A\bar{e}_m)$$
$$= -\omega(\tilde{e}_n, J(-J)\bar{e}_m),$$

where in the last equality we used property (6), $\overline{AB} = \bar{A}\bar{B}$, and property (9), $J = \bar{J}$, so that $-J\bar{A}\bar{e}_m = -J\bar{A}\bar{e}_m = -J\bar{A}\bar{e}_m$. Therefore,

$$(A\#)_{m,n} = -\omega(\tilde{e}_n, J(-J)\bar{e}_m) = \bar{(\tilde{e}_n, -J\bar{A}\bar{e}_m)_{\omega}} = \bar{(-\tilde{e}_n, J\bar{A}\bar{e}_m)_{\omega}}$$
$$= -(J\tilde{e}_n, J\bar{A}\bar{e}_m)_{\omega} = (\tilde{e}_n, J\bar{A}\bar{e}_m)_{\omega} = i\text{sgn}(n)\tilde{e}_n, \overline{A\text{sgn}(m)\tilde{e}_m)_{\omega}}$$
$$= \text{sgn}(mn)(\tilde{e}_n, \tilde{A}\bar{e}_m)_{\omega} = \text{sgn}(mn)(A\tilde{e}_m, \tilde{e}_n)_{\omega} = \text{sgn}(mn)(A)_{n,m}$$
$$= \text{sgn}(mn)A_{-n,-m}.$$
If $A, B \in B(\mathbb{H}_\omega)$, then the block matrix representation for $AB$ is exactly the multiplication of block matrices for $A$ and $B$.

**Proposition 3.5.** Suppose $A \in B(\mathbb{H}_\omega)$ with the matrix $\{A_{m,n}\}_{m,n \in \mathbb{Z}\setminus\{0\}}$. Then the following are equivalent

1. $A = A$;
2. if $u = \bar{u}$, then $Au = \bar{A}u$;
3. $A_{m,n} = \bar{A}_{-m,-n}$ (3.2);
4. as a block matrix, $A$ has the form $\begin{pmatrix} a & b \\ b & \bar{a} \end{pmatrix}$.

**Proof.** Equivalence of (1), (3) and (4) follows from Proposition 3.3 and Notation 3.4. First we show that (1) is equivalent to (2).

$(1) \Rightarrow (2)$. If $u = \bar{u}$, then $Au = \bar{Au} = \bar{A}\bar{u}$.

$(2) \Rightarrow (1)$. Let $u = \bar{\bar{e}}_n + \bar{\bar{e}}_m$, and $v = \bar{\bar{e}}_{-n} + \bar{\bar{e}}_{-m}$. Then $u, v$ are real-valued functions on the circle. Using Proposition 3.3 we have $\bar{\bar{e}}_n = i\bar{\bar{e}}_{-n}$, and therefore $Au = \bar{A}\bar{u}$ and $Av = \bar{A}v$ imply

$$A\bar{e}_n + iA\bar{e}_{-n} = \bar{A}\bar{e}_n - i\bar{A}\bar{e}_{-n}$$

$$A\bar{e}_n - iA\bar{e}_{-n} = -\bar{A}\bar{e}_n + i\bar{A}\bar{e}_{-n}.$$ 

Solving the above two equations for $A\bar{e}_n$, we have

$$A\bar{e}_n = -i\bar{A}\bar{e}_{-n} = \bar{A}\bar{e}_n$$

with this being true for any $n \neq 0$, and so $A = \bar{A}$.

**Proposition 3.6.** Let $A \in B(\mathbb{H}_\omega)$. The following are equivalent:

1. $A$ preserves the form $\omega$;
2. $\omega(Au, Av) = \omega(u, v)$ for any $u, v \in \mathbb{H}_\omega$;
3. $\omega(A\bar{e}_m, A\bar{e}_n) = \omega(\bar{e}_m, \bar{e}_n)$ for any $m, n \neq 0$;
4. $A^TJA = J$;
5. $\sum_{k \neq n} \text{sgn}(mk)A_{k,n}A_{-k,-n} = \delta_{m,n}$ for any $m, n \neq 0$.

If we further assume that $A = \bar{A}$, then the following two are equivalent to the above:

1. $a^T\bar{a} - b^T\bar{b} = 0$ and $a^T\bar{b} - b^T\bar{a} = 0$;
2. $\sum_{k \neq n} \text{sgn}(mk)A_{k,n}A_{-k,n} = \delta_{m,n}$ for any $m, n \neq 0$.

**Proof.** Equivalence of (1), (2) and (3) follows directly from Definition 3.1. Let us check the equivalency of (2) and (4). First assume that (2) holds. By Remark 2.13 we have $\omega(u, v) = (u, Jv)_\omega$, and therefore

$$\omega(Au, Av) = (Au, J\bar{Av})_\omega = (u, A^TJ\bar{A}v)_\omega.$$ 

By assumption, $\omega(Au, Av) = \omega(u, v)$ for any $u, v \in \mathbb{H}_\omega$. So by the non-degeneracy of the inner product $(\cdot, \cdot)_\omega$, we have $A^TJ\bar{A}v = Jv$ for any $v \in \mathbb{H}_\omega$. By definition of $A$, we have $\bar{A}v = \bar{A}v$. So $A^TJ\bar{A}v = Jv$ for any $v \in \mathbb{H}_\omega$, or $A^TJA = J$. Taking conjugation of both sides and using $J = J$, we see that $A^TJA = J$.

Every step above is reversible, therefore we have implication in the other direction as well.
Now we check the equivalency of (3) and (5). First, by Remark 2.13 $\omega(u, v) = (u, J\bar{\omega})_\omega$ and Proposition 3.3

$$\omega(\bar{e}_m, \bar{e}_n) = (\bar{e}_m, J\bar{e}_n)_\omega = -\text{sgn}(m)\delta_{m,-n}. $$

On the other hand, by the continuity of the form $\omega(\cdot, \cdot)$ in both variables, we have

$$\omega(A\bar{e}_m, A\bar{e}_n) = \omega\left(\sum_k A_{k,m}\bar{e}_k, \sum_k A_{k,n}\bar{e}_l\right)$$

$$= \sum_{k,l} A_{k,m}A_{l,n}(\text{sgn}(k))\delta_{k,-l} = -\sum_k \text{sgn}(k)A_{k,m}A_{-k,n}. $$

Now assuming $\omega(A\bar{e}_m, A\bar{e}_n) = \omega(\bar{e}_m, \bar{e}_n)$, we have

$$-\sum_k \text{sgn}(k)A_{k,m}A_{-k,n} = -\text{sgn}(m)\delta_{m,-n}, \text{ for any } m, n \neq 0. $$

By multiplying by $\text{sgn}(m)$ both sides, and replacing $-n$ with $n$, we get (5). Conversely, note that every step above is reversible, therefore we have implication in the other direction.

We have proved equivalence of (1)-(5). Now assume $A = \bar{A}$. To prove equivalence of (4) and (I), just notice that as block matrices, $A, A^T$ and $J$ have the form

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad \begin{pmatrix} a^T & b^T \\ b^T & a^T \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \pi^+ & 0 \\ 0 & -\pi^+ \end{pmatrix}. $$

Equivalence of (5) and (I) can be checked by

$$\text{(1)} \quad A_{\pi^+} - b a^+ = \pi^- \text{ and } b a^+ - a b^+ = 0.$$  

$$\text{(2)} \quad \sum_k \text{sgn}(k)A_{m,k}A_{n,-k} = \delta_{m,n} \text{ for any } m, n \neq 0. $$

\textbf{Proposition 3.7.} Let $A \in B(\mathbb{H}).$ If $A$ preserves the form $\omega$, then the following are equivalent:

1. $A$ is invertible.
2. $AJA^T = J.$
3. $A^T$ preserves the form $\omega.$
4. $\sum_k \text{sgn}(mk)A_{m,k}A_{n,-k} = \delta_{m,n} \text{ for any } m, n \neq 0.$

If we further assume that $A = \bar{A}$, then the following are equivalent to the above:

1. $\bar{a}a^T - b b^T = \pi^- \text{ and } b a^+ - \bar{a} b^+ = 0.$
2. $\sum_k \text{sgn}(mk)A_{m,k}\bar{A}_{n,k} = \delta_{m,n} \text{ for any } m, n \neq 0.$

\textbf{Proof.} We will use several times the fact that if $A$ preserves $\omega$, then $A^TJA = J.$

[1] $\Rightarrow$ [2] Multiplying on the left by $(A^T)^{-1}$ and multiplying on the right by $A^{-1}$ both sides, we get $J = (A^T)^{-1}JA^{-1}$, and so $(A^{-1})^TJA^{-1} = J$. Taking inverse of both sides, and using $J^{-1} = -J$, we have $A^TJA = J.$

[(2) $\Rightarrow$ (1)] As $J$ is injective, so is $A^TJA$, and therefore $A$ is injective. On the other hand, by assumption $AJA^T = J$. As $J$ is surjective, so $AJA^T$ is surjective too. This implies that $A$ is surjective, and therefore $A$ is invertible.

Equivalence of (2) and (3) follows from $(A^T)^T = A$ and Proposition 3.6. Equivalence of (3) and (4) follows directly from Proposition 3.6 and the fact that $(A^T)_{m,n} = A_{-n,m}.$

Now assume that $A = \bar{A}$. Then equivalence of (3) and (I) can be checked by using multiplication of block matrices as in the proof of Proposition 3.6. Finally (4) is equivalent to (II) as if $A = \bar{A}$, then $A_{-n,-m} = \bar{A}_{n,m}$. □
Corollary 3.8. Let $A \in B(\mathbb{H}_\omega)$ and $A = \bar{A}$. Then the following are equivalent:

1. $A$ preserves the form $\omega$ and is invertible;
2. $A^\# A = A^\# \bar{A} = id$;

Proof. By Proposition 3.3

$$(A^\# A)_{m,n} = \sum_{k \neq 0} (A^\#)_{m,k} A_{k,n} = \sum_{k \neq 0} \text{sgn}(mk) A_{k,n} \bar{A}_{k,m},$$

$$(A A^\#)_{m,n} = \sum_{k \neq 0} A_{m,k} (A^\#)_{k,n} = \sum_{k \neq 0} \text{sgn}(mk) A_{m,k} \bar{A}_{n,k}.$$ 

Therefore, by (II) in Proposition 3.6 and (II) in Proposition 3.7 we have equivalence. □

Definition 3.9. Define a (semi)norm $\| \cdot \|_2$ on $B(\mathbb{H}_\omega)$ such that for $A \in B(\mathbb{H}_\omega)$,

$$\|A\|_2^2 = \text{Tr}(b^\dagger b) = \|b\|_{HS},$$

where $b = \pi^+ A \pi^-$. That is, the norm $\|A\|_2$ is just the Hilbert-Schmidt norm of the block $b$.

Definition 3.10. An infinite-dimensional symplectic group $\text{Sp}(\infty)$ is the set of bounded operators $A$ on $H$ such that

1. $A$ is invertible;
2. $A = \bar{A}$;
3. $A$ preserves the form $\omega$;
4. $\|A\|_2 < \infty$.

Remark 3.11. If $A$ is a bounded operator on $H$, then $A$ can be extended to a bounded operator on $\mathbb{H}_\omega$. Therefore, we can equivalently define $\text{Sp}(\infty)$ to be the set of operators $A \in B(\mathbb{H}_\omega)$ such that

1. $A$ is invertible;
2. $A = \bar{A}$;
3. $A$ preserves the form $\omega$;
4. $\|A\|_2 < \infty$.
5. $A$ is invariant on $H$, i.e., $A(H) \subseteq H$.

Remark 3.12. By Corollary 3.8, the definition of $\text{Sp}(\infty)$ is also equivalent to

1. $A = \bar{A}$;
2. $A^\# A = AA^\# = id$;
3. $\|A\|_2 < \infty$.

Proposition 3.13. $\text{Sp}(\infty)$ is a group.

Proof. First we show that if $A \in \text{Sp}(\infty)$, then $A^{-1} \in \text{Sp}(\infty)$. By the assumption on $A$, it is easy to verify that $A^{-1}$ satisfies (1), (2), (3) and (5) in Remark 3.11. We need to show that $A^{-1}$ satisfies the condition (4), i.e. $\|A^{-1}\|_2 < \infty$. Suppose

$$A = \left( \begin{array}{cc} a & b \\ \bar{b} & \bar{a} \end{array} \right) \quad \text{and} \quad A^{-1} = \left( \begin{array}{cc} a' & b' \\ \bar{b'} & \bar{a}' \end{array} \right),$$

where by our assumptions all blocks are bounded operators, and in addition $b$ is a Hilbert-Schmidt operator. We want to prove $b'$ is also a Hilbert-Schmidt operator. $AA^{-1} = I$ and $A^{-1}A = I$ imply that

$$ab' = -b\bar{a}, \quad a'a + b'b = I.$$
The last equation gives $a'ab' + b'b' = b'$, and so

$$b' = a'ab' + b'b' = -a'b\bar{a} + b'b'$$

which is a Hilbert-Schmidt operator as $b$ and $\bar{b}$ are Hilbert-Schmidt. Therefore $\|A^{-1}\|_2 < \infty$ and $A^{-1} \in \text{Sp}(\infty)$.

Next we show that if $A, B \in \text{Sp}(\infty)$, then $AB \in \text{Sp}(\infty)$. By the assumption on $A$ and $B$, it is easy to verify that $AB$ satisfies (1), (2), (3) and (5) in Remark 3.11. We need to show that $AB$ satisfies the condition (4), i.e. $\|AB\|_2 < \infty$. Suppose

$$A = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} c & d \\ \bar{d} & \bar{c} \end{pmatrix},$$

where all blocks are bounded, and $\|b\|_{HS}, \|d\|_{HS} < \infty$. Then

$$AB = \begin{pmatrix} ac + bd & ad + bc \\ \bar{b}c + \bar{a}d & \bar{b}d + \bar{a}c \end{pmatrix}.$$

Then

$$\|AB\|_2^2 = \|ad + bc\|_{HS} \leq \|ad\|_2 + \|bc\|_{HS} < \infty,$$

since both $ad$ and $bc$ are Hilbert-Schmidt operators. Therefore $\|AB\|_2 < \infty$ and $AB \in \text{Sp}(\infty)$.

4. Symplectic Representation of $\text{Diff}(S^1)$

**Definition 4.1.** Let $\text{Diff}(S^1)$ be the group of orientation preserving $C^\infty$ diffeomorphisms of $S^1$. $\text{Diff}(S^1)$ acts on $H$ as follows

$$(\phi,u)(\theta) = u(\phi^{-1}(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u(\phi^{-1}(\theta)) d\theta.$$

Note that if $u \in H$ is real-valued, then $\phi.u$ is real-valued as well.

**Proposition 4.2.** The action of $\text{Diff}(S^1)$ on $H$ gives a group homomorphism

$$\Phi : \text{Diff}(S^1) \to \text{Aut} H$$

defined by $\Phi(\phi)(u) = \phi.u$, for $\phi \in \text{Diff}(S^1)$ and $u \in H$, where $\text{Aut} H$ is the group of automorphisms on $H$.

**Proof.** Let $u \in H$, then $\phi.u$ is a $C^\infty$ function with the mean value 0, and so $\phi.u \in H$. It is also clear that $\phi(u + v) = \phi.u + \phi.v$ and $\phi(\lambda u) = \lambda \phi.u$. So $\Phi$ is well-defined as a map from $\text{Diff}(S^1)$ to $\text{End} H$, the space of endomorphisms on $H$. Now let us check that $\Phi$ is a group homomorphism. Suppose $\phi, \psi \in \text{Diff}(S^1)$ and $u \in H$, then

$$\Phi(\phi\psi)(u)(\theta) = u((\phi\psi)^{-1}(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u((\phi\psi)^{-1}(\theta)) d\theta$$

$$= u((\psi^{-1}\phi^{-1})(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u((\psi^{-1}\phi^{-1})(\theta)) d\theta.$$
On the other hand,
\[
\Phi(\phi)\Phi(\psi)(u)(\theta) = \Phi(\phi) \left[ u(\psi^{-1}(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u(\psi^{-1}(\theta))d\theta \right] \\
= \Phi(\phi) \left[ u(\psi^{-1}(\theta)) \right] = u((\psi^{-1}\phi^{-1})(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u((\psi^{-1}\phi^{-1})(\theta))d\theta.
\]
So \(\Phi(\phi\psi) = \Phi(\phi)\Phi(\psi)\). In particular, the image of \(\Phi\) is in the \(\text{Aut}\ H\).

**Lemma 4.3.** Any \(\phi \in \text{Diff}(S^1)\) preserves the form \(\omega\), that is, \(\omega(\phi.u, \phi.v) = \omega(u, v)\) for any \(u, v \in H\).

**Proof.** By Definition 4.1 \(\phi.u = u(\psi) - u_0, \phi.v = v(\psi) - v_0\), where \(\psi = \phi^{-1}\) and \(u_0, v_0\) are the constants. Then
\[
\omega(\phi.u, \phi.v) = \omega(u(\psi) - u_0, v(\psi) - v_0) \\
= \frac{1}{2\pi} \int_0^{2\pi} (u(\psi(\theta)) - u_0)(v(\psi(\theta)) - v_0)d\theta \\
= \frac{1}{2\pi} \int_0^{2\pi} u(\psi)v'(\psi)\psi'(\theta)d\theta - \frac{1}{2\pi} \int_0^{2\pi} u_0v(\psi)d\theta \\
= \frac{1}{2\pi} \int_0^{2\pi} u(\psi)v'(\psi)d\psi \\
= \omega(u, v).
\]

We are going to prove that a diffeomorphism \(\phi \in \text{Diff}(S^1)\) acts on \(H\) as a bounded linear map, and that \(\Phi(\phi)\) is in \(\text{Sp}(\infty)\). The next lemma is a generalization of a proposition in a paper of G. Segal[9].

**Lemma 4.4.** Let \(\psi \neq id \in \text{Diff}(S^1)\) and \(\phi = \psi^{-1}\). Let
\[
I_{n,m} = (\psi.e^{im\theta}, e^{in\theta}) = \frac{1}{2\pi} \int_0^{2\pi} e^{im\phi-in\theta}d\theta.
\]
Then
\[
(1) \quad \sum_{n>0,m<0} |n||I_{n,m}|^2 < \infty, \quad \text{and} \quad \sum_{m>0,n<0} |n||I_{n,m}|^2 < \infty.
\]
\[
(2) \quad \text{For sufficiently large } |m| \text{ there is a constant } C \text{ independent of } m \text{ such that} \\
\sum_{n \neq 0} |n||I_{n,m}|^2 < C|m|.
\]

**Proof.** Let
\[
m_{\phi'} = \min\{\phi'(\theta)|\theta \in S^1\}; \text{ and } M_{\phi'} = \max\{\phi'(\theta)|\theta \in S^1\}.
\]
Since \(\phi\) is a diffeomorphism, we have \(0 < m_{\phi'} < M_{\phi'} < \infty\).

Take four points \(a, b, c, d\) on the unit circle such that \(a\) corresponds to \(m_{\phi'}\) in the sense \(\tan(a) = m_{\phi'}\), \(b\) corresponds to \(M_{\phi'}\) in the sense \(\tan(b) = M_{\phi'}\), \(c\) is opposite to \(a\), i.e., \(c = a + \pi\), \(d\) is opposite to \(b\), i.e., \(d = b + \pi\). The four points on the circle are arranged in the counter-clockwise order, and \(0 < a < b < \frac{\pi}{2}, \pi < c < d < \frac{3\pi}{2}\).
Let $\tau \in S^1$ such that $\tau \neq \frac{\pi}{2},\frac{3\pi}{2}$. Define a function $\phi_\tau$ on $S^1$ by
\[
\phi_\tau(\theta) = \frac{\cos \tau \cdot \phi(\theta) - \sin \tau \cdot \theta}{\cos \tau - \sin \tau}.
\]

We will show that if $\tau \in (b,c)$ or $\tau \in (d,a)$, then $\phi_\tau$ is an orientation preserving diffeomorphism of $S^1$, where $(b,c)$ is the open arc from the point $b$ to the point $c$, and $(d,a)$ is the open arc from the point $d$ to the point $a$.

Clearly $\phi_\tau$ is a $C^\infty$ function on $S^1$. Also, $\phi_\tau(0) = 0$ and $\phi_\tau(2\pi) = 2\pi$. Taking derivative with respect to $\theta$, we have
\[
\phi'_\tau(\theta) = \frac{\cos \tau \cdot \phi'(\theta) - \sin \tau \cdot \theta}{\cos \tau - \sin \tau}.
\]

By the choice of $\tau$, we can prove that $\phi'_\tau(\theta) > 0$. Therefore, $\phi_\tau$ is an orientation preserving diffeomorphism as claimed.

Let $m, n \in \mathbb{Z}\setminus\{0\}$. Let $\tau_{mn} = \text{Arg}(m + in)$, i.e., the argument of the complex number $m + in$, considered to be in $[0, 2\pi]$. Then we have $m\phi - n\theta = (m-n)\phi_{\tau_{mn}}$.

If $\tau_{mn} \in (b,c)$, then $\phi_{\tau_{mn}}$ is a diffeomorphism. Let $\psi_{\tau_{mn}} = \phi^{-1}_{\tau_{mn}}$. Then
\[
I_{n,m} = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)\phi_{\tau_{mn}}} d\theta = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)\theta} \psi'_{\tau_{mn}}(\theta) d\theta,
\]
where the last equality is by change of variable. On integration by parts $k$ times, we have
\[
I_{n,m} = \left(\frac{1}{i(m-n)}\right)^k \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)\theta} \psi^{(k+1)}_{\tau_{mn}}(\theta) d\theta.
\]

Let $\alpha = [\alpha_0, \alpha_1]$ be a closed arc contained in the arc $(b,c)$. Let $S_\alpha$ be the set of all pairs of nonzero integers $(m,n)$ such that $\alpha_0 < \tau_{mn} < \alpha_1$, where $\tau_{mn} = \text{Arg}(m + in)$. We are going to consider an upper bound of the sum $\sum_{(m,n) \in S_\alpha} |n||I_{n,m}|^2$.

For the pair $(m,n)$, if $|m-n| = p$, the condition $\alpha_0 < \tau_{mn} < \alpha_1$ gives us both an upper bound and a lower bound for $n$:
\[
\frac{m\phi'}{m\phi' - 1}p \leq n \leq \frac{n\phi'}{m\phi' - 1}p.
\]

So $|n| \leq C_1p$ where $C_1$ is a constant which does not depend on the pair $(m,n)$. Also, the number of pairs $(m,n) \in S_\alpha$ such that $|m-n| = p$ is bounded by $C_2p$ for some constant $C_2$. Let $C_3 = \max \left\{|\psi^{(k+1)}_{\tau}(\theta)\} : \theta \in S^1, \tau \in [\alpha_0, \alpha_1]\right\}$. Then
\[
|I_{n,m}| \leq C_3 \left|\frac{1}{i(m-n)}\right|^k \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)\theta} d\theta = C_3 p^{-k}.
\]

Therefore,
\[
\sum_{(m,n) \in S} |n||I_{n,m}|^2 = \sum_p \sum_{(m,n) \in S_n, |m-n|=p} |n||I_{n,m}|^2 \\
\leq \sum_p C_1p \cdot C_3^2 p^{-2k} \cdot C_2p = C_0 \sum_p p^{-(2k-2)},
\]
where the constant $C_0$ depends on the arc $\alpha$. 


Similarly, for a closed arc $\beta = [\beta_0, \beta_1]$ contained in the arc $(d, a)$, we have

$$\sum_{(m,n) \in S_\beta} |n||I_{n,m}|^2 \leq C_\beta \sum_{p} p^{-(2k-2)},$$

where the constant $C_\beta$ depends on the arc $\beta$.

Now let $\alpha = [\pi/2, \pi]$, and $\beta = [3\pi/2, 2\pi]$. Then $\alpha$ is contained in $(b, c)$ and $\beta$ is contained in $(d, a)$. We have

$$\sum_{n>0,m<0} |n||I_{n,m}|^2 = C_\alpha \cdot \sum_{p} p^{-(2k-2)} < \infty$$

and

$$\sum_{n<0,m>0} |n||I_{n,m}|^2 = C_\beta \cdot \sum_{p} p^{-(2k-2)} < \infty,$$

which proves (1) of the lemma.

To prove (2), we let $\alpha = [\alpha_0, \alpha_1]$ be a closed arc contained in the arc $(b, c)$ such that $b < \alpha_0 < \pi/2$ and $\pi < \alpha_1 < c$, and $\beta = [\beta_0, \beta_1]$ be a closed arc contained in the arc $(d, a)$ such that $d < \beta_0 < 3\pi/2$ and $0 < \beta_1 < a$. Then we have

$$\sum_{(m,n) \in S_\alpha} |n||I_{n,m}|^2 + \sum_{(m,n) \in S_\beta} |n||I_{n,m}|^2 \leq C_{\alpha\beta}$$

for some constant $C_{\alpha\beta}$.

Let $m > 0$ be sufficiently large, and $N_m$ be the largest integer less than or equal to $m \tan(\alpha_0)$,

$$\sum_{0<n\leq N_m} |I_{n,m}|^2 \leq \sum_{n \neq 0} |I_{n,m}|^2.$$

Note that $I_{n,m}$ is the $n$th Fourier coefficient of $\psi.e^{im\theta}$. Therefore,

$$\sum_{n \neq 0} |I_{n,m}|^2 = ||\psi.e^{im\theta}||_{L^2}$$

which is bounded by a constant $K$. Therefore,

$$\sum_{0<n\leq N_m} |n||I_{n,m}|^2 \leq K m \tan (\alpha_0).$$

On the other hand,

$$\sum_{n<0} |n||I_{n,m}|^2 + \sum_{n>N_m} |n||I_{n,m}|^2 \leq \sum_{(m,n) \in S_\alpha} |n||I_{n,m}|^2 + \sum_{(m,n) \in S_\beta} |n||I_{n,m}|^2 = C_{\alpha\beta}.$$

Therefore,

$$\sum_{n \neq 0} |n||I_{n,m}|^2 \leq C_{\alpha\beta} + K m \tan(\alpha_0) \leq m C_+,$$

where $C_+$ can be chosen to be, for example, $K \tan(\alpha_0) + C_{\alpha\beta}$, which is independent of $m$.

Similarly, for $m < 0$ with sufficiently large $|m|$,

$$\sum_{n \neq 0} |n||I_{n,m}|^2 \leq m C_-.$$
Let $C = \max\{C_+, C_\cdot\}$. Then we have, for sufficiently large $|m|$, 
\[
\sum_{n \neq 0} |n||I_{n,m}|^2 \leq |m|C,
\]
which proves (2) of the lemma. \(\square\)

**Lemma 4.5.** For any $\psi \in \text{Diff}(S^1)$, $\Phi(\psi) \in B(H)$, the space of bounded linear maps on $H$. Moreover, 
\[
\|\Phi(\psi)\| \leq C, \quad \|\Phi(\psi)\|_2 \leq C,
\]
where $C$ is the constant in Equation 4.1.

**Proof.** First observe that the operator norm of $\Phi(\psi)$ is 
\[
\|\Phi(\psi)\| = \sup\{\|\psi \cdot u\|_\omega \mid u \in H, \|u\|_\omega = 1\}.
\]
For any $u \in H$, let $\tilde{\psi}$ be its Fourier coefficients, that is $\tilde{\psi}(n) = (u, \tilde{e}_n)$, and let $\tilde{u}$ be defined by $\tilde{u} = (u, \tilde{e}_n)$. (2.10.2.12). It can be verified that the relation between $\tilde{u}$ and $\tilde{\psi}$ is: if $n > 0$, then $\tilde{\psi}(n) = \sqrt{n}\tilde{u}(n)$; if $n < 0$, then $\tilde{\psi}(n) = i\sqrt{|n|}\tilde{u}(n)$. We have 
\[
\|u\|_\omega^2 = (u, u) = (\tilde{u}, \tilde{u})_2 = \sum_{n > 0} |\tilde{u}(n)|^2 = \sum_{n \neq 0} |n||\tilde{u}(n)|^2.
\]
Let $\phi = \psi^{-1}$. We have $u(\phi) = \sum_{m \neq 0} \tilde{u}(m)e^{im\phi}$. Using the notation $I_{n,m}$ (4.4), we have 
\[
\|\psi \cdot u\|_\omega^2 = \sum_{n \neq 0} |n||\tilde{\psi} \cdot u(n)|^2 = \sum_{n \neq 0} |n| \left| \frac{1}{2\pi} \int_0^{2\pi} u(\phi(\theta))e^{-in\theta} d\theta \right|^2
\]
\[
= \sum_{n \neq 0} |n| \left| \frac{1}{2\pi} \int_0^{2\pi} \tilde{u}(m)e^{im\phi}e^{-in\theta} d\theta \right|^2
\]
\[
= \sum_{n \neq 0} |n| \left| \sum_{m \neq 0} \tilde{u}(m) \frac{1}{2\pi} \int_0^{2\pi} e^{im\phi-in\theta} d\theta \right|^2
\]
\[
= \sum_{n \neq 0} |n| \left| \sum_{m \neq 0} \tilde{u}(m)I_{n,m} \right|^2
\]
\[
\leq \sum_{m, n \neq 0} |n||\tilde{u}(m)|^2 |I_{n,m}|^2 = \sum_{m \neq 0} |\tilde{u}(m)|^2 \sum_{n \neq 0} |n||I_{n,m}|^2
\]
\[
= \sum_{|n| < M_0} |\tilde{u}(m)|^2 \sum_{n \neq 0} |n||I_{n,m}|^2 + \sum_{|m| > M_0} |\tilde{u}(m)|^2 \sum_{n \neq 0} |n||I_{n,m}|^2,
\]
where the constant $M_0$ in the last equality is a positive integer large enough so that we can apply part (2) of Lemma 4.4. It is easy to see that the first term in the last equality is finite. For the second term we use Lemma 4.4 
\[
\sum_{|n| > M_0} |\tilde{u}(m)|^2 \sum_{n \neq 0} |n||I_{n,m}|^2 \leq C \sum_{|m| > M_0} |\tilde{u}(m)|^2 |m| \leq C.
\]
Thus for any $u \in H$ with $\|u\|_\omega = 1$, $\|\psi \cdot u\|_\omega$ is uniformly bounded. Therefore, $\Phi(\psi)$ is a bounded operator on $H$.

Now we can use Lemma 4.4 again to estimate the norm $\|\Phi(\psi)\|_2$
\[ \|\Phi(\psi)\|_2 = \sum_{n>0, m<0} |(\psi, \varepsilon_m, \varepsilon_n)|^2 = \sum_{n>0, m<0} |n| |(\psi, \varepsilon_m, \varepsilon_n)|^2 = \sum_{n>0, m<0} |n| |I_{n,m}|^2 < \infty. \]

\[ \square \]

**Theorem 4.6.** \( \Phi : \text{Diff}(S^1) \rightarrow \text{Sp}(\infty) \) is a group homomorphism. Moreover, \( \Phi \) is injective, but not surjective.

**Proof.** Combining Lemma 4.3 and Lemma 4.5 we see that for any diffeomorphism \( \psi \in \text{Diff}(S^1) \) the map \( \Phi(\psi) \) is an invertible bounded operator on \( H \), it preserves the form \( \omega \), and \( \|\Phi(\psi)\|_2 < \infty \). In addition, by our remark after Definition 4.1 \( \psi.u \) is real-valued, if \( u \) is real-valued. Therefore, \( \Phi \) maps \( \text{Diff}(S^1) \) into \( \text{Sp}(\infty) \).

Next, we first prove that \( \Phi \) is injective. Let \( \psi_1, \psi_2 \in \text{Diff}(S^1) \), and denote \( \phi_1 = \psi_1^{-1}, \phi_2 = \psi_2^{-1} \). Suppose \( \Phi(\psi_1) = \Phi(\psi_2) \), i.e. \( \psi_1.u = \psi_2.u \), for any \( u \in H \). In particular, \( \psi_1.e^{i\theta} = \psi_2.e^{i\theta} \). Therefore
\[ e^{i\phi_1} - C_1 = e^{i\phi_2} - C_2, \]
where \( C_1 = \frac{1}{2\pi} \int_0^{2\pi} e^{i\phi_1} d\theta \), and \( C_2 = \frac{1}{2\pi} \int_0^{2\pi} e^{i\phi_2} d\theta \). Note that \( e^{i\phi_1} \) and \( e^{i\phi_2} \) have the same image as maps from \( S^1 \) to \( \mathbb{C} \). This implies \( C_1 = C_2 \), since otherwise \( e^{i\phi_1} = e^{i\phi_2} + (C_1 - C_2) \). \( e^{i\phi_1} \) and \( e^{i\phi_2} \) would have had different images. Therefore, we have \( e^{i\phi_1} = e^{i\phi_2} \). But the function \( e^{ix} : S^1 \rightarrow S^1 \) is an injective function, so \( \phi_1 = \phi_2 \). Therefore \( \psi_1 = \psi_2 \), and so \( \Phi \) is injective.

To prove that \( \Phi \) is not surjective, we will construct an operator \( A \in \text{Sp}(\infty) \) which cannot be written as \( \Phi(\psi) \) for any \( \psi \in \text{Diff}(S^1) \). Let the linear map \( A \) be defined by the corresponding matrix \( \{A_{m,n}\}_{m,n \in \mathbb{Z}} \) with the entries
\[
A_{1,1} = A_{-1,-1} = \sqrt{2} \\
A_{1,-1} = i, A_{-1,1} = -i \\
A_{m,m} = 1, \text{ for } m \neq \pm 1
\]
with all other entries being 0.

First we show that \( A \in \text{Sp}(\infty) \). For any \( u \in H \), we can write \( u = \sum_{n \neq 0} \tilde{u}(n)\tilde{e}_n \). Then \( A \) acting on \( u \) changes only \( \tilde{e}_1 \) and \( \tilde{e}_{-1} \). Therefore, \( Au \in H \), and clearly \( A \) is a well-defined bounded linear map on \( H \) to \( H \). Moreover, \( \|A\|_2 < \infty \). It is clear that \( A_{m,n} = \overline{A_{-m,-n}} \), and therefore \( A = \overline{A} \) by Proposition 3.3. Moreover, \( A \) preserves the form \( \omega \) by part (II) of Proposition 3.6, as
\[ \sum_{k \neq 0} \text{sgn}(mk) A_{k,m} \overline{A_{k,n}} = \delta_{m,n}. \]
Finally, \( A \) is invertible, since \( \{A_{m,n}\}_{m,n \in \mathbb{Z}} \) is, with the inverse \( \{B_{k,m}\}_{m,n \in \mathbb{Z}} \) given by
\[
B_{1,1} = B_{-1,-1} = \sqrt{2} \\
B_{1,-1} = -i, B_{-1,1} = i \\
B_{m,m} = 1, \text{ for } m \neq \pm 1
\]
with all other entries being 0. Next we show that $A \neq \Phi(\psi)$ for any $\psi \in \text{Diff}(S^1)$. First observe that if we look at any basis element $\tilde{e}_1 = e^{i\theta}$ as a function from $S^1$ to $\mathbb{C}$, then the image of this function lies on the unit circle. Clearly, when acted by a diffeomorphism $\phi \in \text{Diff}(S^1)$, the image of the function $\phi e^{i\theta}$ is still a circle with radius 1. But if we consider $A \tilde{e}_1$ as a function from $S^1$ to $\mathbb{C}$, we will show that the image of the function $A \tilde{e}_1 : S^1 \to \mathbb{C}$ is not a circle. Therefore, $A \neq \Phi(\psi)$ for any $\psi \in \text{Diff}(S^1)$. Indeed, by definition of $A$ we have

$$A \tilde{e}_1 = \sqrt{2} \tilde{e}_1 - i \tilde{e}_{-1}.$$ 

Let us write it as a function on $S^1$

$$A \tilde{e}_1(\theta) = \sqrt{2} e^{i\theta} - e^{-i\theta} = (\sqrt{2} - 1) \cos \theta + i(\sqrt{2} + 1) \sin \theta,$$

and then we see that the image lies on an ellipse, which is not the unit circle

$$\frac{x^2}{(\sqrt{2} - 1)^2} + \frac{y^2}{(\sqrt{2} + 1)^2} = 1.$$ 

5. The Lie Algebra Associated With $\text{Diff}(S^1)$

Let $\text{diff}(S^1)$ be the space of smooth vector fields on $S^1$. Elements in $\text{diff}(S^1)$ can be identified with smooth functions on $S^1$. The space $\text{diff}(S^1)$ is a Lie algebra with the following Lie bracket

$$[X, Y] = XY' - X'Y, \quad X, Y \in \text{diff}(S^1),$$

where $X'$ and $Y'$ are derivatives with respect to $\theta$.

Let $X \in \text{diff}(S^1)$, and $\rho_t$ be the corresponding flow of diffeomorphisms. We define an action of $\text{diff}(S^1)$ on $H$ as follows: for $X \in \text{diff}(S^1)$ and $u \in H$, $X.u$ is a function on $S^1$ defined by

$$(X.u)(\theta) = \left. \frac{d}{dt} \right|_{t=0} ([\rho_t, u](\theta)],$$

where $\rho_t$ acts on $u$ via the representation $\Phi : \text{Diff}(S^1) \to \text{Sp}(\infty)$.

The next proposition shows that the action is well-defined, and also gives an explicit formula of $X.u$.

**Proposition 5.1.** Let $X \in \text{diff}(S^1)$. Then

$$(X.u)(\theta) = u'(\theta)(-X(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u'(\theta)(-X(\theta)) d\theta,$$

that is, $X.u$ is the function $-u'X$ with the 0th Fourier coefficient replaced by 0.

**Proof.** Let $\rho_t$ be the flow that corresponds to $X$, and $\lambda_t$ be the flow that corresponds to $-X$. Then $\lambda_t$ is the inverse of $\rho_t$ for all $t$.

$$(X.u)(\theta) = \left. \frac{d}{dt} \right|_{t=0} ([\rho_t, u](\theta)] = \left. \frac{d}{dt} \right|_{t=0} [u(\lambda_t(\theta)) - \frac{1}{2\pi} \int_0^{2\pi} u(\lambda_t(\theta)) d\theta].$$
Using the chain rule, we have
\[ \frac{d}{dt} \bigg|_{t=0} u(\lambda_t(\theta)) = u'(\theta)(-\tilde{X}(\theta)), \]
and
\[ \frac{d}{dt} \bigg|_{t=0} \frac{1}{2\pi} \int_0^{2\pi} u(\lambda_t(\theta)) d\theta = \frac{1}{2\pi} \int_0^{2\pi} u'(\theta)(-X(\theta)) d\theta. \]

\[ \square \]

\textbf{Notation 5.2.} We consider $\text{diff}(S^1)$ as a subspace of the space of real-valued $L^2$ functions on $S^1$. The space of real-valued $L^2$ functions on $S^1$ has an orthonormal basis
\[ B = \{ X_l = \cos(m\theta), Y_k = \sin(k\theta), l = 0, 1, ..., k = 1, 2, ... \} \]
which is contained in $\text{diff}(S^1)$.

Let us consider how these basis elements act on $H$.

\textbf{Proposition 5.3.} For any $l = 0, 1, ..., k = 1, 2, ...$ the basis elements $X_l, Y_k$ act on $H$ as linear maps. In the basis $B_\omega$ of $H$, they are represented by infinite dimensional matrices with $(m, n)$th entries equal to
\[ (X_l)_{m,n} = (X_l, \tilde{e}_n, \tilde{e}_m)_\omega = s(m, n) \frac{1}{2} \sqrt{|mn|} (\delta_{m-n,l} + \delta_{n-m,l}) \]
\[ (Y_k)_{m,n} = (Y_k, \tilde{e}_n, \tilde{e}_m)_\omega = s(m, n)(-i) \frac{1}{2} \sqrt{|mn|} (\delta_{m-n,k} - \delta_{n-m,k}) \]
where $m, n \neq 0$,
\[ s(m, n) = \begin{cases} -i & m, n > 0 \\ 1 & m > 0, n < 0 \\ 1 & m < 0, n > 0 \\ i & m, n < 0. \end{cases} \]

\textbf{Proof.} By Proposition 5.1 and a simple verification depending on the signs of $m, n$ we see that
\[ X_l e^{in\theta} = -ine^{in\theta} \cos(l\theta) = -\frac{1}{2} in \left[ e^{i(n+l)\theta} + e^{i(n-l)\theta} \right] \]
\[ Y_k e^{in\theta} = -ine^{in\theta} \sin(k\theta) = -\frac{1}{2} n \left[ e^{i(n+k)\theta} - e^{i(n-k)\theta} \right]. \]
Indeed, recall that a basis element $\tilde{e}_n \in B_\omega$ has the form
\[ \tilde{e}_n = \begin{cases} \frac{1}{\sqrt{n}} e^{in\theta} & n > 0 \\ \frac{i}{\sqrt{|n|}} e^{in\theta} & n < 0. \end{cases} \]
Suppose $m, n > 0$
\[ X_l, \tilde{e}_n = \frac{1}{\sqrt{n}} X_l e^{in\theta} = -\frac{1}{2} i \sqrt{n} \left[ e^{i(n+l)\theta} + e^{i(n-l)\theta} \right], \]
and
\[ (e^{i(n+l)\theta}, \tilde{e}_m)_\omega = \sqrt{m} \delta_{m-n, k}; \quad (e^{i(n-l)\theta}, \tilde{e}_m)_\omega = \sqrt{m} \delta_{n-m, l}. \]
Therefore, 
\[(X_t)_{m,n} = (X_t, e_m, e_n)_\omega = (-i)^2 \sqrt{mn} \left( \delta_{m-n,l} + \delta_{n-m,l} \right).\]

All other cases can be verified similarly. □

**Remark 5.4.** Recall that \(H_\omega\) is the completion of \(H\) under the metric \(\langle \cdot, \cdot \rangle_\omega\). The above calculation shows that the trigonometric basis \(X_t, Y_k\) of \(\text{diff}(S^1)\) act on \(H_\omega\) as unbounded operators. They are densely defined on the subspace \(H \subseteq H_\omega\).

### 6. Brownian Motion on \(\mathfrak{sp}(\infty)\)

**Notation 6.1.** As in [1], let \(\mathfrak{sp}(\infty)\) be the set of infinite-dimensional matrices \(A\) which can be written as block matrices of the form
\[
\begin{pmatrix}
  a & b \\
  b & \bar{a}
\end{pmatrix}
\]
such that \(a + a^\dagger = 0\), \(b = b^T\), and \(b\) is a Hilbert-Schmidt operator.

**Remark 6.2.** The set \(\mathfrak{sp}(\infty)\) has a structure of Lie algebra with the operator commutator as a Lie bracket, and we associate this Lie algebra with the group \(\text{Sp}(\infty)\).

**Proposition 6.3.** Let \(\{A_{m,n}\}_{m,n \in \mathbb{Z} \setminus \{0\}}\) be the matrix corresponding to an operator \(A\). Then any \(A \in \mathfrak{sp}(\infty)\) satisfies (1) \(A_{m,n} = \overline{A_{-m,-n}}\); (2) \(A_{m,n} + A_{n,m} = 0\), for \(m, n > 0\); (3) \(A_{m,n} = A_{-n,-m}\), for \(m > 0, n < 0\).

Moreover, \(A \in \mathfrak{sp}(\infty)\) if and only if (1) \(A = \overline{A}\); (2) \(A^+ A^- = 0\).

**Proof.** The first part follows directly from definition of \(\mathfrak{sp}(\infty)\). Then we can use this fact and the formula for the matrix entries of \(A^\#\) in Proposition 3.3 to prove the second part. □

**Definition 6.4.** Let \(HS\) be the space of Hilbert-Schmidt matrices viewed as a real vector space, and \(\mathfrak{sp}_{\text{HS}} = \mathfrak{sp}(\infty) \cap HS\).

The space \(HS\) as a real Hilbert space has an orthonormal basis
\[B_{\text{HS}} = \{e^{Re}_{mn} : m, n \neq 0\} \cup \{e^{Im}_{mn} : m, n \neq 0\},\]
where \(e^{Re}_{mn}\) is a matrix with \((m, n)\)-th entry 1 all other entries 0, and \(e^{Im}_{mn}\) is a matrix with \((m, n)\)-th entry 1 all other entries 0.

The space \(\mathfrak{sp}_{\text{HS}}\) is a closed subspace of \(HS\), and therefore a real Hilbert space. According to the symmetry of the matrices in \(\mathfrak{sp}_{\text{HS}}\), we define a projection \(\pi : HS \to \mathfrak{sp}_{\text{HS}}\), such that
\[
\begin{align*}
\pi(e^{Re}_{mn}) &= \frac{1}{2} (e^{Re}_{mn} - e^{Re}_{-m,-n} + e^{Re}_{-m,n} - e^{Re}_{m,-n}), & \text{if } \text{sgn}(mn) > 0 \\
\pi(e^{Im}_{mn}) &= \frac{1}{2} (e^{Im}_{mn} + e^{Im}_{-m,-n} - e^{Im}_{-m,n} - e^{Im}_{m,-n}), & \text{if } \text{sgn}(mn) > 0 \\
\pi(e^{Re}_{mn}) &= \frac{1}{2} (e^{Re}_{mn} + e^{Re}_{-m,-n} - e^{Re}_{-m,n} + e^{Re}_{m,-n}), & \text{if } \text{sgn}(mn) < 0 \\
\pi(e^{Im}_{mn}) &= \frac{1}{2} (e^{Im}_{mn} - e^{Im}_{-m,-n} + e^{Im}_{-m,n} - e^{Im}_{m,-n}), & \text{if } \text{sgn}(mn) < 0
\end{align*}
\]
Notation 6.5. We choose $B_{sp_{HS}} = \pi(B_{HS})$ to be the orthonormal basis of $sp_{HS}$.

Clearly, if $A \in sp_{HS}$, then $|A|_{sp_{HS}} = |A|_{HS}$.

Definition 6.6. Let $W_t$ be a Brownian motion on $sp_{HS}$ which has the mean zero and covariance $Q$, where $Q$ is assumed to be a positive symmetric trace class operator on $H$. We further assume that $Q$ is diagonal in the basis $B_{sp_{HS}}$.

Remark 6.7. $Q$ can also be viewed as a positive function on the set $B_{sp_{HS}}$, and the Brownian motion $W_t$ can be written as

$$W_t = \sum_{\xi \in B_{sp_{HS}}} \sqrt{Q(\xi)} B^\xi_t \xi,$$

where $\{B^\xi_t\}_{\xi \in B_{sp_{HS}}}$ are standard real-valued mutually independent Brownian motions.

Our goal now is to construct a Brownian motion on the group $Sp(\infty)$ using the Brownian motion $W_t$ on $sp_{HS}$. This is done by solving the Stratonovich stochastic differential equation

$$\delta X_t = X_t \delta W_t.$$

This equation can be written as the following Itô stochastic differential equation

$$dX_t = X_t dW_t + \frac{1}{2} X_t D dt,$$

where $D = \text{Diag}(D_m)$ is a diagonal matrix with entries

$$D_m = -\frac{1}{4} \text{sgn}(m) \sum_k \text{sgn}(k) \left[ Q^{Re}_{mk} + Q^{Im}_{mk} \right]$$

with $Q^{Re}_{mk} = Q(\pi(e^{Re}_{mk}))$ and $Q^{Im}_{mk} = Q(\pi(e^{Im}_{mk}))$.

Notation 6.8. Denote by $sp_{Q_{HS}} = Q^{1/2}(sp_{HS})$ which is a subspace of $sp_{HS}$. Define an inner product on $sp_{Q_{HS}}$ by $\langle u, v \rangle_{sp_{Q_{HS}}} = \langle Q^{-1/2}u, Q^{-1/2}v \rangle_{sp_{HS}}$. Then $B_{sp_{Q_{HS}}} = \{ \xi = Q^{1/2}\xi : \xi \in B_{sp_{HS}} \}$ is an orthonormal basis of the Hilbert space $sp_{Q_{HS}}$.

Notation 6.9. Let $L^2_2$ be the space of Hilbert-Schmidt operators from $sp_{Q_{HS}}$ to $sp_{HS}$ with the norm

$$||\Phi||_{L^2_2}^2 = \sum_{\xi \in B_{sp_{Q_{HS}}}} |\Phi_\xi|^2_{sp_{HS}} = \sum_{\xi, \zeta \in B_{sp_{HS}}} Q(\xi)(\Phi_\xi, \zeta)_{sp_{HS}}^2 = \text{Tr}[\Phi^* Q \Phi],$$

where $Q(\xi)$ means $Q$ evaluated at $\xi$ as a positive function on $B_{sp_{HS}}$.

Lemma 6.10. If $\Psi \in L(sp_{HS}, sp_{HS})$, a bounded linear operator from $sp_{HS}$ to $sp_{HS}$, then $\Psi$ restricted on $sp_{Q_{HS}}$ is a Hilbert-Schmidt operator from $sp_{Q_{HS}}$ to $sp_{Q_{HS}}$, and $||\Psi||_{L^2_2} \leq \text{Tr}(Q)||\Psi||^2$, where $||\Psi||$ is the operator norm of $\Psi$. 
has a unique solution, up to equivalence, among the processes satisfying

\[ \text{The stochastic differential equation} \]

Theorem 6.12.

Define \( B : \mathfrak{sp}_{\text{HS}} \rightarrow L_2^0 \) by \( B(Y)A = (I + Y)A \) for \( A \in \mathfrak{sp}_{\text{HS}}^0 \), and \( F : \mathfrak{sp}_{\text{HS}} \rightarrow \mathfrak{sp}_{\text{HS}} \) by \( F(Y) = \frac{1}{2}(I + Y)D \).

Note that \( B \) is well-defined by Lemma 6.10. Also \( D \in \mathfrak{sp}_{\text{HS}} \), and so \( F(Y) \in \mathfrak{sp}_{\text{HS}} \) and \( F \) is well-defined as well.

Theorem 6.12. The stochastic differential equation

\[ dY_t = B(Y_t)dW_t + F(Y_t)dt \]
\[ Y_0 = 0 \]

has a unique solution, up to equivalence, among the processes satisfying

\[ P \left( \int_0^T |Y_s|^2_{\mathfrak{sp}_{\text{HS}}} ds < \infty \right) = 1. \]

Proof. To prove this theorem we will use Theorem 7.4 from the book by G. DaPrato and J. Zabczyk [3] as it has been done in [6, 7]. It is enough to check

1. \( B \) is a measurable mapping.
2. \( |B(Y_1) - B(Y_2)|_{L_2^0} \leq C_1|Y_1 - Y_2|_{\mathfrak{sp}_{\text{HS}}} \) for \( Y_1, Y_2 \in \mathfrak{sp}_{\text{HS}} \);
3. \( |B(Y)|_{L_2^0} \leq K_1(1 + |Y|^2_{\mathfrak{sp}_{\text{HS}}}) \) for any \( Y \in \mathfrak{sp}_{\text{HS}} \);
4. \( F \) is a measurable mapping.
5. \( |F(Y_1) - F(Y_2)|_{\mathfrak{sp}_{\text{HS}}} \leq C_2|Y_1 - Y_2|_{\mathfrak{sp}_{\text{HS}}} \) for \( Y_1, Y_2 \in \mathfrak{sp}_{\text{HS}} \);
6. \( |F(Y)|_{\mathfrak{sp}_{\text{HS}}} \leq K_2(1 + |Y|^2_{\mathfrak{sp}_{\text{HS}}}) \) for any \( Y \in \mathfrak{sp}_{\text{HS}} \).

Proof of 1. By the proof of 2, \( B \) is a continuous mapping, therefore it is measurable.

Proof of 2.

\[ |B(Y_1) - B(Y_2)|_{L_2^0} = \sum_{\xi \in B_{r+p}^{s_{\text{HS}}}} |(Y_1 - Y_2)\xi|_{\mathfrak{sp}_{\text{HS}}}^2 = \sum_{\xi \in B_{r+p}^{s_{\text{HS}}}} Q(\xi)|(Y_1 - Y_2)\xi|_{\mathfrak{sp}_{\text{HS}}}^2 \]

\[ \leq \sum_{\xi \in B_{r+p}^{s_{\text{HS}}}} Q(\xi)||\xi||^2|Y_1 - Y_2|^2_{\mathfrak{sp}_{\text{HS}}} \leq \max_{\xi \in B_{r+p}^{s_{\text{HS}}}} ||\xi||^2 \left( \sum_{\xi \in B_{r+p}^{s_{\text{HS}}}} Q(\xi) \right)|Y_1 - Y_2|^2_{\mathfrak{sp}_{\text{HS}}} \]

\[ = \text{Tr} Q \left( \max_{\xi \in B_{r+p}^{s_{\text{HS}}}} ||\xi||^2 \right)|Y_1 - Y_2|^2_{\mathfrak{sp}_{\text{HS}}} = C_1^2|Y_1 - Y_2|^2_{\mathfrak{sp}_{\text{HS}}}, \]

where \( ||\xi|| \) is the operator norm of \( \xi \), which is uniformly bounded for all \( \xi \in B_{\mathfrak{sp}_{\text{HS}}} \).
Proof of 3.
\[
|B(Y)|^2 \leq \sum_{\xi \in B_{sp}HS} |(I + Y)\xi|_{sp}^2 = \sum_{\xi \in B_{sp}HS} Q(\xi)(I + Y)|\xi|_{sp}^2 \\
\leq |(I + Y)\xi|_{sp}^2 \sum_{\xi \in B_{sp}HS} Q(\xi)|\xi|^2 \leq (1 + |Y|_{sp}^2) \cdot K_1.
\]

Proof of 4. By the proof of 5, \( F \) is a continuous mapping, therefore it is measurable.

Proof of 5.
\[
|F(Y_1) - F(Y_2)|_{sp} = |\frac{1}{2}(Y_1 - Y_2)D|_{sp} \leq \frac{1}{2}D||Y_1 - Y_2||_{sp}.
\]

Proof of 6.
\[
|F(Y)|_{sp}^2 = \frac{1}{2}(I + Y)D_{sp}^2 \leq \frac{1}{2}D^2I + Y_{sp}^2 \leq K_2(1 + |Y|_{sp}^2).
\]

\[\square\]

**Notation 6.13.** Let \( B^\# : sp_{HS} \to L^2_0 \) be the operator \( B^\#(Y)A = A^\#(I + Y) \), and \( F^\# : sp_{HS} \to sp_{HS} \) be the operator \( F^\#(Y) = \frac{1}{2}D^\#(Y + I) \).

**Proposition 6.14.** If \( Y_t \) is the solution to the stochastic differential equation
\[
dX_t = B(X_t)dW_t + F(X_t)dt \\
X_0 = 0,
\]
where \( B \) and \( F \) are defined in Notation 6.11, then \( Y_t^\# \) is the solution to the stochastic differential equation
\[
dX_t = B^\#(X_t)dW_t + F^\#(X_t)dt \quad (6.6) \\
X_0 = 0,
\]
where \( B^\# \) and \( F^\# \) are defined in Notation 6.13.

**Proof.** This follows directly from the property \((AB)^\# = B^\#A^\#\) for any \( A \) and \( B \), which can be verified by using part (5) of Proposition 3.3. \[\square\]

**Lemma 6.15.** Let \( U \) and \( H \) be real Hilbert spaces. Let \( \Phi : U \to H \) be a bounded linear map. Let \( G : H \to H \) be a bounded linear map. Then
\[
Tr_H(G\Phi^{\ast}) = Tr_U(\Phi^{\ast}G\Phi)
\]

**Proof.**
\[
Tr_H(G\Phi^{\ast}) = \sum_{i,j \in U; k \in U} G_{ij}^{\ast}\Phi_{j,k}^{\ast} = \sum_{i,j \in H; k \in U} G_{ij}\Phi_{j,k}^{\ast}
\]
\[
Tr_U(\Phi^{\ast}G\Phi) = \sum_{i,j \in U; k \in U} (\Phi^{\ast})_{k,i}G_{ij}\Phi_{j,k} = \sum_{i,j \in U; k \in U} G_{ij}\Phi_{j,k}^{\ast}.
\]
Therefore \( Tr_H(G\Phi^{\ast}) = Tr_U(\Phi^{\ast}G\Phi). \) \[\square\]

**Lemma 6.16.**
\[
\sum_{\xi \in B_{sp}HS} (Q^{1/2}\xi)(Q^{1/2}\xi)^\# = -D
\]
Proof. If $\xi \in B_{\mathfrak{sp}_{HS}}$, then $\xi \in \mathfrak{sp}(\infty)$, so $\xi^\# = -\xi$. We will use the fact that
$$(e_{ij}^{Re}e_{kl}^{Re})_{pq} = \delta_p\delta_{jk}\delta_{i+l}$$
where $e_{ij}^{Re}$ is the matrix with the $(i,j)$th entry being 1 and all other entries being zero. Using this fact, we see

(1) for $\xi = \frac{1}{2}(e_{mn}^{Re} - e_{nm}^{Re} + e_{-m,-n}^{Re} - e_{n,m}^{Re})$ with $\text{sgn}(mn) > 0$,

$$(Q^{1/2}\xi)(Q^{1/2}\xi)^\# = -\frac{1}{4}Q_{mn}^{Re}[-e_{mn}^{Re} - e_{nm}^{Re} - e_{-m,-n}^{Re} - e_{n,m}^{Re}]$$

(2) for $\xi = \frac{1}{2}(e_{mn}^{Im} + e_{nm}^{Im} - e_{-m,-n}^{Im} - e_{n,m}^{Im})$ with $\text{sgn}(mn) > 0$,

$$(Q^{1/2}\xi)(Q^{1/2}\xi)^\# = -\frac{1}{4}Q_{mn}^{Im}[-e_{mn}^{Re} - e_{nm}^{Re} - e_{-m,-n}^{Re} - e_{n,m}^{Re}]$$

(3) for $\xi = \frac{1}{2}(e_{mn}^{Re} - e_{nm}^{Re} + e_{-m,-n}^{Re} + e_{n,m}^{Re})$ with $\text{sgn}(mn) < 0$,

$$(Q^{1/2}\xi)(Q^{1/2}\xi)^\# = -\frac{1}{4}Q_{mn}^{Re}[e_{mn}^{Re} + e_{nm}^{Re} + e_{-m,-n}^{Re} + e_{n,m}^{Re}]$$

(4) for $\xi = \frac{1}{2}(e_{mn}^{Im} + e_{nm}^{Im} - e_{-m,-n}^{Im} - e_{n,m}^{Im})$ with $\text{sgn}(mn) < 0$,

$$(Q^{1/2}\xi)(Q^{1/2}\xi)^\# = -\frac{1}{4}Q_{mn}^{Im}[e_{mn}^{Re} + e_{nm}^{Re} + e_{-m,-n}^{Re} + e_{n,m}^{Re}]$$

Each of the above is a diagonal matrix. The lemma can be proved by looking at the diagonal entries of the sum. \qed

**Theorem 6.17.** Let $Y_t$ be the solution to Equation 6.5. Then $Y_t + I \in \text{Sp}(\infty)$ for any $t > 0$ with probability 1.

**Proof.** The proof is adapted from papers by M. Gordina [6, 7]. Let $Y_t$ be the solution to Equation (6.5) and $Y^\#_t$ be the solution to Equation (6.6). Consider the process $Y_t = (Y_t, Y^\#_t)$ in the product space $\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$. It satisfies the following stochastic differential equation

$$dY_t = (B(Y_t), B^\#(Y^\#_t))dW + (F(Y_t), F^\#(Y^\#_t))dt.$$  

Let $G$ be a function on the Hilbert space $\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$ defined by $G(Y_1, Y_2) = \Lambda((Y_1 + I)(Y_2 + I))$, where $\Lambda$ is a nonzero linear real bounded functional from $\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$ to $\mathbb{R}$. We will apply Itô’s formula to $G(Y_t) = G(Y_t, Y^\#_t)$. Then $(Y_t + I)(Y^\#_t + I) = I$ if and only if $\Lambda((Y_1 + I)(Y^\#_t + I) - I) = 0$ for any $\Lambda$.

In order to use Itô’s formula we must verify that $G$ and the derivatives $G_t, G_Y, G_Y Y$ are uniformly continuous on bounded subsets of $[0, T] \times \mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$, where $G_Y$ is defined as follows

$$G_Y(Y)(S) = \lim_{\epsilon \to 0} \frac{G(Y + \epsilon S) - G(Y)}{\epsilon}$$

for any $Y, S \in \mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$ and $G_Y Y$ is defined as follows

$$G_{Y Y}(Y)(S \otimes T) = \lim_{\epsilon \to 0} \frac{G_Y(Y + \epsilon T)(S) - G_Y(Y)(S)}{\epsilon}$$

for any $Y, S, T \in \mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$. Let us calculate $G_t, G_Y, G_{Y Y}$. Clearly, $G_t = 0$. It is easy to verify that for any $S = (S_1, S_2) \in \mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$

$$G_Y(Y)(S) = \Lambda(S_1(Y_2 + I) + (Y_1 + I)S_2)$$
and for any $S = (S_1, S_2) \in \mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$ and $T = (T_1, T_2) \in \mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$

$$G_{YY}(Y)(S \otimes T) = \Lambda(S_1 T_2 + T_1 S_2).$$

So the condition is satisfied.

We will use the following notation

$$G_Y(Y)(S) = \langle \hat{G}_Y(Y), S \rangle_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}}$$

$$G_{YY}(Y)(S \otimes T) = \langle \hat{G}_{YY}(Y), S, T \rangle_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}},$$

where $\hat{G}_Y(Y)$ is an element of $\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$ corresponding to the functional $G_Y(Y)$ in $(\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS})^*$ and $\hat{G}_{YY}(Y)$ is an operator on $\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}$ corresponding to the functional $G_{YY}(Y) \in ((\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}) \otimes (\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}))^*$.

Now we can apply Itô’s formula to $G(Y_t)$

$$G(Y_t) - G(Y_0) = \int_0^t \langle \hat{G}_Y(Y_s), (B(Y_s)dW_s, B^#(Y^#_s)dW_s) \rangle_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}} ds$$

$$+ \int_0^t \langle \hat{G}_Y(Y_s), (F(Y_s), F^#(Y^#_s)) \rangle_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}} ds$$

$$+ \int_0^t \frac{1}{2} \text{Tr}_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}} \left[ \hat{G}_{YY}(Y_s) \left( B(Y_s)Q^{1/2}, B^#(Y^#_s)Q^{1/2} \right) \right] ds.$$

Let us calculate the three integrands separately. The first integrand is

$$\langle \hat{G}_Y(Y_s), (B(Y_s)dW_s, B^#(Y^#_s)dW_s) \rangle_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}}$$

$$= \left( B(Y_s)dW_s \right) (Y^#_s + I) + (Y_s + I) \left( B^#(Y^#_s)dW_s \right)$$

$$= (Y_s + I)dW_s(Y^#_s + I) + (Y_s + I)dW^#_s(Y^#_s + I) = 0.$$

The second integrand is

$$\langle \hat{G}_Y(Y_s), (F(Y_s), F^#(Y^#_s)) \rangle_{\mathfrak{sp}_{HS} \times \mathfrak{sp}_{HS}}$$

$$= F(Y_s)(Y^#_s + I) + (Y_s + I)F^#(Y^#_s)$$

$$= \frac{1}{2}(Y_s + I)D(Y^#_s + I) + \frac{1}{2}(Y_s + I)D^#(Y^#_s + I)$$

$$= \frac{1}{2}(Y_s + I)(D + D^#)(Y^#_s + I)$$

$$= (Y_s + I)D(Y^#_s + I),$$

where we have used the fact that $D = D^#$, since $D$ is a diagonal matrix with all real entries.
The third integrand is

\[
\frac{1}{2} \text{Tr}_{\mathfrak{sp} \times \mathfrak{sp}} [G_{YY}(Y_s) \left( B(Y_s)Q^{1/2}, B^#(Y^#_s)Q^{1/2} \right)^* \left( B(Y_s)Q^{1/2}, B^#(Y^#_s)Q^{1/2} \right)]
\]

\[
= \frac{1}{2} \text{Tr}_{\mathfrak{sp} \times \mathfrak{sp}} \left[ \left( B(Y_s)Q^{1/2}, B^#(Y^#_s)Q^{1/2} \right)^* G_{YY}(Y_s) \left( B(Y_s)Q^{1/2}, B^#(Y^#_s)Q^{1/2} \right) \right]
\]

\[
= \frac{1}{2} \sum_{\xi \in B_{\mathfrak{sp}}} G_{YY}(Y_s) \left( \left( B(Y_s)Q^{1/2}\xi, B^#(Y^#_s)Q^{1/2}\xi \right) \otimes \left( B(Y_s)Q^{1/2}\xi, B^#(Y^#_s)Q^{1/2}\xi \right) \right)
\]

\[
= \sum_{\xi \in B_{\mathfrak{sp}}} \left( B(Y_s)Q^{1/2}\xi \right) \left( B^#(Y^#_s)Q^{1/2}\xi \right)
\]

\[
= \sum_{\xi \in B_{\mathfrak{sp}}} \left( Y_s + I \right) \left( Q^{1/2}\xi \right) \left( Q^{1/2}\xi \right)^# \left( Y^#_s + I \right)
\]

\[
= -(Y_s + I)D(Y^#_s + I),
\]

where the second equality follows from Lemma 6.15, and the last equality follows from Lemma 6.16.

The above calculations show that the stochastic differential of $G$ is zero. So $G(Y_t) = G(Y_0) = \Lambda(I)$ for any $t > 0$ and any nonzero linear real bounded functional $\Lambda$ on $\mathfrak{sp} \times \mathfrak{sp}$. This means $(Y_t + I)(Y^#_t + I) = I$ almost surely for any $t > 0$. Similarly we can show $(Y^#_t + I)(Y_t + I) = I$ almost surely for any $t > 0$. Therefore $Y_t + I \in \text{Sp}(\infty)$ almost surely for any $t > 0$. \qed

References

MARIA GORDINA: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CONNECTICUT, STORRS, CT 06269, U.S.A.
E-mail address: gordina@math.uconn.edu

MANG WU: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CONNECTICUT, STORRS, CT 06269, U.S.A.
E-mail address: mvu@math.uconn.edu
ANALYSIS OF COMPLEX BROWNIAN MOTION

YUH-JIA LEE* AND KUANG-GHIEH YEN

Abstract. A theory of generalized functions based on the complex Brownian motion \( \{ Z(t) : t \in \mathbb{R} \} \), for which each \( Z(t) \) is \( N(0, |t|) \), is established on the probability space \( (S_c', B(S_c'), \nu(dz)) \), where \( S_c' \) is the dual of the Schwartz space \( S \), \( S_c' \) is the complexification of \( S' \), identified as the product space \( S' \times S' \), \( B(S_c') \) the Borel field of \( S' \times S' \) and \( \nu(dz) \) denotes the product measure \( \mu_1(dx) \mu_1(dy) \).

Using the representation of the complex Brownian motion

\[
Z_t(x, y) = \frac{1}{\sqrt{2}} \left( (x, h_t) + i(y, h_t) \right),
\]

where

\[
h_t = \begin{cases} 
1\{0,t\}, & t > 0, \\
-1\{t,0\}, & t < 0,
\end{cases}
\]

and employing the technique of white noise calculus initiated by Hida (see, e.g. [2] and [4]), we analyze functionals of complex Brownian motion. To define generalized complex Brownian functionals, we adopt the space of CKS entire functionals as test functions. As applications, the stochastic integral with respect to a complex Brownian motion are defined and studied. The Itô formula for complex Brownian functionals is obtained and it is shown that the evaluation of stochastic integral with respect to a complex Brownian motion follows the rule of Stratonovich integral.

\[\text{1. Introduction}\]

In this paper we are devoted to a systematic study of the complex Brownian functionals, by which we mean functions of complex Brownian motion given by

\[
Z(t, \omega) = \frac{1}{\sqrt{2}} [B_1(t, \omega) + iB_2(t, \omega)],
\]

where \( B_1 \) and \( B_2 \) are independent real-valued standard Brownian motions. Clearly \( Z(t) \) is normally distributed with mean zero and variance parameter \( |t| \).

As the calculus of the complex Brownian motion will play the main role, we need to represent \( Z(t) \) as a function on a certain probability space. In this paper, we choose \( (S_c', B(S_c'), \nu(dz)) \) as the underlying probability space, where \( S \) is the Schwartz space with dual space \( S' \), \( S_c' \) is the complexification of \( S' \) which is identified as the product space \( S' \times S' \), \( B(S_c') \) the Borel field of \( S' \times S' \) and \( \nu(dz) \) denotes

* This research is supported by the National Science Council of Taiwan.
the product measure \( \mu_1(dx)\mu_2(dy) \), where \( \mu_t \) represents the Gaussian measure defined on \( S' \) with characteristic function given by

\[
C(\xi) = \int_{S'} e^{(x,\xi)} \mu_t(dx) = e^{-|\xi|^2/2}.
\]

One sees easily that the complex Brownian motion on \((S'_c, B(S'_c), \nu(dz))\) may be represented by

\[
Z_t(x, y) = \frac{1}{\sqrt{2}} \left( (x, h_t) + i(y, h_t) \right),
\]

where

\[
h_t = \begin{cases} 1_{[0,t]} & , t > 0, \\ -1_{[t,0]} & , t < 0. \end{cases}
\]

The calculus of complex Brownian functionals is then performed with respect to the measure \( \mu(dz) \). For example, let \( f : \mathbb{C} \to \mathbb{C} \) be an entire function of exponential growth and \( Z(t) = \frac{1}{\sqrt{2}} (x + iy, h_t) \) the complex Brownian motion. Then we have

\[
E[|f(Z(t))|^2] = \int_{S'_c} \int_{S'_c} |f((x + iy, h_t))|^2 \mu_{1/2}(dx)\mu_{1/2}(dy).
\]

The above identity gives a connection between the calculus of complex Brownian motion and the Segal–Bargmann entire functionals (see §3 for definition).

Being motivated by the (real) white noise analysis initiated by Hida (see, e.g. [4] and [2]), it is desirable to develop a theory of generalized complex functionals by white noise calculus approach. The main results are given as follows:

2. A Stochastic integral with respect to complex Brownian motion is defined and studied (for related work, we referred the reader to [3]).
3. An Itô formula for complex Brownian functionals is derived.

As an example, it is shown that, for any Segal–Bargmann entire function \( F \), the Itô formula is given by

\[
F(Z(b)) - F(Z(a)) = \int_a^b F'(Z(t))dZ(t).
\]

The formula is then extended to a generalized CKS entire function \( F \) by using Hitsuda–Skorokhod integral given below:

\[
\frac{d}{dt} \langle F(Z(t)), \phi \rangle_c = \langle \partial_t^* F'(Z(t)), \phi \rangle_c,
\]

where \( \langle \cdot, \cdot \rangle_c \) denotes the pairing of generalized functional and test functional, \( \partial_t = D_{\xi} \) and \( \partial_t^* \) is its adjoint, and \( \phi \) is any CKS entire functional.

2. White Noise Calculus

The (Gaussian) white noise is generally understood in the engineering literature as a stationary stochastic process with constant spectral density. In mathematics, it can be shown that the white noise is the time derivative \( \{ \dot{B}(t) : t > 0 \} \) of a Brownian motion. Since almost all sample paths of a Brownian motion are differentiable nowhere, the (Gaussian) white noise is realized in mathematics as a
generalized process. In this paper, the Brownian motion denoted by \( \{ B(t) : t > 0 \} \) will be regarded as a regular generalized functional on \( S' \). Hence for almost all \( x (\mu_t) \) the white noise \( t \mapsto \dot{B}_t(x) \) is a generalized function in \( S' \). Thus \( S' \) is regarded as the state space of white noise.

\( S' \) also has a nuclear space structure described as follows:

Let \( A \) denote the operator \( A = 1 + t^2 - (d/dt)^2 \) with domain \( \mathcal{D}(A) \subset L^2 = L^2(\mathbb{R}) \) and \( \mathcal{D}(A) \) contains a CONS \( \{ e_n : n \in \mathbb{N} \} \) of \( L^2 \) consisting of eigenfunctions of \( A \) with corresponding eigenvalues \( \{ 2n + 2 : n = 1, 2, \ldots \} \). \( \{ e_n \} \) are known as Hermite functions.

For \( p \geq 0 \), let \( S_p = D(A^p) \). Its dual space is given by \( S_p^* = S_{-p} \). Then we have \( S = \cap_p S_p \) which is equipped with the projective limit topology and the dual space of \( S(\mathbb{R}^1) \) is given by \( S' = \cup_p S_p \) which is equipped with the inductive topology. Moreover, the spaces \( S \subset L^2(\mathbb{R}^1) \subset S' \) form a Gel'fand triple.

It is well-known that \( S' \) carries a standard Gaussian measure \( \mu_t \) with variance parameter \( t > 0 \) (for details, we refer the reader to [5]) and \( \mu_t \)'s run through all finite dimensional orthogonal projections on \( H \).

Denote the class of Segal–Bargmann entire functions on \( H \) by \( \mathcal{SB}_t[H] \) and define \( \| f \|_{\mathcal{SB}_t[H]} = \sqrt{M_f} \). Then \( (\mathcal{SB}_t[H], \| \cdot \|_{\mathcal{SB}_t[H]}) \) is a Hilbert space.

It follows immediately from [6] that we have

\[
\| f \|_{\mathcal{SB}_t[H]}^2 = \sum_{k=0}^{\infty} \frac{(2t)^k}{k!} \left( \sum_{i_1, \ldots, i_k = 1}^{N} |D^k f(0)e_{i_1} \cdots e_{i_k}|^2 \right) = \sum_{k=0}^{\infty} \frac{(2t)^k}{k!} \| D^k f(0) \|_{\mathcal{HS}^2[H]}^2,
\]

(3.1)

3. Test and Generalized Functions

Definition 3.1. (Segal–Bargmann space[6]) Let \( H \) be a real separable Hilbert space and \( H_c \) the complexification of \( H \). A single-valued function \( f \) defined on \( H_c \) is called a Segal–Bargmann entire function if it satisfies the following conditions:

(i) \( f \) is analytic in \( H_c \).

(ii) The number

\[
M_f := \sup_{P} \int_H \int_H |f(Px + iPy)|^2 n_t(dx)n_t(dy)
\]

is finite, where \( n_t \) denotes as the Gaussian cylinder measure on \( H \) with variance parameter \( t > 0 \) (for details, we refer the reader to [5]) and \( P's \) run through all finite dimensional orthogonal projections on \( H \).

Denote the class of Segal–Bargmann entire functions on \( H \) by \( \mathcal{SB}_t[H] \) and define \( \| f \|_{\mathcal{SB}_t[H]} = \sqrt{M_f} \). Then \( (\mathcal{SB}_t[H], \| \cdot \|_{\mathcal{SB}_t[H]}) \) is a Hilbert space.
where $\|S\|_{H^S(H)}$ denotes the Hilbert–Schmidt norm of an $n$-linear operator $S \in L^n(H)$ defined by

$$\|S\|_{H^S(H)} := \left( \sum_{i_1, \ldots, i_k = 1}^{\infty} |S e_{i_1} \cdots e_{i_k}|^2 \right)^{1/2}$$

which is independent of the choice of CONS $\{e_i\}$ of $H$.

Next, we introduce the CKS entire functionals [1].

Let $\alpha(n), n \geq 0$, be a sequence of real numbers satisfying the following conditions:

(a) $\alpha(0) = 1$ and $\inf_{n \geq 0} \alpha(n) > 0$,
(b) $\lim_{n \to \infty} n^{-1} \alpha(n)^{1/n} = 0$,
(c) $\gamma_{n+2}/\gamma_{n+1} \leq \gamma_{n+1}/\gamma_n$, for all $n \geq 0$, where $\gamma_n = \alpha(n)/n!$.

Definition 3.2. (Infinite-dimensional CKS entire functionals) For each $p \in \mathbb{R}$, define

$$\|\phi\|_{\alpha,p} = \left( \sum_{n=0}^{\infty} \alpha(n) \frac{\|D^n \phi(0)\|_{H^S[S]}^2}{n!} \right)^{1/2}$$

and set

$$SB_{p,\alpha} = \{ \phi \in SB_{1/2}[H] : \|\phi\|_{\alpha,p} < \infty \}.$$

Let $SB_\alpha$ be the projective limit of $SB_{p,\alpha}$ for $p \geq 0$ and let $SB_\alpha'$ be the dual space of $SB_\alpha$. Then $SB_\alpha$ is a nuclear space and we have the following continuous inclusions:

$$SB_\alpha \subset SB_{p,\alpha} \subset SB[L^2] \subset SB'_{p,\alpha} \subset SB'_\alpha,$$


The space $SB_\alpha$, which is referred to as the CKS entire functionals on $S'_{c}$, will serve as test functionals and $SB'_\alpha$ is referred to as the generalized complex Brownian functionals.

The space $SB'_{p,\alpha}$ may be identified as the space of entire functions defined on $S_{p,c}$ such that $\|\phi\|_{\alpha-1,-p} < \infty$ and the pairing of $SB'_{\alpha}$ and $SB_\alpha$ is defined by

$$\langle \langle \Phi, \varphi \rangle \rangle = \sum_{n=0}^{\infty} \frac{1}{n!} \langle \langle D^n \Phi(0), D^n \varphi(0) \rangle \rangle_{H^S},$$

where

$$\langle \langle D^n \Phi(0), D^n \varphi(0) \rangle \rangle_{H^S} := \sum_{i_1, \ldots, i_n = 1}^{\infty} \left[ D^n \Phi(0) e_{i_1} \cdots e_{i_n} D^n \varphi(0) e_{i_1} \cdots e_{i_n} \right].$$

Definition 3.3. (One-dimensional CKS entire functions) If $\phi(z)$ can be represented by a formal power series $\sum_{n=0}^{\infty} a_n z^n$, we define

$$\|\phi\|_\alpha = \left( \sum_{n=0}^{\infty} \alpha(n)n! |a_n|^2 \right)^{1/2}$$

and let

$$SB_\alpha(\mathbb{R}) = \{ \phi : \|\phi\|_{\alpha,p} < \infty \}.$$
If $\phi(z)$ is a formal power series represented by $\sum_{n=0}^{\infty} b_n z^n$, we define

$$\|\phi\|_{\alpha-1} = \left( \sum_{n=0}^{\infty} \frac{n!|b_n|^2}{\alpha(n)} \right)^{1/2}.$$ 

Then the dual space of $SB_\alpha$ is characterized by

$$SB_\alpha'(\mathbb{R}) = \{ \phi : \|\phi\|_{\alpha-1} < \infty \}.$$ 

**Remark 3.4.** Let $k$ be any positive integer and let $f(w) = w^k$ ($w \in \mathbb{C}$). Then, for any $h \in H$, the functional $\Phi = f(\langle \cdot, h \rangle)$ is clearly a Segal–Bargmann entire function defined on $H$. $\Phi$ is also defined on $S$ c.a.e. ($\nu$) and we have, for $\phi \in SB_\alpha$,

$$\langle \langle \Phi, \phi \rangle \rangle_c = \int_S \int_S f(\langle x + iy, h \rangle)\phi(x + iy)\mu_1(dx)\mu_2(dy).$$

The above identity will be used in the proof of Theorem 5.7.

### 4. Examples of Complex Brownian Functionals

The following lemma will be used frequently in computation.

**Lemma 4.1.** For the functional representation of a complex Brownian motion $Z(t)$ given by

$$Z(t, x, y) = \frac{\langle x + iy, h_t \rangle}{\sqrt{2}},$$

we have

(i) $$\langle \langle Z(t), 1 \rangle \rangle_c = \int_S \int_S (x + iy, h_t)\mu_{1/2}(dx)\mu_{1/2}(dy) = 0.$$ 

(ii) $$\langle \langle Z(t)^n, Z(t)^n \rangle \rangle_c = n!\delta_{m,n}t^n.$$  

**Proof.** (i) is clear and the proof of (ii) follows immediately from integration by parts formula (see, for example [6]). $\square$

**Example 4.2.** If $\xi \in S$, then clearly $(\cdot, \xi) \in SB_\alpha$. For any $\varphi \in SB_\alpha$, it follows immediately from the definition of the $\langle \langle \cdot, \cdot \rangle \rangle_c$ pairing that we have

$$\langle \langle (\cdot, \xi), \varphi \rangle \rangle_c = D_{\xi}\varphi(0) = D\varphi(0)\xi,$$

where $D\varphi(0)$ denotes the Fréchet derivative of $\varphi$ at 0. Note that the right hand side of the the identity in Equation (4.1) remains meaningful even for $\xi \in S'$. $(\cdot, \xi)$ defined a continuous linear functional on $SB_\alpha$ so that $(\cdot, \xi) \in SB_\alpha'$. It is worthwhile to note that, for $\varphi \in SB[L^2, L^2[\mathbb{R}]$—differentiable i.e. $\varphi$ is Fréchet differentiable in the directions of $L^2[\mathbb{R}]$. If, for $\xi \in L^2[\mathbb{R}]$, we interpret $D_{\xi}\varphi(0)$ as the $L^2[\mathbb{R}]$—derivative of $\varphi$ at 0 in the direction of $\xi$, then the identity in Equation (4.1) remains valid. It follows that one can apply this identity with $\xi = (h_{t+\epsilon} - h_t)/\epsilon$ and $\varphi \in SB_\alpha$ to get

$$\langle \langle (\cdot, (h_{t+\epsilon} - h_t)/\epsilon), \varphi \rangle \rangle_c = D\varphi(0)[(h_{t+\epsilon} - h_t)/\epsilon].$$
Let $\epsilon \to 0$ in the above identity. We obtain the definition of the complex white noise $\dot{Z}(t)$ as follows:

$$\langle\langle \dot{Z}(t), \varphi \rangle \rangle_c = D\varphi(0)\delta_t.$$ 

Clearly $\dot{Z}(t) \in \mathcal{SB}_\alpha'$.

**Example 4.3.** Given $h \in L^2[\mathbb{R}]$, we have

$$\langle\langle e^{\langle \cdot, h \rangle}, \varphi \rangle \rangle_c = \sum_{n=0}^{\infty} \frac{1}{n!} \langle\langle h \otimes_n, D^n\varphi(0) \rangle \rangle = \phi(h).$$

If $h = h_t$ then we obtain

$$\langle\langle e^{Z(t)}, \varphi \rangle \rangle_c = \phi(h_t). \quad (4.2)$$

It is easy to see that $e^{Z(t)} \in \mathcal{SB}_\alpha'$.

**Example 4.4.** Let $h, k \in L^2[\mathbb{R}]$ such that $\langle h, k \rangle = 0$. Then it follows from Equations (4.1) and (4.2) that

$$\langle\langle \langle \cdot, h \rangle e^{\langle \cdot, k \rangle}, \varphi \rangle \rangle_c = \langle h, k \rangle \langle\langle 1, \varphi \rangle \rangle_c + \langle\langle e^{\langle \cdot, k \rangle}, \varphi \rangle \rangle_c = \langle\langle e^{\langle \cdot, k \rangle}, \varphi \rangle \rangle_c. \quad (4.3)$$

Taking $h = (h_t + \epsilon - h_t)/\epsilon$ and $k = h_t$ in Equation (4.3) and then letting $\epsilon \to 0$, we are led to the definition of $\dot{Z}(t) \exp(Z(t))$ given in the following

**Definition 4.5.** For $\varphi \in \mathcal{SB}_\alpha$, we define

$$\langle\langle \dot{Z}(t)e^{Z(t)}, \varphi \rangle \rangle_c = D\varphi(h_t)\delta_t. \quad (4.4)$$

Clearly $\dot{Z}(t)e^{Z(t)} \in \mathcal{SB}_\alpha'$.

**Example 4.6.** If $Z(t)$ is a complex Brownian motion, then it follows from Example 4.2 and Definition 4.5 that we have

$$\frac{d}{dt}e^{Z(t)} = e^{Z(t)} \dot{Z}(t).$$

**Example 4.7.** (Composition of generalized function with a complex Brownian motion) Let $f \in \mathcal{SB}_\alpha'(\mathbb{R})$ be an one dimensional generalized CKS entire function represented by $f(z) = \sum_{n=0}^{\infty} b_n z^n$. Assume that $\sum_{n=0}^{\infty} \alpha(n)^{-1} n! |b_n|^2 < \infty$. Then, for $\psi \in \mathcal{SB}_\alpha$, we define

$$\langle\langle f(Z(t)), \psi \rangle \rangle_c = \sum_{n=0}^{\infty} b_n D^n\psi(0)h^n_t. \quad (4.5)$$

It will be proved in the next section that $f(Z(t)) \in \mathcal{SB}_\alpha'$. 

5. Itô Formula

We first prove that the composition $f(Z(t))$ defined in Example 4.7 is in fact a generalized Segal–Bargmann functional.

**Theorem 5.1.** Let $f \in SB'_\alpha(\mathbb{R})$ be an one dimensional generalized CKS entire function represented by $f(z) = \sum_{n=0}^\infty b_n z^n$. Assume that that $\{b_n\}$ satisfies $\sum_{n=0}^\infty n! |b_n|^{-1} n! |b_n|^2 < \infty$. Then $f(Z(t))$, defined by Equation (4.5), is a member of $SB'_\alpha$.

**Proof.** Recall that $\langle\langle f(Z(t)), \psi \rangle\rangle_c = \sum_{n=0}^\infty b_n D^n \psi(0) h_t^n$.

For each $t > 0$, there exists $p$ such that $|h_t|_{-p} \leq 1$. Thus

$$|\langle\langle f(Z(t)), \psi \rangle\rangle_c| = \sum_{n=0}^\infty \left( b_n \left( \frac{n!}{\alpha(n)} \right)^{1/2} \right) \left( \frac{D^n \psi(0) h_t^n}{\sqrt{n!}} \right) (\alpha(n))^{1/2}$$

$$\leq \|f\|_{\alpha^{-1}} \sqrt{\sum_{n=0}^\infty \frac{|D^n \psi(0)|^2}{n!} h_t^n \alpha(n)} $$

$$\leq \|f\|_{\alpha^{-1}} \sqrt{\sum_{n=0}^\infty \frac{\|D^n \psi(0)\|_{HS[S, S^{-p}]}^2}{n!} |h_t|^{2n-p} \alpha(n)}$$

$$\leq \|f\|_{\alpha^{-1}} \|\psi\|_{\alpha, \alpha'}.$$ 

This proves that $f(Z(t)) \in SB'_\alpha$. $\square$

Now we are ready to derive the Itô formula. Let $f \in SB'_\alpha(\mathbb{R})$. Then we have

$$\frac{d}{dt} \langle\langle f(Z(t)), \phi \rangle\rangle_c = \sum_{n=0}^\infty b_n D^n \phi(0) h_t^{n-1} \delta_t$$

$$= \sum_{n=0}^\infty b_n n D^{n-1} (D\phi(0)) h_t^{n-1}$$

$$= \sum_{n=0}^\infty b_n n D^{n-1} (\partial_t \phi(0)) h_t^{n-1}$$

$$= \langle\langle \partial_t^* f'(Z(t)), \phi \rangle\rangle_c,$$

where $\partial_t = \partial_{h_t}$ and $\partial_t^*$ is the adjoint operator of $\partial_t$. It follows that

$$\frac{d}{dt} f(Z(t)) = \partial_t^* f'(Z(t)),$$

This proves the Itô formula for a complex Brownian motion. As a summary, we state the above result as a theorem.
Theorem 5.2. (Itô formula) Let \( f \in \mathcal{SB}_r'(\mathbb{R}) \). Then we have
\[
\frac{d}{dt} f(Z(t)) = \partial_t^* f'(Z(t)),
\]
or in the integral form,
\[
f(Z(b)) - f(Z(a)) = \int_a^b \partial_t^* f'(Z(t)) dt.
\]

As in the case of a real Brownian motion, the term on the right hand side may be interpreted as a stochastic integral as shown below.

**Definition 5.3.** Suppose that \( f \in \mathcal{SB}_r' \). Define the stochastic integral \( f(Z(t)) \) as follows:
\[
\left\langle \left\langle \int_a^b f(Z(t)) dZ(t), \phi \right\rangle \right\rangle_c := \lim_{\|\Delta_n\| \to 0} \left\langle \left\langle \sum_{i=1}^n f(Z(t_{i-1}))(Z(t_i) - Z(t_{i-1})), \phi \right\rangle \right\rangle_c,
\]

where \( a = t_0 < t_1 < t_2 < \cdots < t_n = b \) and \( \|\Delta_n\| = \max_j |t_j - t_{j-1}|. \)

**Lemma 5.4.** Let \( \tilde{h} = \langle x + iy, h \rangle \) and \( \tilde{k} = \langle x + iy, k \rangle \), then
\[
\left\langle \left\langle \tilde{h}, \phi \right\rangle \right\rangle_c = \langle D^2 \phi(0), h \otimes k \rangle.
\]

**Proof.** The proof is straightforward. \( \square \)

**Corollary 5.5.** Let \( \tilde{h}_t = \langle x + iy, h_t \rangle \), then \( \left\langle \left\langle \tilde{h}_t^n, \phi \right\rangle \right\rangle_c = \langle D^n \phi(0), h_t^{\otimes n} \rangle. \)

**Example 5.6.** To compare the stochastic integrals for complex Brownian motions with the stochastic integral for real Brownian motions, we first show that
\[
\int_a^b Z(t) dZ(t) = \frac{1}{2} (Z^2(b) - Z^2(a)).
\]
In fact, it follows from Definition 5.3 that, for any test functional \( \varphi \), we have
\[
\left\langle \left\langle \int_a^b Z(t) dZ(t), \varphi \right\rangle \right\rangle_c
= \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n \left\langle \left\langle Z(t_{i-1})(Z(t_i) - Z(t_{i-1})), \varphi \right\rangle \right\rangle_c
= \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n D^2 \varphi(0) h_{t_{i-1}} h_{t_i - t_{i-1}}
= \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n \frac{1}{2} \left\{ D^2 \varphi(0)(h_{t_i}^{\otimes 2} - h_{t_{i-1}}^{\otimes 2}) + \left[ D^2 \varphi(0)(h_{t_i} - h_{t_{i-1}}) \otimes Z^2 \right] \right\}
= \lim_{\|\Delta_n\| \to 0} \left\{ (I) + (II) \right\}.
\]

It is easy to see that
\[
(I) = \left\langle \left\langle \frac{1}{2} (Z^2(b) - Z^2(a)), \phi \right\rangle \right\rangle_c
\]
and \( (II) \to 0 \) as \( \| \Delta_n \| \to 0 \). Therefore we have
\[
\int_a^b Z(t) dZ(t) = \frac{1}{2} (Z^2(b) - Z^2(a)). \tag{5.1}
\]
On the other hand,
\[
\langle\int_a^b \partial_t^* Z(t)dt, \phi\rangle_c = \int_a^b \langle\partial_t^* Z(t), \phi\rangle_c dt
\]
\[
= \int_a^b (D^2 \phi(0), h_t \otimes \delta_t) dt
\]
\[
= \frac{1}{2} \int_a^b \frac{d}{dt} (D^2 \phi(x + iy), h_t \otimes h_t) dt
\]
\[
= \frac{1}{2} \langle\langle [Z^2(b) - Z^2(a)], \phi\rangle\rangle_c
\]  
(5.2)

It follows from Equations (5.1) and (5.2) that
\[
\int_a^b Z(t) dZ(t) = \int_a^b \partial_t^* Z(t) dt.
\]

**Theorem 5.7.** Let \( f \in SB_\alpha(\mathbb{R}) \) and \( \phi \in SB_\alpha \) then
\[
\langle\int_a^b f(Z(t))dZ(t), \phi\rangle_c = \langle\int_a^b \partial_t^* f(Z(t)) dt, \phi\rangle_c.
\]

**Proof.** Since \( f \) can be represented by a formal power series, it is sufficient to prove the theorem for \( f(z) = z^k \) for arbitrary non-negative integer \( k \). By definition,
\[
\langle\int_a^b f(Z(t))dZ(t), \varphi\rangle_c = \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n \langle f(Z(t_{i-1}))(Z(t_i) - Z(t_{i-1})), \varphi\rangle_c
\]
for any \( \varphi \in SB_\alpha \). According to Remark 3.4, for each \( i \), we have
\[
\langle f(Z(t_{i-1}))(Z(t_i) - Z(t_{i-1})), \varphi\rangle_c
\]
\[
= \int_{S^2} f((x - iy, h_{t_i}))(x - iy, h_{t_i} - h_{t_{i-1}}) \varphi(x + iy) \mu_2(dx) \mu_2(dy)
\]
\[
= \int_{S^2} f((x - iy, Ph_{t_i}))(x - iy, P[h_{t_i} - h_{t_{i-1}}]) \varphi(x + iy) \mu_2(dx) \mu_2(dy).
\]  
(5.3)

Next, apply integration by parts formula, Equation (5.3) becomes
\[
\int_{S^2} \int_{S^2} f((x - iy, h_{t_{i-1}}))(x - iy, h_{t_i} - h_{t_{i-1}}) \varphi(x + iy) \mu_2(dx) \mu_2(dy)
\]
\[
= \int_{S^2} \int_{S^2} f((x - iy, h_{t_{i-1}}))(x, h_{t_i} - h_{t_{i-1}}) \varphi(x + iy) \mu_2(dx) \mu_2(dy)
\]
\[
- i \int_{S^2} \int_{S^2} f((x - iy, h_{t_{i-1}}))(y, h_{t_i} - h_{t_{i-1}}) \varphi(x + iy) \mu_2(dx) \mu_{1/2}(dy)
\]
\[
= \frac{1}{2} \int_{S^2} \int_{S^2} (f((x - iy, h_{t_i}))(D\varphi(x + iy), h_{t_i} - h_{t_{i-1}}) \mu_2(dx) \mu_2(dy)
\]
\[
- i (f((x - iy, h_{t_{i-1}}))(D\varphi(x + iy), P[h_{t_i} - h_{t_{i-1}}]) \mu_2(dx) \mu_{1/2}(dy)
\]
\[
= \int_{S^2} \int_{S^2} f((x - iy, h_{t_{i-1}}))(D\varphi(x + iy), h_{t_i} - h_{t_{i-1}}) \mu_2(dx) \mu_2(dy).
\]
Then we have
\[
\langle f(Z(t_{i-1}))(Z(t_i) - Z(t_{i-1})), \varphi \rangle_c
= \int_{S'} \int_{S'} f((x - iy, h_{t_{i-1}}))(D\varphi(x + iy), h_{t_i} - h_{t_{i-1}}) \mu_{\frac{1}{2}}(dx) \mu_{\frac{1}{2}}(dy)
= \int_{S'} \int_{S'} f((x - iy, h_{t_{i-1}}))(D\varphi(x + iy)) \frac{h_{t_i} - h_{t_{i-1}}}{t_i - t_{i-1}}(t_i - t_{i-1}). \quad (5.4)
\]
Finally, by Equation (5.4), we have
\[
\sum_{i=1}^{n} \langle f(Z(t_{i-1}))(Z(t_i) - Z(t_{i-1})), \varphi \rangle_c
= \sum_{i=1}^{n} (f(Z(t_{i-1}))(D\varphi(x + iy), \frac{h_{t_i} - h_{t_{i-1}}}{t_i - t_{i-1}}(t_i - t_{i-1})
\rightarrow \langle \int_{a}^{b} \partial_1^* f(Z(t))dt, \phi \rangle_c \quad \text{as} \quad \| \Delta_n \| \rightarrow 0.
\]
The last step follows from the fact that \( \frac{h_{t_i} - h_{t_{i-1}}}{t_i - t_{i-1}} \) converges to \( \delta_t \), uniformly with respect to the norm of \( S_p \) for \( p > \frac{1}{2} \). We complete the proof. \( \square \)

Remark 5.8. The reason why the second derivative term, from the classic Itô formula, disappears may be explained as follows:
Let \( F(z) = P(x, y) + iQ(x, y) \), where \( z = x + iy \). Since \( F \) is analytic, from the Cauchy-Riemann equations \( \frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y} \) and \( \frac{\partial P}{\partial y} = -\frac{\partial Q}{\partial x} \), it follows that both \( P \) and \( Q \) are harmonic functions. Thus \( \Delta P = \Delta Q = 0 \). Because of this fact, the second derivative term in the Itô formula for \( F(Z(b)) - F(Z(a)) \) becomes:
\[
\frac{1}{2} \int_{a}^{b} |\Delta(B_1(t), B_2(t)) + i\Delta Q(B_1(t), B_2(t))| dt
+ \int_{a}^{b} \left[ \frac{\partial^2 P}{\partial x \partial y}(B_1(t), B_2(t)) + \frac{i\partial^2 Q}{\partial x \partial y}(B_1(t), B_2(t)) \right] d(B_1, B_2)_t.
\]
Since \( B_1 \) and \( B_2 \) are independent processes, we get \( \langle B_1, B_2 \rangle_t = 0 \). Thus the second derivative term in Itô formula, for holomorphic functions, is vanishing.

Remark 5.9. Conclusion: We have shown that the stochastic integral with respect to a complex Brownian motion behaves exactly in the same manner as Stratonovich integral. Since the Segal–Bargmann entire functions play the same role as the U-functionals in white noise analysis, one can take the inverse Segal–Bargmann transform \( S^{-1} \) to obtain the corresponding white noise functional, for example,
\[
S^{-1} \left\{ \int_{a}^{b} Z(t)^n dZ(t) \right\} = \int_{a}^{b} B(t)^n dB(t),
\]
where
\[
S^{-1} \varphi(x) = \int_{S'} \varphi(\sqrt{2}(x + iy)) \mu(dy).
\]
It is worthwhile to mention here that $S^{-1}\varphi(x)$ is nothing but the conditional expectation of $\varphi(\sqrt{2}z)$ with respect to the real Brownian motion.

As an application, we consider the linear Itô stochastic differential equation
\begin{equation}
X_t = X_0 + c \int_0^t X_s \, ds + \sigma \int_0^t X_s \, dB_s, \quad t \in [0, T],
\end{equation}
for given constants $c$ and $\sigma > 0$. Let us first consider the corresponding stochastic differential equation with $B_t$ being replaced by the complex Brownian motion $\sqrt{2}Z(t)$, i.e.
\begin{equation}
\frac{dX^c_t}{dt} = cX^c_t + \sqrt{2}\sigma X^c_t \dot{Z}(t), \quad X^c_0 = X_0.
\end{equation}
We can solve Equation (5.6) by employing the method of usual ordinary differential equation,
\[X^c_t = X_0 e^{ct + \sqrt{2}\sigma Z_t}.
\]
Finally, taking conditional expectation for $X^c_t$ with respect to the real Brownian motion, we obtain
\[X_t = X_0 e^{ct} \int_{S^1} e^{\sigma(x+iy,h_1)} \mu(dy) = X_0 e^{(c-\frac{1}{2}\sigma^2)t + \sigma(x,h_1)} = X_0 e^{(ct-\frac{1}{2}\sigma^2)t + \sigma B_t}, \quad t \in [0, T].\]

Note that $X_t$ is a geometric Brownian motion which solves the stochastic differential equation in Equation (5.5).

**Acknowledgment.** The authors would like to thank Professor Aurel Stan for making numerous comments and suggestions, which lead to the improvement of this paper. In particular, Remark 5.8 is added to explain the reason why the second derivative term in the classical Itô formula disappears.

**References**


Yuh-Jia Lee: Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung, Taiwan 811
E-mail address: yjlee@nuk.edu.tw

Kuang-Ghieh Yen: Institute of Statistics, National University of Kaohsiung, Kaohsiung, Taiwan 811
E-mail address: d0964401@nuk.edu.tw
AN INFINITE DIMENSIONAL STOCHASTIC ANALYSIS
APPROACH TO LOCAL VOLATILITY DYNAMIC MODELS

R. CARMONA AND S. NADTOCHIY

Abstract. The difficult problem of the characterization of arbitrage free
dynamic stochastic models for the equity markets was recently given a new
life by the introduction of market models based on the dynamics of the local
volatility. Typically, market models are based on Itô stochastic differential
equations modeling the dynamics of a set of basic instruments including, but
not limited to, the option underliers. These market models are usually recast
in the framework of the HJM philosophy originally articulated for Treasury
bond markets. In this paper we streamline some of the recent results on
the local volatility dynamics by employing an infinite dimensional stochastic
analysis approach as advocated by the pioneering work of L. Gross and his
students.

1. Introduction and Notation

The difficult problem of the characterization of arbitrage free dynamic stochastic
models for the equity markets was recently given a new life in [2] by the intro-
duction of market models based on the dynamics of the local volatility surface.
Market models are typically based on the dynamics of a set of basic instruments
including, but not limited to, the option underliers. These dynamics are usually
given by a continuum of Itô’s stochastic differential equations, and the first order
of business is to check that such a large set of degrees of freedom in the model
specification does not introduce arbitrage opportunities which would render the
model practically unacceptable.

Market models originated in the groundbreaking original work of Heath, Jarrow
and Morton [11] in the case of Treasury bond markets. These authors modeled the
dynamics of the instantaneous forward interest rates and derived a no-arbitrage
condition in the form of a drift condition. This approach was extended to other
fixed income markets and more recently to credit markets. The reader interested
in the HJM approach to market models is referred to the recent review article
[1]. However, despite the fact that they were the object of the first success of
the mathematical theory of option pricing, the equity markets have offered the
strongest resistance to the characterization of no-arbitrage in dynamic models.
This state of affairs is due to the desire to accommodate the common practice of
using the Black-Scholes implied volatility to code the information contained in the
prices of derivative instruments. Indeed, while defining stochastic dynamics for the

2000 Mathematics Subject Classification. Primary 91B24.
Key words and phrases. Volatility, market models, abstract Wiener space.
implied volatility surface is rather natural (see, for example [4, 5, 8]), deriving no-arbitrage conditions is highly technical and could be only done in specific particular cases [7, 14, 16, 15].

In the present paper, we streamline some of the recent results on local volatility dynamics by employing an infinite dimensional stochastic analysis approach as advocated by the pioneering work of L. Gross and his students.

One of the main technical results of [2] is the semi-martingale property of call option prices corresponding to a local volatility surface which evolves over time according to a set of Itô's stochastic differential equations. We denote by $C_t(T, K)$ the price at time $t$ of a European call option with maturity $T \geq t$ and strike $K > 0$. For each fixed $t > 0$ we have

$$\frac{\partial}{\partial T}C_t(T, K) = \frac{1}{2} a^2_t(T, K) K^2 \frac{\partial^2}{\partial K^2}C_t(T, K), \quad t < T$$

(1.1)

To be more specific, if for each maturity $T > 0$ and strike $K > 0$, we have

$$d a_t(T, K) = \alpha_t(T, K)dt + \beta_t(T, K) \cdot dW_t,$$

(1.2)

the result we revisit here says that the solution of the Dupire PDE (1.1) is a semi-martingale whenever the second order term coefficient $a^2_t(T, K)$ has for each $T > 0$ and $K > 0$, a stochastic Itô's differential of the form (1.2).

The goal of this paper is to simplify the proof of this result, while at the same time extending it to the case of infinitely many driving Wiener processes $W_t$. Our new proof uses the general framework of infinite dimensional analysis. It streamlines the main argument and gets rid of a good number of technical lemmas proved in [2]. The theoretical results from functional analysis and infinite dimensional stochastic analysis which are needed in this paper can be found in Kuo's original Lecture Notes in Mathematics [13], and in the more recent book by Carmona and Tehranchi [3]. Already, this book was dedicated to Leonard Gross for his groundbreaking work on abstract Wiener spaces and the depth of his contribution to infinite dimensional stochastic analysis. Contributing the present paper to a volume in the honor of his 70th birthday is a modest way to show our deep gratitude.

2. Solutions of the Pricing Equations

As explained in the introduction, we denote by $C_t(T, K)$ the price at time $t$ of a European Call option with strike $K$ and maturity $T$. It is a random variable measurable with respect to the σ-field $\mathcal{F}_t$ of the natural filtration of a Wiener process $W = \{W_t\}_t$. Throughout the paper, we use the notation $\tau = T - t$ for the time to maturity, and we find it convenient to use the notation $x = \log(K/S)$.

2.1. Pricing PDEs. We will find convenient to use the notation

$$\bar{C}_t(\tau, x) := \frac{1}{S_t} C_t(t + \tau, S_t e^x), \quad \tau > 0, x \in \mathbb{R}.$$

for call prices and

$$D_x := \frac{1}{2} (\partial_{xx}^2 - \partial x), \quad D_x^+ := \frac{1}{2} (\partial_{xx}^2 + \partial x)$$
for partial differential operators which we use throughout the paper. Then, if we consider the local volatility \( a_2^2(T, K) \) as given, and introduce the notation

\[
\tilde{a}_2^2(t + \tau, S_t e^x),
\]

then we can conclude that the call price \( \tilde{C}_t(\ldots) \) satisfies the following initial-value problem

\[
\begin{cases}
\partial_\tau w = \tilde{a}_2^2(\tau, x) D_x w(\tau, x) \\
w(0, x) = (1 - e^x)^+. 
\end{cases}
\] (2.1)

We will introduce more notation later in the paper, but for the time being we denote by \( p(\tilde{a}_2^2; \tau, x; u, y) \), with \( \tau > u \), the fundamental solution of the forward partial differential equation (PDE for short) in (2.1) with coefficient \( \tilde{a}_2^2 \). Similarly, we introduce \( q(\tilde{a}_2^2; u, y; \tau, x) \), with \( u < \tau \), the fundamental solution of the backward equation

\[
\partial_u w = -\tilde{a}_2^2(u, y) D_y w(u, y),
\] (2.2)

which is, in a sense, dual to (2.1). We will sometimes drop the argument \( \tilde{a}_2^2 \) of the fundamental solutions \( p \) and \( q \), when the coefficient \( \tilde{a}_2^2 \) is assumed to stay the same. Notice that, if \( w \) is the solution of (2.1), we have

\[
D_x w(\tau, x) = \frac{1}{2} e^x q(0, 0; \tau, x).
\] (2.3)

This equality will be used later in the paper.

2.2. Fréchet Differentiability. For each fixed \( \bar{\varepsilon} > 0 \) and integers \( k, m \geq 1 \), and for any smooth function \( (\tau, x) \mapsto f(\tau, x) \) defined in the strip \( S = [0, \bar{\varepsilon}] \times \mathbb{R} \), we define the norm

\[
\|f\|_{C^{k,m}(S)} = \sup_{(\tau,x) \in S} \left( \sum_{i=0}^k |\partial_{\tau}^i f(\tau, x)| + \sum_{j=1}^m |\partial_x^j f(\tau, x)| \right).
\]

Next we denote by \( \tilde{B} \) the space of functions \( f \) on \( S \) which are continuously differentiable in the first argument and five times continuously differentiable in the second argument, and for which the norm \( \|\cdot\|_{C^{1,5}(S)} \) is finite. We subsequently denote \( \|\cdot\|_{\tilde{B}} := \|\cdot\|_{C^{1,5}(S)} \).

Now we fix \( \varepsilon > 0 \), and we define the strip \( S_\varepsilon \) by \( S_\varepsilon = [\varepsilon, \bar{\varepsilon}] \times \mathbb{R} \). We then define \( \tilde{W}_\varepsilon = C^{1,2}(S_\varepsilon) \) and the mapping

\[
F_\varepsilon : \tilde{B} \hookrightarrow \tilde{W}_\varepsilon,
\]

where, for any \( h \in \tilde{B} \), the image \( F_\varepsilon(h) \) is the restriction to \( S_\varepsilon \) of the solution of (2.1) with \( e^h \) in lieu of the coefficient \( \tilde{a}_2^2 \). Notice that \( e^h \in \tilde{B} \), and that it is bounded away from zero, implying that \( F_\varepsilon(h) \) is well defined.

We are ready to state and prove the main functional analytic result of the paper. This result is technical in nature, but it should be viewed as the work horse for the paper.
Proposition 2.1. The mapping $F_{ξ} : \tilde{B} \mapsto \tilde{W}_ξ$ defined above is twice continuously Fréchet differentiable and for any $h, h', h'' \in \tilde{B}$, we have

$$ F_{ξ}'(h)[h'](τ, x) = \frac{1}{2} \int_0^τ \int_\mathbb{R} h'(u, y)e^{h(u, y)} + yp(e^h; τ, x; u, y)q(e^h; 0, 0; u, y)dydu, $$

and

$$ F_{ξ}''(h)[h', h''](τ, x) = \frac{1}{2} \int_0^τ \int_\mathbb{R} h'(u, y)e^{h(u, y)} + y. $$

$$ \left( \int_\mathbb{R} p(e^h; τ, x; v, z)e^{h(v, z)}h''(v, z)D_zp(e^h; v, z; u, y)dvdv \right) q(e^h; 0, 0; u, y) $$

$$ -p(e^h; τ, x; u, y) \left( \int_0^u q(e^h; 0, 0; v, z)e^{h(v, z)}h''(v, z)D_zq(e^h; v, z; u, y)dvdv \right) dydu $$

Proof. Our proof is based on a systematic use of uniform estimates on the fundamental solutions of the parabolic equations (2.1) and (2.2), and their derivatives. These estimates are known as Gaussian estimates. Typically, they hold when the second order coefficients are uniformly bounded together with a certain number of its derivatives. As a preamable to the technical details of the proof, we first state the Gaussian estimates on the fundamental solutions that we will use in this paper. If $Γ$ denotes the fundamental solution of (2.1) or (2.2), then the following estimate holds

$$ \left| \partial^m_{x} \partial^k_{y} \Gamma(τ, x; u, y) \right| \leq \frac{C}{τ - u^{(1+m+k)/2}} \exp \left( -c \frac{(x - y)^2}{τ - u} \right), \tag{2.4} $$

and consequently

$$ \left| \partial^l_{x} \partial^k_{y} \Gamma(τ, x + y; u; y) \right| \leq \frac{C}{τ - u^{(1+m+k)/2}} \exp \left( -c \frac{x^2}{τ - u} \right), \tag{2.5} $$

for $0 \leq k + m \leq 4$, $i = 0, 1$, $τ \neq u \in [0, 7]$ and $x, y \in \mathbb{R}$. Here, the constants $c$ and $C$ depend only upon the lower bound of $\tilde{a}^2(τ, x)$ and the norm $\|\tilde{a}^2\|_{C^{1,ε}(S)}$, where $\tilde{a}^2$ is the coefficient in the PDEs (2.1) and (2.2).

Inequality (2.4) is derived on pp. 251-261 of [9]. The comments on the dependence of constants $c$ and $C$ on $\tilde{a}^2$ are given in [12].

Fix $h \in \tilde{B}$. Estimate (2.4) holds for $p(e^{h+h'}; τ, x; u, y)$ and $q(e^{h+h'}; τ, x; u, y)$, uniformly over $h'$ varying in a neighborhood of zero, say $U(0) \subset \tilde{B}$. In the following we consider only $h' \in U(0)$.

We now extend the properties of the fundamental solutions to a larger class of functions. For each integer $s \geq 0$ we introduce the space $\tilde{G}^s$:

Definition 2.2. We say that a family of functions $Γ = \{Γ(λ; ...; ...)\}_{λ \in A}$ belongs to $\tilde{G}^s(Λ)$ if, for each $λ \in Λ$, the function $Γ(λ; τ, x; u, y)$ is defined for all $0 \leq u < τ \leq 7$, $τ, x, y \in \mathbb{R}$, and:

1. $Γ$ is $s$ times differentiable in $(x, y)$, and its derivatives are jointly continuous in $(τ, x; u, y)$, moreover, $Γ$ satisfies estimates (2.4), for $0 \leq k + m \leq s$, uniformly over $λ \in Λ$;
Definition 2.3. The family of functions $\Gamma$ is said to belong to class $\mathcal{G}^s(\Lambda)$, for some integer $s \geq 0$, if it belongs to $\tilde{\mathcal{G}}^s(\Lambda)$, and, in addition, satisfies the following: if $s \geq 2$, then $\Gamma$ is continuously differentiable in $\tau$, and, for all $\lambda \in \Lambda$,

$$\partial_\tau \Gamma(\lambda; \tau, x; u, y) = \sum_{i=0}^{2} f_i(\lambda; \tau, x) \partial^i_x \tilde{\Gamma}_1(\lambda; \tau, x; u, y),$$

where each $\tilde{\Gamma}_1 \in \tilde{\mathcal{G}}^s$, and each $\|f_i(\lambda; \ldots)\|_{C^1,\ldots,2(\mathcal{S})}$ is bounded over $\lambda \in \Lambda$.

For the most part of this proof we assume that the functions are parameterized by the set $\Lambda = U(0) \subset \mathcal{B}$, and therefore drop the argument $\Lambda$ of the class $\mathcal{G}^s$. Notice that the families of fundamental solutions

$$\{p(e^{h^i}; \ldots)\}_{h^i \in U(0)} \quad \text{and} \quad \{q(e^{h^i}; \ldots)\}_{h^i \in U(0)}$$

belong to $\mathcal{G}^4$.

We now derive some important properties of the classes of functions introduced above. Let us consider $\Gamma_1, \Gamma_2 \in \tilde{\mathcal{G}}^s$ with $s \geq 2$, let us fix integers $i, k, j, m$ satisfying

$$0 \leq i + k + j + m \leq s + 1, \quad (i + k) \lor (j + m) \leq s$$

and let $f \in C^{i+k-1+j+m-1}(\mathcal{S})$. Then, for all $\lambda_1, \lambda_2 \in \Lambda$, $x_1, x_2 \in \mathbb{R}$, and $0 \leq \tau_1 < \tau_2 < \tau$, we have:

$$\int_{\tau_1}^{\tau_2} \int_{\mathcal{R}} \frac{\partial^{i+k}}{\partial x_1^i \partial y^k} \Gamma_2(\lambda_2; \tau_2, x_2; u, y) f(u, y) \frac{\partial^{j+m}}{\partial x_1^j \partial y^m} \Gamma_1(\lambda_1; u, y; \tau_1, x_1) dy \, du$$

$$= \int_{\tau_1}^{\tau_2} \int_{\mathcal{R}} \frac{\partial^{m+1}}{\partial y^m} \left( \frac{\partial}{\partial x_1} + \frac{\partial}{\partial y} \right)^j \Gamma_1(\lambda_1; u, y; \tau_1, x_1) \left[ \frac{\partial^{m-1}}{\partial y^{m-1}} \Gamma_2(\lambda_2; \tau_2, x_2; u, y) f(u, y) \frac{\partial^{i+k}}{\partial x_1^i \partial y^k} \right] dy \, du$$

$$+ \int_{\tau_1}^{\tau_2} \int_{\mathcal{R}} \frac{\partial^{k+1}}{\partial y^k} \left( \frac{\partial}{\partial x_2} + \frac{\partial}{\partial y} \right)^i \Gamma_2(\lambda_2; \tau_2, x_2; u, y) \left[ \frac{\partial^{k-1}}{\partial y^{k-1}} \Gamma_1(\lambda_1; u, y; \tau_1, x_1) \right] dy \, du$$

$$+ \int_{\tau_1}^{\tau_2} \int_{\mathcal{R}} \frac{\partial^{i+k}}{\partial x_1^i \partial y^k} \left[ f(u, y) \frac{\partial^{j+m}}{\partial x_1^j \partial y^m} \Gamma_1(\lambda_1; u, y; \tau_1, x_1) \right] dy \, du.$$
Now, fix some $s \geq 2$, choose some $\Gamma_1, \Gamma_2 \in \hat{G}^s$ and any family of functions $\{f(\lambda; \cdot, \cdot) \in C^{1,s-1}(\mathcal{S})\}_{\lambda \in \Lambda}$, and define

\[
I[\Gamma_2, f, \Gamma_1](\lambda; \tau_2, x_2; \tau_1, x_1) := \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) f(\lambda, u, y) D_y \Gamma_1(\lambda; u, y; \tau_1, x_1) dy du.
\]

We are going to show that

\[
I[\Gamma_2, f, \Gamma_1](\lambda; \tau_2, x_2; \tau_1, x_1) = \|f(\lambda)\| \|\Gamma_3(\lambda; \tau_2, x_2; \tau_1, x_1)\|
\]

for some $\Gamma_3 \in \hat{G}^{s-1}$.

If, for some $\lambda \in \Lambda$, $f(\lambda) \equiv 0$, the statement of the claim is obvious. Therefore we will assume that $\|f(\lambda)\| > 0$. The smoothness of $\Gamma_3$ in $(x_1, x_2)$, and estimate (2.4) follow from (2.6), after we integrate by parts in the definition of $I$. To obtain inequality (2.5), we only need to make a shift of the integration variable and proceed as in (2.6).

We now verify the second condition of Definition 2.2. Pick some $g \in C^1_0(\mathbb{R})$, and, assuming that $s \geq 2$, proceed as follows

\[
\left| \int_{\mathbb{R}} g(x_1) \Gamma_3(\lambda; \tau_2, x_2; \tau_1, x_1) dx_1 \right| \\
= \left| \int_{\mathbb{R}} g(x_1) \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) f(\lambda; u, y) \frac{\partial}{\partial f(\lambda)} \frac{\partial}{\partial g(x_1)} \Gamma_1(\lambda; u, y; \tau_1, x_1) dy du dx_1 \right|
\]
\[
\begin{align*}
    \leq c_6 \int_{\mathbb{R}} (|g(x)| + |g'(x)|) \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) \frac{|f(\lambda; u, y)|}{\|f(\lambda)\|} \, dy \, du \\
    = c_6 \int_{\mathbb{R}} (|g(x)| + |g'(x)|) \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) \frac{|f(\lambda; u, y)|}{\|f(\lambda)\|} \, dy \, du
\end{align*}
\]

\[\sum_{i=0}^{2} \sum_{j=0}^{1} \left| \partial_{y} \partial_{x_i} \partial_{x_j} \frac{\Gamma_1(\lambda; u, y; \tau_1, x_1)}{\|f(\lambda)\|} \right| \, dy \, du \, dx_1
\]

which goes to zero as \( \tau_1 \to \tau_2 \). We integrated by parts in \( x_1 \), and applied estimates (2.4), (2.5) to obtain the above inequality. The interchangeability of integration and differentiation is justified by (2.6) (just notice that, as it is clear from the first line of (2.6), the integrals are, sometimes, understood as iterated rather than double integrals). The above estimate proves that \( \Gamma_3 \) satisfies the second condition in Definition 2.2.

Now, assume that, in addition, \( \Gamma_1 \) and \( \Gamma_2 \) belong to \( \mathcal{G}^s \). We claim that, in this case, \( \Gamma_3 \) is in \( \mathcal{G}^{s-1} \). We only need to verify the additional property in the Definition 2.3. Assume \( s - 1 \geq 2 \), then, using the expression for the \( \tau_2 \) derivatives of \( \Gamma_2 \), and the fact that \( \Gamma_3 \in \mathcal{G}^{s-1} \), we obtain the following

\[
\frac{\partial}{\partial \tau_2} \left[ \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) \frac{f(\lambda; u, y)}{\|f(\lambda)\|} \, dy \, du \right]
\]

\[= c_6 \int_{\mathbb{R}} (|g(x)| + |g'(x)|) \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) \frac{|f(\lambda; u, y)|}{\|f(\lambda)\|} \, dy \, du
\]

where each \( f_i(\lambda; \ldots) \) is in \( C^{1,s-2}(\mathcal{S}) \), and the \( \Gamma_i \)'s belong to \( \mathcal{G}^s \). The above decomposition completes the proof of the claim: \( \Gamma_3 \in \mathcal{G}^{s-1} \).

It is easy to see, integrating by parts, that the operator \( J \) defined by

\[J[\Gamma_2, f, \Gamma_1](\lambda; \tau_2, x_2; \tau_1, x_1) := \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} D_y \Gamma_2(\lambda; \tau_2, x_2; u, y) f(\lambda, u, y) \Gamma_1(\lambda; u, y; \tau_1, x_1) \, dy \, du
\]

has the same properties as \( I \).

Similarly, for any \( \{f(\lambda; \ldots) \in C^{1\cdot 2}(\mathcal{S})\}_{\lambda \in \Lambda} \), and \( \Gamma_1, \Gamma_2 \in \mathcal{G}^2 \), we define the function \( K[\Gamma_2, f, \Gamma_1] \) by:

\[K[\Gamma_2, f, \Gamma_1](\lambda; \tau_2, x_2; \tau_1, x_1) := \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \Gamma_2(\lambda; \tau_2, x_2; u, y) e^{y-x_1} f(\lambda, u, y) \Gamma_1(\lambda; u, y; \tau_1, x_1) \, dy \, du
\]

and, using (2.6) and (2.7), we obtain the estimate:

\[|\partial_{\tau_2} K| + \sum_{j=0}^{2} |\partial_{x_j} K| \leq c_8 \|f(\lambda)\|(\tau_2 - \tau_1)^{-3/2} \exp \left(-c_9 \frac{(x_2 - x_1)^2}{\tau_2 - \tau_1} \right),
\]

where the constants \( c_8, c_9 \) depend on \( \Gamma_1 \) and \( \Gamma_2 \), but not on \( \lambda \).
We now proceed with the proof of the proposition. Writing the initial value problem (2.1) twice, first with \(e^h\), and then with \(e^{h+h'}\), and subtracting one from another, we can, formally, apply the Feynman-Kac formula and obtain

\[
\mathbf{F}_\varepsilon(h+h')(\tau, x) = \mathbf{F}_\varepsilon(h)(\tau, x) + \frac{1}{2} \int_0^\tau \int_\mathbb{R} \left\lfloor p(e^{h+h'}; \tau, x; u, y) e^{h'(u, y)} (e^{h'(u, y)} - 1) q(e^h, 0, 0; u, y) \right\rfloor dy du.
\]

This representation follows from the uniqueness of weak solution of (2.1), see, for example, [6] for details. Applying the same technique to the fundamental solution \(p\), we get

\[
\Delta p(\tau, x; u, y) := p(e^{h+h'}; \tau, x; u, y) - p(e^h; \tau, x; u, y) = \int_0^\tau \int_\mathbb{R} p(e^{h+h'}; \tau, x; u, y) e^{h(u, y)} (e^{h(u, y)} - 1) D_x p(e^h; v, z; u, y) dy dz dv
\]

(2.10)

Since all the families of functions considered in this part of the proof are parameterized by \(h' \in U(0)\), we use the shorter notation \(f(h')\) instead of \(\{f(h')\}_{h' \in U(0)}\), for the arguments of operator \(I\).

We define \(\Delta q\) in a similar way. Next we rewrite (2.10) as

\[
\mathbf{F}_\varepsilon(h+h') = \mathbf{F}_\varepsilon(h) + \mathbf{F}_\varepsilon(h')[h'] + r_1 + r_2,
\]

with

\[
r_1(\tau, x) = \frac{1}{2} \int_0^\tau \int_\mathbb{R} p(e^{h+h'}; \tau, x; u, y) e^{h(u, y)} (e^{h(u, y)} - 1) q(e^h, 0, 0; u, y) dy du
\]

\[
= \frac{1}{2} K \left[p(e^{h+h'}), e^h (e^h - 1 - h') q(e^h)\right](h'; \tau, x; 0, 0),
\]

and

\[
r_2(\tau, x) = \frac{1}{2} \int_0^\tau \int_\mathbb{R} \Delta p(\tau, x; u, y) e^{h(u, y)} q(e^h; 0, 0; u, y) dy du
\]

\[
= \frac{1}{2} K \left[I \left[p(e^{h+h'}), e^h (e^h - 1), p(e^h)\right], e^h h', q(e^h)\right](h'; \tau, x; 0, 0).
\]

Because of the properties of the operator \(I\) derived earlier, it is easy to see that the function \(I \left[p(e^{h+h'}), e^h (e^h - 1), p(e^h)\right]\) belongs to \((e^{h'} - 1) \cdot G^3\). Therefore, using estimate (2.8), we have immediately that for \(i = 1, 2\),

\[
\|r_i\|_{\mathcal{W}_\varepsilon} \leq c_{10} \|h'\|_{\mathcal{B}}^3
\]

and this implies that \(\mathbf{F}_\varepsilon\) is Fréchet differentiable, with Fréchet derivative as given in the statement of the proposition. The fact that \(\mathbf{F}_\varepsilon'(h)[.]\) is bounded on the unit ball of \(\mathcal{B}\) follows, again, from (2.8).

We now compute the Fréchet derivative of \(\mathbf{F}_\varepsilon'(.)\) using the same technique as in the first part of the proof.
We fix \( h \in \mathcal{B} \) and we consider families of functions parameterized by \((h', h'') \in \Lambda := \mathcal{B} \times U(0)\). We redefine \( \Delta p \), using \( h'' \) instead of \( h' \) in (2.10). Then we have

\[
(F^\prime(h + h'') - F^\prime(h))[h'](\tau, x) = 
\frac{1}{2} \int_0^\tau \int_{\mathbb{R}} h'(u, y)e^{h(u,y) + y} \Delta p(\tau, x; u, y) q(e^h; 0, 0; u, y) dydu 
+ \frac{1}{2} \int_0^\tau \int_{\mathbb{R}} h'(u, y)e^{h(u,y) + y} p(e^h; \tau, x; u, y) \Delta q(0, 0; u, y) dydu 
+ \frac{1}{2} \int_0^\tau \int_{\mathbb{R}} h'(u, y)e^{h(u,y) + y} \Delta p(\tau, x; u, y) q(0, 0; u, y) dydu 
+ \frac{1}{2} \int_0^\tau \int_{\mathbb{R}} h'(u, y)e^{h(u,y) + y} (e^{h''(u,y)} - 1) p(e^{h + h''}; \tau, x; u, y) 
\cdot q(e^h; 0, 0; u, y) dydu 
\tag{2.11}
\]

Next, we decompose the first integral in (2.11)

\[
\int_0^\tau \int_{\mathbb{R}} h'(u, y)e^{h(u,y) + y} \Delta p(\tau, x; u, y) q(e^h; 0, 0; u, y) dydu 
= K \left[I \left[p(e^h), e^{h''} - 1 - h'', p(e^h)\right], e^{h''} - 1, p(e^h)\right] (h', h''; \tau, x; 0, 0) 
+ K \left[I \left[p(e^h), e^{h''} - 1 - h'', p(e^h)\right], e^{h''} - 1, p(e^h)\right] (h', h''; \tau, x; 0, 0) 
+ K \left[I \left[p(e^h), e^{h''} - 1, p(e^h)\right], e^{h''} - 1, p(e^h)\right] (h', h''; \tau, x; 0, 0). 
\]

The first term in the right hand side of the above expression is linear in \( h'' \). It is the first component of \( \mathbf{F}'' \). Using the properties of the operator \( I \), we conclude that

\[
I \left[p(e^h), e^{h''} - 1 - h'', p(e^h)\right] 
+ I \left[p(e^h), e^{h''} - 1, p(e^h)\right], e^{h''} - 1, p(e^h)\right] 
= \|h''\| G\|h''\|_G^2 \Gamma, 
\]

where \( \Gamma \in \mathcal{G}^2 \). Therefore, using estimate (2.8), we conclude that the \( \| \cdot \|_{\mathcal{W}_G} \) norms of the last two terms in the right hand side of (2.12) are bounded by a constant times \( \|h''\| G\|h''\|_G^2 \).

A similar decomposition holds true for the second integral in the right hand side of (2.11) provided the operator \( I \) is replaced by \( J \). Moreover, the \( \| \cdot \|_{\mathcal{W}_G} \) - norms of last two integrals in (2.11) are also bounded by a constant times \( \|h''\| G\|h''\|_G^2 \).

to see this, recall (2.10) and write its analog for \( \Delta q \), then apply the properties operators \( I \) and \( J \), and use estimate (2.8). This yields the existence of \( \mathbf{F}''(h) \), as given in the proposition.

To show the continuity of the second derivative, fix any \( h' \) and \( h'' \) in \( \mathcal{B} \) and consider any \( \Delta h \in U(0) \). We only show the continuity of the first component of \( \mathbf{F}''(\cdot)[h', h''] \) at \( h \), uniformly over \( h' \) and \( h'' \) in a bounded set. The proof for the second component is the same. We introduce the difference
\[ \Delta K := K \left[ I \left( p(e^{h+\Delta h}), h''e^{h+\Delta h}, p(e^{h+\Delta h}) \right), e^{h+\Delta h}h', q(e^h + \Delta h) \right] - K \left[ I \left( p(e^h), h''e^h, p(e^h) \right), e^h h', q(e^h) \right] = K \left[ I \left( p(e^h), h''e^h, p(e^h) \right), e^h h', q(e^h) \right] + K \left[ I \left( p(e^h), h''e^h, p(e^h) \right), e^{h+\Delta h}h', q(e^{h+\Delta h}) \right] + K \left[ I \left( p(e^h), h''e^h, p(e^h) \right), h' e^{h+\Delta h} - 1, q(e^{h+\Delta h}) \right] + K \left[ I \left( p(e^h), h''e^h, p(e^h) \right), h' e^h, -\beta \left( q(e^{h+\Delta h}), e^h(e^{\Delta h} - 1), q(e^h) \right) \right]. \]

And, as before, using the properties of \( I, J \) and \( K \), we conclude that
\[ \| \Delta K \|_{\mathcal{W}_\epsilon} \leq c_{11} \| h' \|_{\mathcal{W}_\epsilon} \| h'' \|_{\mathcal{W}_\epsilon} \| \Delta h \|_{\mathcal{W}_\epsilon}. \]

which completes the proof of the proposition. \( \square \)

Recall that the price \( C_T(T, x) \) at time \( t \) of an European call option is given by \( w(\tilde{a}^2; T - t, x + \log S_t) \), where \( w(\tilde{a}^2; \ldots) \) is the solution of (2.1). Therefore, in order to get to the Fréchet differentiability of the price of a call option from the above result, we will need to compose \( F_\epsilon \) with another mapping. This justifies the introduction, for each \( T \in [\epsilon, \tau] \) and \( x \in \mathbb{R} \) of the mapping
\[ \delta_{T,x} : [0, T - \epsilon] \times \mathcal{W}_\epsilon \times \mathbb{R} \hookrightarrow \mathbb{R} \]
defined by
\[ \delta_{T,x}(t, w, y) = w(T - t, x + y). \]

We have:

**Proposition 2.4.**

1. For each \((w, y) \in \mathcal{W}_\epsilon \times \mathbb{R}, \delta_{T,x}(., w, y)\) is continuously differentiable, and the partial derivative \( \partial \delta_{T,x}/\partial t \) is a continuous functional on \([0, T - \epsilon] \times \mathcal{W}_\epsilon \times \mathbb{R} \).

2. For each \( t \in [0, T - \epsilon], \delta_{T,x}(t, \ldots) \) is twice Fréchet differentiable and for any \( w, w', w'' \in \mathcal{W}_\epsilon \) and \( y, y', y'' \in \mathbb{R} \), its derivatives satisfy
\[ \delta_{T,x}(t, w, y)[w', y'] = w'(T - t, x + y) + y' \partial_x w(T - t, x + y) \]
and
\[ \delta_{T,x}(t, w, y)[(w', y'), (w'', y'')] = y'' \partial_x w'(T - t, x + y) + y' \partial_x w'(T - t, x + y) + y'' \partial_x w''(T - t, x + y). \]

Moreover, \( \delta_{T,x} \) and \( \delta_{T,x}'' \) are continuous operators from \([0, T - \epsilon] \times \mathcal{W}_\epsilon \times \mathbb{R} \) into \( \mathcal{W}_\epsilon \times \mathbb{R} \) and \( L \left( \mathcal{W}_\epsilon \times \mathbb{R}, \mathcal{W}_\epsilon \times \mathbb{R} \right) \) respectively.

**Proof.** Let us fix \((w, y) \in \mathcal{W}_\epsilon \times \mathbb{R}\). Then, for any \( t \in [0, T - \epsilon] \), we have
\[ \partial \delta_{T,x}(t, w, y)/\partial t = -\partial_t w(T - t, x + y). \]
We first show that this functional is continuous in \((t, w, y) \in [0, T-\varepsilon] \times \tilde{W}_\varepsilon \times \mathbb{R}\). Consider any \((t', w', y') \in [0, T-\varepsilon] \times \tilde{W}_\varepsilon \times \mathbb{R}\), then

\[
|\partial_t w(T-t, x+y) - \partial_t w'(T-t', x+y')| = |\partial_x w(T-t, x+y) - \partial_x w(T-t', x+y')| + |\partial_y w(T-t', x+y') - \partial_y w'(T-t', x+y')| \tag{2.12}
\]

The first difference in the right hand side above can be made as small as we want by choosing \((t, x)\) and \((t', x')\) close enough. The second difference is bounded by \(\|w - w'\|_{\tilde{W}_\varepsilon}\). This implies continuity of the partial derivative \(\partial_\delta T_{_x}/\partial t\), proving the first statement of the proposition.

Let us now compute the derivatives of \(\delta_{T,x}\). We will keep \((t, w, y) \in [0, T-\varepsilon] \times \tilde{W}_\varepsilon \times \mathbb{R}\) fixed, and consider \((w', y') \in U(0) \subset \tilde{W}_\varepsilon \times \mathbb{R}\), where \(U(0)\) is a neighborhood of zero. Notice that

\[
\delta_{T,x}(t, w + w', y + y') - \delta_{T,x}(t, w, y) = w'(T-t, x+y+y') - w'(T-t, x+y) + y'\partial_x w'(T-t, x+y) + y'\partial_y w'(T-t, x+y + \xi y'),
\]

for some \(\xi \in [0,1]\), and that

\[
|\partial_y(y') + y'\partial_x w'(T-t, x+y + \xi y')| = \tilde{\delta} \left( |y'|^2 + \|w'||_{\tilde{W}_\varepsilon}^2 \right).
\]

Therefore, we have obtained the expression for \(C_{T,x}'\), as given in the proposition.

Now consider \((w', y') \in \tilde{W}_\varepsilon \times \mathbb{R}\) and \((w'', y'') \in U(0) \subset \tilde{W}_\varepsilon \times \mathbb{R}\), the rest of parameters being fixed. Then:

\[
\delta_{T,x}'(t, w + w', y + y') - \delta_{T,x}'(t, w, y) = w''(T-t, x+y+y'') - w''(T-t, x+y) + y''\partial_x w''(T-t, x+y+y'')
\]

\[
+ y'\partial_x w'(T-t, x+y+y') - y'\partial_x w(T-t, x+y) = y''\partial_x w''(T-t, x+y) + (y'')^2 \partial_x^2 w'(T-t, x+y + \xi y'') + y''\partial_x w''(T-t, x+y) + y'(y'' \partial_x^2 w(T-t, x+y) + \tilde{\delta}(y'')),
\]

for some \(\xi, \xi' \in [0,1]\). Again, noticing that

\[
|\partial(y'')^2 \partial_x^2 w''(T-t, x+y + \xi y'') + y''\partial_x^2 w''(T-t, x+y + \xi y'') + y'\tilde{\delta}(y'')| \leq \sqrt{|y'|^2 + \|w'||_{\tilde{W}_\varepsilon}^2} \tilde{\delta} \left( |y''|^2 + \|w''|^2_{\tilde{W}_\varepsilon} \right).
\]

we get the desired expression for \(\delta_{T,x}''\).

In order to show the continuity of \(\delta_{T,x}'\) and \(\delta_{T,x}''\), we fix \((w', y'), (w'', y'') \in \tilde{W}_\varepsilon \times \mathbb{R}\), and we prove the continuity of \(\delta_{T,x}'(\ldots)[w', h']\) and \(\delta_{T,x}''(\ldots)[(w', h'), (w'', h'')]\) by, essentially, repeating the argument of (2.13). Finally, notice that the continuity is uniform over \((w', y'), (w'', y'')\) when they are restricted to a bounded set. \(\square\)

Now, consider the composition of the two operators introduced above. For each \(T \in (\varepsilon, \bar{T}]\) and \(x \in \mathbb{R}\), we have

\[
C_{T,x} : [0, T-\varepsilon] \times \tilde{E} \times \mathbb{R} \to \mathbb{R}
\]

\[
C_{T,x}(t, h, y) = \delta_{T,x}(t, \mathbf{F}_\varepsilon(h), y)
\]
As a composition of twice Fréchet differentiable operators, $C_{T,x}(t,\ldots)$ is, clearly twice Fréchet differentiable, for each $t \in [0, T - \delta]$. Due to the continuity of $F^m_{\tau}(\cdot)$, $\delta^m_{T,x}(...)$, and $\delta''_{T,x}(...)$, the Fréchet derivatives of $C_{T,x}(t, h, y)$ are also continuous in $(t, h, y)$. Finally, $C_{T,x}$, clearly, satisfies the first statement of Proposition 2.4. Thus, applying the chain rule we obtain the following

**Proposition 2.5.** For each $t \in [0, T - \delta]$, functional $C_{T,x}(t,\ldots)$ is twice Fréchet differentiable, such that, for any $h, h', h'' \in \tilde{B}$ and $y, y', y'' \in \mathbb{R}$, we have

$$C''_{T,x}(t, h, y)[h', y'] = F^m_{\tau}(h)[h', h''](T - t, x + y) + y'\partial_x F(\tau)(h)(T - t, x + y),$$

and

$$C''_{T,x}(t, h, y)[(h'', y''), (h''', y''')] = F^m_{\tau}(h)[h', h''](T - t, x + y) + y'\partial_x F(\tau)(h)(T - t, x + y),$$

and $C_{T,x}, C''_{T,x}$ are continuous operators from $[0, T - \delta] \times \tilde{B} \times \mathbb{R}$ into $\tilde{B}^* \times \mathbb{R}$ and $L\left(\tilde{B} \times \mathbb{R}, \tilde{B}^* \times \mathbb{R}\right)$ respectively.

3. Using Itô’s Formula in Infinite Dimension

The purpose of this section is to extend the proof of the semi-martingale property given in [2] to the case of infinitely many driving Wiener processes.

We denote by $\tilde{B}$ the cylindrical Brownian motion constructed on the canonical cylindrical Gaussian measure of some separable Hilbert space $\tilde{H}$. The reader can think of $\tilde{H} = l_2$ - the space of square - sumable sequences but the specific form of this Hilbert space is totally irrelevant for what we are about to do.

The first step is to construct a Hilbert subspace of $\tilde{B}$. For each functions $f$ and $g$ with enough derivatives square integrables and for each non-negative integers $k$ and $m$, we define the scalar product

$$< f, g >_{\tilde{W}^{k,m}(S)} = \sum_{i=0}^{k} \partial^i_x f(0,0)\partial^i_x g(0,0) + \sum_{j=0}^{m} \partial^j_x f(0,0)\partial^j_x g(0,0)$$

$$+ \int_S \nabla (\partial^k_x f(\tau, x)) \nabla (\partial^k_x g(\tau, x)) + \nabla (\partial^m_x f(\tau, x)) \nabla (\partial^m_x g(\tau, x)) dx d\tau.$$

Now we fix a compact set $K$ contained in $S$ and containing the origin $(0,0)$, and we consider the space of functions on $S$ which are constant outside $K$, namely whose derivatives vanish outside $K$. For the sake of definiteness we will choose $K = [0, \tau] \times [-M, M]$ for a positive (large) number $M$. Equipped with the scalar product $<\ldots>_{\tilde{W}^{k,m}(S)}$, defined above, this space of functions (more precisely of equivalence classes of functions) is a Hilbert space which we denote $H$. It is clearly contained in $\tilde{B}$. Define by $\tilde{B}$, the completion of $H$ in the $\|\cdot\|_{C^{k,m}(S)}$ norm. Thus, the pair $(H, \tilde{B})$ forms a conditional Banach Space.

Clearly, $\tilde{B}$ is a subspace of $\tilde{B}$, and therefore, Proposition 2.5 holds for the restriction of $C_{T,x}$ to $\tilde{B}$ as well.
For any given real separable Banach space $\mathcal{G}$ we denote by $\mathcal{L}(\mathcal{G})$ the space of all non-anticipative random processes in $\mathcal{G}$ (measurable mappings $X : \Omega \times [0, \infty) \rightarrow \mathcal{G}$) such that

$$E \int_0^t \|X_u\|^2_\mathcal{G} du < \infty,$$

for all $t \geq 0$. Where $\mathcal{G}$ is a Banach space. Also, we denote by $L^2(\mathcal{H})$ the space of all Hilbert-Schmidt operators on $\mathcal{H}$.

Next, we choose $\alpha \in L(B)$ and $\beta \in L^2(L^2(\mathcal{H}, \mathcal{H}))$, and we model dynamics of $h_t$, the logarithm of the squared local volatility at time $t$, $\tilde{a}^2_t$, by the infinite dimensional Itô’s stochastic differential

$$dh_t = \alpha_t dt + \beta_t dB_t,$$

which together with an initial condition $h_0 \in B$, defines a random process in $B$.

Also, we assume the following dynamics for the logarithm of the underlying

$$d\log S_t = -\frac{1}{2} \sigma^2_t dt + \sigma_t dB_t, \quad \log S_0,$$

where $\sigma$ is $\mathbb{R}$-valued random process with $E \int_0^t \sigma_u^2 du < \infty$ almost surely, for any $t \geq 0$, and $e_1 \in \mathcal{H}$ is a fixed unit vector.

Now, thanks to Proposition 2.5, we can apply Itô’s formula (see, for example, [13], p. 200) to $(C^T_x(t, h_t, \log S_t))_{t \in [0, T - \bar{\epsilon}]}$. We get that for any $T \in (\bar{\epsilon}, \bar{\tau}]$ and $x \in \mathbb{R}$, we have, almost surely, for all $t \in [0, T - \bar{\epsilon})$,

$$C^T_x(t, h_t, \log S_t) = C^T_x(0, h_0, \log S_0)$$

$$+ \int_0^t \left( \frac{\partial}{\partial t} C^T_x(u, h_u, \log S_u) + C^T_x(u, h_u, \log S_u)[\alpha_u, -\frac{1}{2} \sigma^2_u] \right.$$

$$+ \frac{1}{2} \text{Tr} \left( (\beta_u, \sigma_u e_1)^* \circ C''_{T,x}(u, h_u, \log S_u) \circ (\beta_u, \sigma_u e_1) \right) \bigg) du$$

$$+ \int_0^t C''_{T,x}(u, h_u, \log S_u) \circ (\beta_u, \sigma_u e_1) dB_u$$

where $C''_{T,x}$ and $C''_{T,x}$ are given in Proposition 2.5.

**Remark 3.1.** Since $\bar{\epsilon}$ can be made as small as we want, the above representation holds for any $T \in (0, \bar{\tau}]$, and all $t \in [0, T)$. Then, since we choose $\bar{\tau}$ as large as we want, the above representation holds for any $T > 0$, and all $t \in [0, T)$.

We now restate the above result after choosing a complete orthonormal basis $\{e_n\}_n$ of $\mathcal{H}$. Notice that without any loss of generality we can assume that the first element $e_1$ of this basis is in fact the unit vector entering the equation for the dynamics (3.1) of the logarithm of the underlying spot price. As it should be clear, fixing a basis is essentially assuming that $\mathcal{H} = l_2$. If we consider that $\beta_i$ is given by the sequence $\{\beta_i(\cdot) \in \mathcal{H}\}_{n=1}^\infty$ of its components on the basis vectors, then we have the following theorem.
Theorem 3.2. For any $T > 0$ and $x \in \mathbb{R}$, we have, almost surely, for all $t \in [0, T)$,
\[
C_{T,x}(t, h_t, \log S_t) = C_{T,x}(0, h_0, \log S_0) \\
+ \int_0^t \left[ F_\varepsilon'(h_u)[\alpha_u] \right] du - \frac{1}{2} \sigma_u^2 \partial^2_x F_\varepsilon(h_u) \partial_t F_\varepsilon(h_u) \\
+ \sigma_u \partial_x F_\varepsilon'(h_u)[\beta_u] + \frac{1}{2} \sum_{n=1}^{\infty} F_\varepsilon''(h_u)[\beta^n_u, \beta^n_u] (T - u, x + \log S_u) \\
+ \int_0^t \left[ \sum_{n=1}^{\infty} F_\varepsilon'(h_u)[\beta^n_u] + \sigma_u \partial_x F_\varepsilon(h_u) \right] (T - u, x + \log S_u) dB^n_u,
\]

if we use the notation $\{B^n\}_n$ for the sequence of independent standard one-dimensional Brownian motions $B^n_t = < e^n, B_t >$. $F_\varepsilon'$ and $F_\varepsilon''$ are given in Proposition 2.1.

This is the infinite dimensional version of the semi-martingale result of [2].

References
R. Carmona: Bendheim Center for Finance, ORFE, Princeton University, Princeton, NJ 08544, USA
E-mail address: rcarmona@princeton.edu

S. Nadtochy: Bendheim Center for Finance, ORFE, Princeton University, Princeton, NJ 08544, USA
E-mail address: snadtoch@princeton.edu
HEDGING CLAIMS WITH FEEDBACK JUMPS IN THE PRICE PROCESS

KISEOP LEE AND PHILIP PROTTER

Abstract. We study a hedging and pricing problem of a model where the price process of a risky asset has jumps with instantaneous feedback from the most recent asset price. We model these jumps with a doubly stochastic Poisson process with an intensity function depending on the current price. We find a closed form expression of the local risk minimization strategy using Föllmer and Schweizer decomposition and Feynman-Kac type integro-differential equation. The possibility that the jumps depend on the most recent price is new for this type of model.

1. Introduction

The original Black-Scholes paradigm of modelling an asset price process with a geometric Brownian motion is still widely used, although there is widespread dissatisfaction with it. Many alternatives have been proposed, including stochastic volatility models, general (non-linear) stochastic differential equations, and also replacing Brownian noise with noise coming from a Lévy process. An advantage of stochastic volatility models is the presence of ‘heavy tails’, an advantage shared by Levy noise models, and the Lévy noise models have the additional attribute of incorporating jumps into price process. All of these theories are well established and well known.

In this article, we propose a new type of price process, which incorporates jumps, but unlike other models the random jump times do not arise by a prior specified exogenous distributional hypotheses. In more simple traditional cases one can model the arrival times of the jumps of the noise via a point process $N$ as follows:

$$N_t = \sum_{i=1}^{\infty} 1_{\{t \geq T_i\}}$$

where $N_t - \int_0^t \lambda(s, \omega) ds = \text{martingale}$,

and the times $T_i$ are the arrival times, with $\lambda_s$ the arrival intensity. A standard feature of these models is that the arrival intensity process $\lambda$ is specified \textit{a priori}, and is either non-random (as for example in the Poisson case, and more generally the Lévy case), or is a given stochastic process. A price model with this variety of
jumps can be written as the solution $S$ of the following type of equation:

$$dS_t = \sigma(S_t)dB_t + b(S_t-)(N_t - \int_0^t \lambda_s ds) + \mu(S_t)dt.$$  \hspace{1cm} (1.1)

In this article we propose a new way of modelling the arrival intensity of the jump times, one where the instantaneous arrival intensity is a function of the past history of the asset price process up to that point. This makes the jumps intrinsic to the pricing process and the past noise itself. Three examples where such a price process construction could be used, are to model (1) large trader behavior where a stop order kicks in at a certain level thereby causing a (perhaps small) jump in the asset price, or (2) a change in the stock price due to a creditor calling a loan, caused by the asset price falling below a certain level, or (3) aggregate behavior of many traders acting in concert due to a run-up or dramatic decline in an asset price, such as (again, for example) a sell-off on a Friday due to a rumor later proved false with a consequent repurchase on the following Monday. We can express this model in heuristic notation (to which we later give a rigorous meaning) as follows:

$$dS_t = \sigma(S_t)dB_t + b(S_t-)(N_t - \int_0^t \lambda(S_t; r \leq s) ds) + \mu(S_t)dt.$$  

We emphasize that the difference in the two types of equations is the instantaneous feedback loop present in the second model, which takes the jump arrival intensity $\lambda$ to be a functional of the past paths of the solution of the stochastic differential equation $S$, the model for the asset price process.

The mathematics involved of making sense of this idea is non trivial, and we will rely on earlier work of J.Jacod and P.Protter [17]. The inspiration to consider models of this type came form work of R.Frey [14].

We will not only make sense of this idea for the price process, but we will find a closed form expression for the hedging strategy for a class of such asset price processes. We will also construct its minimal martingale measure in the Föllmer and Schweizer [11] sense.

On the hedging problem in incomplete markets, the local risk minimization and the mean-variance hedging have been two major quadratic approaches. The local risk minimization sacrifices the self-financing property, but its terminal value is the same as the payoff of a contingent claim. The mean-variance hedging, on the other hand, focuses on the self-financing property. Föllmer and Sondermann [12] studied the risk minimization when the asset price process is a martingale under the original measure, and later, Föllmer and Schweizer [11] and Schweizer [31] studied the local risk minimization for a general semimartingale case. Schweizer [32] provided the solution to the mean-variance hedging for general claims with continuous price processes.

While mean-variance hedging gives a control over the total risk, the local risk minimization often gives a simpler hedging strategy. (See Heath, Platen, and Schweizer [16], for example.) There have been many studies on the above quadratic criteria since they had been proposed. To name a few, Frey [14] studied a risk minimizing strategy when the price process is a pure jump process with a stochastic jump rate and a martingale under the original measure. Chan [7] found a local risk minimizing strategy when the price process is driven by general Lévy processes.
Lim [21] studied a closed form expression of the mean-variance hedging strategy in a specific jump diffusion model, using a backward stochastic differential equations and stochastic optimal control theory. Our suggested model has non-Lévy type jumps as introduced in Frey [14]. Another advantage of our model is that it allows asymmetric return distributions. We can obtain this asymmetry by controlling the jump size distribution \( \nu \), as long as it has mean 0 and a finite second moment. This flexibility of the model gives us a better fit of real stock market data.

The outline for this paper is as follows. In Section 2.1, we introduce our feedback jump model and some technical assumptions. We discuss the Markov property of the model and the minimal martingale measure, dynamics of processes under the changed measure in Section 2.2. We construct a Feynman-Kac type integro-differential equation and show the representation property in Section 2.3. We have our main theorem, which gives us the hedging strategy in Section 2.4. We apply this result to liquidity modelling in Section 3, and we conclude in Section 4.

2. Hedging of Options in an Incomplete Market

2.1. The Model. We consider a market which consists of a risky asset and a riskless asset. For simplicity, we assume that the price of the riskless asset is always 1, which implies that the interest rate is 0. A portfolio \((\xi, \eta)\) is a vector process where \( \xi_t \) is a unit amount of the risky asset at time \( t \) and \( \eta_t \) is a unit amount of the riskless asset at time \( t \). Therefore, the value \( V \) of a portfolio \((\xi, \eta)\) at time \( t \) is given by
\[
V_t = \int_0^t \xi_s dS_s + \eta_t.
\]

We are given a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})\) satisfying the usual conditions where \( T \) is a fixed time. \( \mathbb{P} \) represents the statistical or empirical probability measure. \((\mathcal{F}_t)_{0 \leq t \leq T}\) is a filtration which makes all the processes in the model adapted. We define our price process of a risky asset \( S \) as a solution of the stochastic differential equation:
\[
dS_t = f(S_{t-})dB_t + g(S_{t-})dR_t + h(S_{t-})dt, \quad 0 \leq t \leq T, \tag{2.1}
\]
where \( B \) is a standard Brownian motion,
\[
R_t = \sum_{n=1}^{N_t} U_n, \tag{2.2}
\]
and \( N \) is a doubly stochastic Poisson process with a bounded intensity function \( \lambda(S_{t-}) \), in other words,
\[
N_t - \int_0^t \lambda(S_{s-}) ds = \text{a local martingale under } \mathbb{P}. \tag{2.3}
\]
We notice that \( N_t \) denotes the number of jumps up to time \( t \), and \( U_n \) denotes the size of \( n \)-th jump. For the details of doubly stochastic Poisson processes, readers can consult [4], [35]. Since the intensity of \( N \) is a function of the left continuous version of the current stock price \( S_{t-} \), the jump process \( R \) gets an instantaneous feedback from the most recent stock price.
Here, $U_t$'s are i.i.d. random variables with mean 0 and a finite second moment $\sigma^2$ with density function $\nu(dx)$, and $f, g, h$ are bounded measurable Lipschitz functions. We need the following technical condition to guarantee an existence of the unsigned minimal martingale measure.

$$\frac{h(x)g(x)}{f(x)^2 + g(x)^2 \lambda(x)\sigma^2} < 1$$

(2.4)

for all $x$. Without this condition, the minimal martingale measure may not exist under the current model, since it becomes a signed measure. For more discussion on this issue, readers can consult [33].

It follows from the calculation of the compensator of the random measure $\tilde{p}(dt, dx)$ in Theorem 2.6 that $\tilde{R}$ is a local martingale, and hence by the definition of $\tilde{S}$, one can easily see that $\tilde{S}$ is a special semimartingale with the canonical decomposition

$$\tilde{S}_t = \tilde{S}_0 + \int_0^t f(S_s)dB_s + \int_0^t h(S_s)ds + \int_{\mathbb{R}} g(S_s)\tilde{U}_s \{0 \leq x \leq \lambda(S_s)\} \, dm(ds, dx),$$

(2.6)
and for $N_t$, one can take the process

$$N_t = \int_0^t \int_\mathbb{R} 1_{\{0 \leq x \leq \lambda(s)\}} m(ds, dx). \quad (2.7)$$

Using the previous lemma and a standard iteration method, we can show the Markov property under the original measure $P$. Later, we can show the Markov property still holds under the minimal martingale measure, which will be introduced in Theorem 2.

**Theorem 2.2.** Let $S_t$ be as in equation (2.1). Then $S_t$ is Markov under $P$.

**Proof.** Define $K_t = \int_0^t \int_\mathbb{R} xm(dx, ds)$ and $T'$ be an $\mathcal{F}_t$-stopping time, $T' < \infty$ a.s. Define

$$G^{T'} = \sigma\{K_{T'+u} - K_{T'}, B_{T'+u} - B_{T'}, u > 0\}.$$  

Then $G^{T'}$ is independent of $\mathcal{F}_{T'}$. For $u \geq 0$, define inductively

$$Y^0(x, T', u) = x,$$

$$Y^{n+1}(x, T', u) = x + \int_{T'}^{T'+u} \int_\mathbb{R} f(Y^n(x, T', s)) dB_s + \int_{T'}^{T'+u} h(Y^n(x, T', s)) ds + \int_{T'}^{T'+u} \int_\mathbb{R} U_s g(Y^n(x, T', s)) 1_{\{0 \leq y \leq \lambda(y)\}} m(dy, ds).$$

Then, by using induction and after some standard work, we can easily see that

$$E^x \{h(S(S_0, 0, T' + u))|\mathcal{F}_{T'}\} = E^x \{h(S(S_0, 0, T' + u))|S(S_0, 0, T')\}.$$

for any bounded, Borel function $h$.

Next, we find the minimal martingale measure. The minimal martingale measure can be used for local risk minimization, which will be explained in Section 2.4. We recall the definition of it for readers’ convenience. For more discussion on the minimal martingale measure, readers can consult [11, 33]. A martingale measure $Q \approx P$ is called **minimal** if $Q = P$ on $\mathcal{A}_0$, and if any square-integrable $P$-martingale $L$ that satisfies $(L, M) = 0$ remains a martingale under $Q$, where $M$ is the martingale part of $S$ in the canonical decomposition under $P$.

**Theorem 2.3.** Suppose that $f, g, h, \lambda$ satisfy (2.4). Let us define

$$Z_t = 1 - \int_0^t \left( \frac{Z_s - h(S_s)}{f(S_s -)dB_s + g(S_s -)dR_s} \right) ds.$$  

Then, $Z_t > 0$ and $E(Z_t) = 1$ for all $t \in [0, T]$. Furthermore, $Q$ defined by $\frac{dQ}{dP} = Z_T$ is the unique minimal martingale measure of $S$.

**Proof.** Suppose that there exists a minimal martingale measure and let us denote it by $P^*$. Define $Z_t$ by

$$Z_t = E[\frac{dP^*}{dP}|\mathcal{F}_t].$$  

(2.8)

$M_t = \int_0^t (f(S_s -)dB_s + g(S_s -)dR_s)$ denotes the martingale part of the Doob-Meyer decomposition of $S_t$ and $A_t = \int_0^t h(S_s -) ds$ denotes its predictable part under $P$.  

Under $\mathbb{P}^*$, the Doob-Meyer decomposition of $M_t$ is given by $M_t = S_t - A_t$. But the Girsanov-Meyer theorem (Section 3.8 of [27]) shows that $A_t$ also satisfies

$$-A_t = \int_0^t \frac{1}{Z_s} d\langle M, Z \rangle_s.$$  

Since $\langle M, Z \rangle \ll \langle M, M \rangle = \langle S, S \rangle$, there exists a predictable process $\alpha_t$ such that

$$A_t = \int_0^t \alpha_s d\langle S, S \rangle_s.$$  

On the other hand, there exists some $\beta_t$ such that

$$Z_t = 1 + \int_0^t \beta_s dM_s + L_t,$$  

where $L_t$ is a square integrable martingale under $\mathbb{P}$ orthogonal to $M$. Moreover,

$$-A_t = \int_0^t \frac{1}{Z_s} d\langle M, Z \rangle_s = \int_0^t \frac{1}{Z_s} \beta_s d\langle S, S \rangle_s$$

gives us the relation

$$\alpha_t = -\frac{\beta_t}{Z_t}.$$  

Since $\mathbb{P}^*$ is minimal and $L_t$ is a square integrable martingale under $\mathbb{P}$ orthogonal to $M$, $L$ is a martingale under $\mathbb{P}^*$. We observe that $\langle L, Z \rangle = 0$, since $L$ is both a $\mathbb{P}$ and a $\mathbb{P}^*$ martingale. (see the Lemma on p109, [27] for details.) Now we get

$$\langle L, L \rangle = \langle L, Z \rangle = 0,$$

and

$$Z_t = 1 + \int_0^t Z_s(-\alpha_s) dM_s,$$  

where

$$\alpha_t = -\frac{\beta_t}{Z_t}.$$  

It remains to calculate $\alpha$. We can do it easily by

$$\alpha_s = \frac{dA_s}{d\langle S, S \rangle_s}$$

$$= \frac{h(S_s)}{f(S_s)^2 + g(S_s)^2 \sigma^2 \lambda(S_s)}.$$  

Therefore, we get

$$Z_t = 1 - \int_0^t \frac{Z_s \cdot h(S_s)}{f(S_s)^2 + g(S_s)^2 \sigma^2 \lambda(S_s)} (f(S_s) dB_s + g(S_s) dR_s).$$  

Since there is a unique solution of the equation (2.11), if there exists a minimal martingale measure, it is unique. We can easily check that (2.11) is also a sufficient condition. Let $M'_t$ be a $\mathbb{P}$ martingale such that $\langle M', M \rangle = 0$. What we should show is that $M'_t$ is $\mathbb{P}^*$ martingale, which is true since

$$\langle M', Z \rangle = \langle M', M \rangle = 0.$$  

(2.12)
To show that $Z_t > 0$ and $E(Z_t) = 1$ for all $t$, notice that

$$Z_t = 1 + \int_0^t Z_s - dY_s,$$

and

$$Z_t = \exp(Y_t - \frac{1}{2}[Y,Y]_t) \prod_{0<\Delta Y_s \leq t} (1 + \Delta Y_s) \exp(-\Delta Y_s).$$

By the condition (2.4), we notice that $Z_t > 0$. By (2.11), $Z_t$ is a $\mathbb{P}$-local martingale.

Since $Z_0 = 1$, it suffices to show that $E([Z,Z]_t) < \infty$ to get $E(Z_t) = 1$ for all $t$.

For some constant $M$, we have

$$E([Z,Z]_t) < ME\left(\int_0^t Z_s^2 ds + \int_0^t Z_s^2 dN_s\right)$$

$$= ME\left(\int_0^t Z_s^2 ds + \int_0^t Z_s^2 \lambda(S_s-) ds\right)$$

$$< ME\int_0^t Z_s^2 ds$$

$$< M\int_0^t E(Z_s^2) ds$$

Let $Y_t = -\int_0^t \frac{h(S_s-)}{f(S_s-)^2 + g(S_s-)^2 \lambda(S_s-)} \sigma^2 (f(S_s-) dB_s + g(S_s-) dR_s)$. Since $f, g, h$ are bounded, and from direct calculation of expected values under normal and Poisson distributions, we get

$$E(Z_t^2) < E \exp(2Y_t) < N,$$

for some constant $N$. \hfill \Box

We study the new Brownian motion and compensated jump measure under the minimal martingale measure $Q$ in next two theorems. The next theorem tells us on the new Brownian motion under $Q$ is a shift of the old Brownian motion by some drift term.

**Theorem 2.4.** Let $Q$ be as in Theorem 2. Under $Q$,

$$\tilde{B}_t = B_t + \int_0^t \frac{h(S_s-)}{f(S_s-)^2 + g(S_s-)^2 \lambda(S_s-)} \sigma^2 (f(S_s-) dB_s + g(S_s-) dR_s)$$

is a Brownian Motion.

**Proof.** By the Girsanov-Meyer theorem, we know that

$$N_t = B_t - \int_0^t \frac{1}{Z_s-} d\langle Z, B \rangle_s$$

is a $Q$ local martingale. However,

$$\langle Z, B \rangle_t = -\int_0^t \frac{Z_s- h(S_s-)}{f(S_s-)^2 + g(S_s-)^2 \lambda(S_s-)} \sigma^2 (f(S_s-) dB_s + g(S_s-) dR_s)$$

which gives $N_t = \tilde{B}_t$. Moreover,

$$\langle \tilde{B}, \tilde{B} \rangle_t = \langle B, B \rangle_t = t.$$

Therefore, by Lévy’s theorem (see p.86 of [27]), $\tilde{B}$ is a $\mathbb{Q}$ Brownian motion. $\square$

By $p^R(dt, dx)$ we denote the counting measure associated with the process $R_t$; $p^R(dt, dx)$ is a random measure on $[0, T] \times \mathbb{R}$ such that for functions $W : \Omega \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$,

$$
\int_0^t \int_{\mathbb{R}} W(\omega; s, x)p^R(ds, dx) = \sum_{n=1}^\infty W(\omega; T_n(\omega), U_n(\omega))1_{(T_n(\omega) \leq t)},
$$

when $T_n$ is $n$-th jump time of $N_t$. In the next theorem, we study the change of the compensated measure of the random measure $p^R(dt, dx)$ of $R_t$ under the minimal martingale measure. To do this, we need a lemma. It is a part of Theorem 3.17 of Chapter III (page 157), Jacod and Shiryaev [18] and a version of Girsanov theorem for random measure. We need some definitions. $M^\mu_\sigma$ is the positive measure defined by $M^\mu_\sigma(W) = E(W \ast \mu_\sigma)$ for all measurable nonnegative function $W$, where $W \ast \mu_\sigma = \int_{[0,\infty]} \int_E W(\omega, s, x)\mu(\omega; ds, dx)$ where $\mu$ is a random measure. By the predictable $\sigma$-field $\tilde{\mathcal{F}}$, we mean a $\sigma$-field on $\Omega \times \mathbb{R}$ such that $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{B}$ where $\mathcal{B}$ is a Borel $\sigma$-field. There is a notion of conditional expectation relative to $M^\mu_\sigma$ with respect to the predictable $\sigma$-field $\tilde{\mathcal{F}}$: for every nonnegative measurable function $W$, let $W' = M^\mu_\sigma(W|\tilde{\mathcal{F}})$ denote the $M^\mu_\sigma$-a.s. unique $\tilde{\mathcal{F}}$-measurable function such that $M^\mu_\sigma(WU) = M^\mu_\sigma(W'U)$ for all nonnegative $\tilde{\mathcal{F}}$-measurable $U$. For more details and examples, we refer to Chapter 2 and 3 of Jacod and Shiryaev [18].

**Lemma 2.5.** Assume that $\mathbb{Q} \ll \mathbb{P}$ and let $Z$ be the density process. Let $\mu = \mu(\omega; dt, dx)$ be an integer-valued random measure on $\mathbb{R} \times E$, and denote by $\nu = \nu(\omega; dt, dx)$ its $\mathbb{P}$-compensator of $\mu$. Let $Y$ be any nonnegative version of $M^\mu_\sigma(\frac{Z}{Z|\mathcal{F}_t}1_{\{Z > 0\}}|\tilde{\mathcal{F}})$ and $\nu'$ be a version of the $\mathbb{Q}$ compensator. Then $\nu'(\omega; dt, dx) = Y(\omega; t, x)\nu(\omega; dt, dx)$ $\mathbb{P}$-a.s.

**Theorem 2.6.** Let $\mathbb{Q}$ be as in Theorem 2. Under $\mathbb{Q}$, the compensated measure of $p^R(dt, dx)$ is given by

$$
q^*(dt, dx) = p^R(dt, dx) - (1 - \frac{h(S_{t-})g(S_{t-})}{f(S_{t-})^2 + g(S_{t-})^2 \sigma^2} x)\lambda(S_{t-})\nu(dx)dt.
$$

**Proof.** Since $R_t = \sum_{n=1}^{N_t} U_n$, $N_t$ is a doubly stochastic Poisson process with intensity function $\lambda(S_{t-})$, and $U_n$ has a density $\nu(dx)$, the random measure $p^R(dt, dx)$ admits $(\mathbb{P}, \mathcal{F}_t)$ local characteristics $(\lambda(S_{t-}), \nu(dx))$ (see Definition 5, p236 of Brémaud [4] for a definition of local characteristics). Then, Corollary 15 on page 247 of Brémaud [4] tells us that the compensator of $p^R(dt, dx)$ under $\mathbb{P}$ is $\lambda(S_{t-})\nu(dx)dt$ (p218-219 Frey [14] for details). We refer to Chapter 8 of Brémaud [4] for further discussion on properties of local characteristics. Using the notations of Lemma 2, $\mu(\omega; dt, dx) = p^R(\omega; dt, dx)$ and $\nu(\omega; dt, dx) = \lambda(S_{t-})\nu(dx)dt$. By Lemma 2, to find $M^\mu_\sigma(\frac{Z}{Z|\mathcal{F}_t}1_{\{Z > 0\}}|\tilde{\mathcal{F}})$ is enough. Note that

$$
Z_t = 1 + \int_0^t Z_{s-}dX_s,
$$
where
\[ X_t = -\int_0^t \frac{h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2\lambda(S_{s-})\sigma^2}(f(S_{s-})dB_s + g(S_{s-})dR_s). \]

Therefore,
\[ \frac{Z_t}{Z_{t-}} = 1 + \Delta X_t = 1 - \frac{h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2\lambda(S_{s-})\sigma^2}g(S_{s-})\Delta R_s, \]
and
\[ M^\mu_p(\frac{Z}{Z^-}1_{\{Z > 0\}}) = 1 - \frac{h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2\lambda(S_{s-})\sigma^2}g(S_{s-})M^\mu_p(\Delta R_s|\hat{\mathbb{P}}). \]

Observe that
\[ M^\mu_p(\Delta R_s) = E \int_0^\infty \int_\mathbb{R} x1_{\{\Delta R_s > 0, U_s = x\}}p^R(\omega; ds, dx). \]

For any nonnegative \( \hat{\mathbb{P}} \)-measurable \( W(\omega; t, x) \), since \( p^R(\omega; ds, dx) \) is zero unless \( \Delta R_s > 0, U_s = x \), we get
\[ M^\mu_p(\Delta R_s W) = E \int_0^\infty \int_\mathbb{R} x1_{\{\Delta R_s > 0, U_s = x\}}W(\omega; s, x)p^R(\omega; ds, dx) \]
\[ = E \int_0^\infty \int_\mathbb{R} xW(\omega; s, x)p^R(\omega; ds, dx) \]
\[ = M^\mu_p(xW). \]

Since the function \( (\omega; s, x) \rightarrow x \) is \( \hat{\mathbb{P}} \)-measurable, we get
\[ M^\mu_p(\Delta R_s|\hat{\mathbb{P}}) = x. \]

Therefore, by (2.13),(2.14) and Lemma 2, the compensator of \( p^R \) under the minimal martingale measure \( \mathbb{Q} \) is given by
\[ (1 - \frac{g(S_{\tau^-})h(S_{\tau^-})}{f(S_{\tau^-})^2 + g(S_{\tau^-})^2\lambda(S_{\tau^-})\sigma^2})\lambda(S_{\tau^-})\nu(dx)dt. \]

In Section 2.2, we showed that \( S_t \) is Markov under \( \mathbb{P} \). But what we really need is the Markov property under the changed measure \( \mathbb{Q} \), not under the original measure \( \mathbb{P} \). This Markov property of \( S \) under \( \mathbb{Q} \) follows immediately from Theorem 3 and 4, following analogous steps of the proof of Theorem 1.

2.3. The Integrals-Differential Equation. Let us define \( V_t = E^\mathbb{Q}[H(S_T)|\mathcal{F}_t] \), where \( H \) is a European style contingent claim. By the Markov property of \( S \) under \( \mathbb{Q} \),
\[ V_t = E^\mathbb{Q}[H(S_T)|\mathcal{F}_t] = E^\mathbb{Q}[H(S_T)|S_t] = v(t, S_t), \]
for some function \( v = v(t, x) \).

In order to use Itô’s formula, we need \( v \) to be \( C^{1,2} \). We can define \( v \) by (using standard Markov process notation):
\[ v(t, x) = E^\mathbb{Q}[H(S_{T-t})]. \]
Then using Markov process theory (see, eg, [34]) we have

\[ v(t, x) = E^x[H(S_{T-t})] \]

implies

\[ v(t, S_t) = E^{S_t}[H(S_{T-t})] \]

\[ = E^x[H(S_{T-t}) \circ \theta_t|\mathcal{F}_t] \]

\[ = E^x[H(S_T)|S_t], \]

which is intuitively equal to

\[ = E[H(S_T)|S_t = x]. \]

We now have the following result.

**Theorem 2.7.** Let \( S \) be as in equation (2.1), and we assume that the coefficients \( f, g \) and \( h \), and also \( \lambda \), are all bounded and at least \( bC^2 \), and that their second derivatives are Lipschitz continuous. We further assume that \( H \) is bounded, and also in \( bC^2 \). Then \( v = v(t, x) \) as defined in equation (2.15) above, is \( C^{1,2} \).

**Proof.** This theorem follows essentially from the theory of flows of stochastic differential equations, as presented (for example) in [27]. Nevertheless, because of the presence of the feedback term, we give its proof. Our candidate first derivative of \( v \) in \( x \) is given by

\[ D_t = 1 + \int_0^t f'(S_u^x)D_u dB_u + \int_0^t g'(S_{u-}^x)D_u dR_u + \int_0^t h'(S_u^x)D_u du. \quad (2.16) \]

where \( S_u^x \) denotes \( S \) starting at the point \( x \), at time \( t \). Note that there is no problem with the existence of \( D_t \): it is a well defined stochastic exponential of the semimartingale \( S \). Since the coefficients are globally Lipschitz, we have that \( x \mapsto S_u^x \) is \( C^2 \) in \( x \) by Theorem 40, on page 310, of [27].

Using the above, we now have by \( H \in bC^2 \) and dominated convergence that

\[ \frac{\partial}{\partial x} E[H(S_t^x)] = E[\frac{\partial}{\partial x} H(S_t^x)] \]

\[ = E[H'(S_t^x) \frac{\partial}{\partial x} S_t^x] = E[H'(S_t^x) D_t^x] \]

To repeat the argument to get \( C^2 \) in \( x \) we need to control \( D_t^x \) in order to use dominated convergence, since we are assuming that \( H' \) is bounded. However \( D_t^x \) is the stochastic exponential of a semimartingale, so this follows by Theorem 55, page 326, of [27]. Therefore repeating the argument one more time yields that \( x \mapsto v(t, x) \) is \( C^2 \) in \( x \). It remains to show that \( t \mapsto v(t, x) \) is \( C^1 \). To simplify notation, we take \( x = 0 \) and suppress the \( x \) variable. Thus we want to show that \( t \mapsto E_Q[H(S_t)] \) is \( C^1 \), where \( Q \) is a risk neutral measure for \( S \), such that \( S \) is Markov and a true martingale (and not only a local martingale); we have already constructed such a risk neutral measure \( Q \). But this follows from Theorem 2.6 and our hypotheses on \( h, g, f \) and \( \lambda \), since we know the form of the compensation of \( p^R(dt, dx) \).

\[ \square \]

Note that our hypotheses are not best possible in Theorem 2.7 in order for \( v \) to be \( C^{1,2} \). For example, let us consider the linear function \( H(x) = ax + b \). Then we get \( v(t, x) = ax + b \) since \( S_t \) is a martingale under \( Q \) and clearly it is \( C^{1,2} \).
When \( v \) is \( C^{1,2} \), by Itô's formula,

\[
v(t, S_t) = v(0, S_0) + \int_0^t v_t(s, S_{s-}) ds + \int_0^t f(S_{s-}) v_x(s, S_{s-}) d\tilde{B}_s
\]

\[
+ \int_0^t \int_{\mathbb{R}} \{ v(s, S_{s-} (1 + x \frac{g(S_{s-})}{S_{s-}})) - v(s, S_{s-}) \} q^*(dx, ds)
\]

\[
+ \int_0^t v_x(s, S_{s-}) h(S_{s-}) ds + \frac{1}{2} \int_0^t v_{xx}(s, S_{s-}) f(S_{s-})^2 ds
\]

\[
+ \int_0^t f(S_{s-}) v_x(s, S_{s-}) \frac{h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2 \lambda(S_{s-}) \sigma^2} f(S_{s-}) ds
\]

\[
- \int_0^t \int_{\mathbb{R}} \{ v(s, S_{s-} (1 + x \frac{g(S_{s-})}{S_{s-}})) - v(s, S_{s-}) \}
\]

\[
\times (1 - \frac{g(S_{s-}) h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2 \lambda(S_{s-}) \sigma^2}) x \lambda(S_{s-}) \nu(dx) ds.
\]  

(2.17)

Since \( v(t, S_t) \) is a \( \mathbb{Q} \) - martingale, the right side of (2.17) also should be a \( \mathbb{Q} \) - martingale. Therefore, since a continuous process with finite variation can be a martingale only if it is constant, we need following conditions:

\[
v_t(s, S_{s-}) + v_x(s, S_{s-}) h(S_{s-}) + \frac{1}{2} v_{xx}(s, S_{s-}) f(S_{s-})^2
\]

\[
- f(S_{s-}) v_x(s, S_{s-}) \frac{h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2 \lambda(S_{s-}) \sigma^2} f(S_{s-})
\]

\[
+ \int_{\mathbb{R}} \{ v(s, S_{s-} (1 + x \frac{g(S_{s-})}{S_{s-}})) - v(s, S_{s-}) \}
\]

\[
\times (1 - \frac{g(S_{s-}) h(S_{s-})}{f(S_{s-})^2 + g(S_{s-})^2 \lambda(S_{s-}) \sigma^2}) x \lambda(S_{s-}) \nu(dx) = 0
\]  

(2.18)

for almost all \( s \), a.s. and

\[
v(T, S_T) = H(S_T).
\]  

(2.19)

This gives us an integro-differential equation which has a form of Feynman-Kac type differentiation operator plus an integration part, which arises from the jumps. We showed that when

\[
v(t, x) = E_Q[H(S_T)|S_t = x],
\]

where \( S \) is in (2.1) and \( v \) is \( C^{1,2} \), then \( v \) is a solution of the integro-differential equation (2.18, 2.19).

We can show the opposite direction as well. If a solution of the integro-differential equation (2.18, 2.19) exists, it has a stochastic representation

\[
v(t, x) = E_Q^{t,x}(H(S_T)),
\]
where \( S \) is given by the solution of (2.1). To see this, we just do the same thing backward. From (2.17), Itô’s formula applied to \( v(T, S_T) \) gives us

\[
v(T, S_T) = v(t, S_t) + \int_t^T v_t(s, S_s) \, ds + \int_t^T \mathcal{A}v(s, S_s) \, ds
\]

\[
+ \int_t^T f(S_s) \, v_x(s, S_s) \, d\tilde{B}_s
\]

\[
+ \int_t^T \int_{\mathbb{R}} \left[ v(s, S_s - (1 + x \frac{g(S_s - \lambda \sigma)}{S_s - \lambda \sigma})) - v(s, S_s - \lambda \sigma) \right] q^*(dx, ds).
\]

By (2.21) and martingale properties of the last two terms, we get

\[
v(t, x) = E^t_x[v(T, S_T)] = E^t_x(H(S_T)). \tag{2.20}
\]

Let us assume that \( v \) is \( C^{1,2} \). Then we have \( v \) satisfies (2.18, 2.19) and any other solution of (2.18, 2.19) must have the stochastic representation \( v(t, x) = E^t_x(H(S_T)) \). Therefore, \( v \) is the unique classical solution of the integro-differential equation (2.18, 2.19).

If we weaken the hypotheses of Theorem 2.7 so that we do not know that \( v \) is \( C^{1,2} \), then we do not know if the solution exists. But, we know that any solution of (2.18, 2.19) must have the stochastic representation \( v(t, x) = E^t_x(H(S_T)) \). Therefore, we have \( v(t, x) = E^t_x(H(S_T)) \) as an \textit{a priori} estimator of the integro-differential equation (2.18, 2.19).

Note that we can denote the integro-differential equation (2.18, 2.19) using an infinitesimal generator of the Markov process \( S \). Let \( \mathcal{A} \) denote the infinitesimal generator of the Markov process \( S_t \).

\[
\mathcal{A}v(x) = \lim_{t \to 0} \frac{E^t_x[v(S_t)] - v(S_0)}{t}
\]

\[
= \lim_{t \to 0} \frac{E^t_x(v(S_t)) - v(x)}{t}
\]

\[
= d \frac{d}{dt} E^x_t(v(S_t))|_{t=0}.
\]

By Itô’s formula applied to \( v(S_t) \),

\[
v(S_t) = v(S_0) + \int_0^t f(S_s) \, v_x(S_s) \, d\tilde{B}_s
\]

\[
+ \int_0^t \int_{\mathbb{R}} \left[ v(S_s - (1 + x \frac{g(S_s - \lambda \sigma)}{S_s - \lambda \sigma})) - v(S_s - \lambda \sigma) \right] q^*(dx, ds)
\]

\[
+ \int_0^t v_x(S_s - \lambda \sigma) h(S_s - \lambda \sigma) \, ds + \frac{1}{2} \int_0^t v_x x(S_s) f(S_s - \lambda \sigma)^2 \, ds
\]

\[
+ \int_0^t f(S_s) v_x(S_s - \lambda \sigma) \frac{h(S_s - \lambda \sigma)}{f(S_s - \lambda \sigma)^2 + g(S_s - \lambda \sigma)^2} \, ds
\]
- \int_0^t \{ v(s, S_s - (1 + x \frac{g(S_s - )}{S_s - }) - v(S_s - )) \\
\times (1 - \frac{g(S_s - ) h(S_s - )}{f(S_s - )^2 + g(S_s - )^2 \lambda(S_s - ) \sigma^2} x) \lambda(S_s - ) \nu(dx)ds. \)
Then we take an expectation under Q and get
\[ \frac{d}{dt} E_Q^x(v(S_t))|_{t=0} = v'(x) h(x) + \frac{1}{2} v''(x) f(x)^2 \]
\[ - v'(x) f(x)^2 \frac{h(x)}{f(x)^2 + g(x)^2 \lambda(x) \sigma^2} \]
\[ + \int_\mathbb{R} \{ v(x(1 + z \frac{g(x)}{x})) - v(x) \} \]
\[ \times (1 - \frac{g(x) h(x)}{f(x)^2 + g(x)^2 \lambda(x) \sigma^2} z) \lambda(x) \nu(dz) \]
\[ = Av(x). \]
Now, the previous integro-differential equation can be written in the form
\[ -v_t = Av, \quad (2.21) \]
where \( v_t \) is a partial derivative and \( v(T, x) = H(x) \).

2.4. The Explicit Hedging Strategy. For any portfolio \((\xi, \eta)\), the cost process \( C \) is defined by \( C_t = V_t - \int_0^t \xi_s dS_s, (0 \leq t \leq T) \). If \( C \) is positive, the value of the portfolio is bigger than the cumulative gain from the portfolio, which means we have to inject some money to keep the portfolio. If it is negative, we can withdraw some money, since we have some overflows. If the portfolio is self-financing, we notice that \( C \) is always constant, which means there is no cash-flow after the initial payment. By the local risk minimization strategy, we mean a portfolio whose cost process is a square integrable martingale orthogonal to \( M \), where \( M \) is the martingale part of \( S \) under \( P \). Readers can consult [11, 31, 16] for more detailed discussion on the local risk minimization. Föllmer and Schweizer [11] suggested useful sufficient conditions for the existence of the local risk minimization strategy which are easier to calculate. They found these conditions when the price process is a continuous semimartingale. Although our model is not continuous, we can show that their result still works as long as the price process is \( H^2 \) semimartingale.

Suppose that \( H(S_T) \) is our contingent claim such as a European call option or put option, and \( M \) denotes the martingale part of \( S \) under \( P \). We also assume that \( V_t = E_Q[H(S_T)|\mathcal{G}_t] \) has a decomposition
\[ V_t = V_0 + \int_0^t \xi_s^H dS_s + L_t, \quad (2.22) \]
where \( L \) is a square-integrable \( P \)-martingale that is orthogonal to \( M \) under \( P \), in other words, \( \langle L, M \rangle = 0 \). For a given hedging strategy to be the local risk minimization strategy, its cost process should be a square-integrable martingale that is orthogonal to \( M \) under \( P \). In other words, \( \xi^H \) in the above decomposition (2.22) is, in fact, the local risk minimization strategy, since \( L_t + V_0 \) becomes the cost process and is a square-integrable \( \mathbb{P} \)-martingale that is orthogonal to \( M \) under
\( \xi^H \) can be computed further in the following way. From \( V_t = V_0 + \int_0^t \xi^H_s \, dS_s + L_t \), we get
\[
\langle V, S \rangle_t = \int_0^t \xi^H_s \, d\langle S, S \rangle_s + \langle L, S \rangle_t.
\]
Since \( L \) is orthogonal to \( M \) under \( \mathbb{P} \),
\[
\langle L, S \rangle_t = \langle L, M \rangle_t + \langle L, A \rangle_t = 0.
\]
Therefore, we have \( \langle V, S \rangle_t = \int_0^t \xi^H_s \, d\langle S, S \rangle_s \) and thus,
\[
\xi^H_t = \frac{d\langle V, S \rangle_t}{d\langle S, S \rangle_t}.
\] (2.23)
where the quadratic variations are calculated under \( \mathbb{P} \). For more details, we refer Föllmer and Schweizer [11], Chan [7]. Thus, as long as we know the existence of the decomposition of \( V \), we can calculate the closed form local risk minimization strategy from the equation (2.23).

Notice that the quadratic variations are calculated under \( \mathbb{P} \). In Föllmer and Schweizer [11], we did not need to specify a measure under which the bracket processes, \( \langle \cdot, \cdot \rangle \), are calculated, since the bracket processes are invariant under a change of measure if processes are continuous semimartingales. If processes are discontinuous, in general, the bracket processes are different if we calculated them under different measures. For the definition of the bracket processes when processes are discontinuous semimartingales and properties, readers can consult Section 3.5 of [27].

First, we assume the decomposition (2.22) exists. Later, in Theorem 6, we can show that this decomposition always exists under our model. The following is the main theorem of our paper.

**Theorem 2.8.** If the decomposition of \( V_t = E_{\mathbb{Q}}[H(S_T)|\mathcal{F}_t] \) exists, the locally risk minimizing strategy is given by
\[
\xi^H_t = \frac{j(t, S_{t-})}{f(S_{t-})^2 + g(S_{t-})^2 \sigma^2 \lambda(S_{t-})},
\] (2.24)
where
\[
j(t, S_{t-}) = \lambda(S_t) \int_{\mathbb{R}} \{v(t, S_{t-}(1 + g(S_{t-})) - v(t, S_{t-})) + v(t, S_{t-}) \} x \nu(dx) g(S_{t-})
\]
\[+ v_x(t, S_{t-}) f(S_{t-})^2.
\]

**Proof.** Recall that the optimal strategy is given by
\[
\xi^H_t = \frac{d\langle V, S \rangle_t}{d\langle S, S \rangle_t}.
\]
We get
\[
d\langle R, R \rangle_t = \sigma^2 \lambda(S_{t-}) dt,
\] (2.25)
and
\[
d\langle S, S \rangle_t = f(S_{t-})^2 dt + g(S_{t-})^2 \sigma^2 \lambda(S_{t-}) dt.
\] (2.26)
Observe that from the equation (2.17), then we have

\[ V_t = v(t, S_t) = v(0, S_0) + \int_0^t v_t(s, S_{s-}) ds + \int_0^t f(S_{s-})v_x(s, S_{s-}) d\tilde{B}_s \]
\[ + \int_0^t \int_\mathbb{R} \{ v(s, S_{s-}(1 + xg(S_{s-}))) - v(s, S_{s-})\} q^*(dx, ds). \]

From this, we can calculate \( d\langle V, R \rangle_t \) and \( d\langle V, B \rangle_t \).

\[ d\langle V, R \rangle_t = dt\lambda(S_t) \int_\mathbb{R} \{ v(t, S_{t-}(1 + xg(S_{t-}))) - v(t, S_{t-})\} x\nu(dx), \]
\[ d\langle V, B \rangle_t = v_x(t, S_{t-}) f(S_{t-}) dt. \]

From (2.26) and

\[ d\langle V, S \rangle_t = g(S_{t-}) d\langle V, R \rangle_t + f(S_{t-}) d\langle V, B \rangle_t, \]
we get the result. \( \square \)

In Theorem 5, we assumed the existence of decomposition of \( V_t = E^Q[H(S_T)|\mathcal{F}_t] \). To guarantee the existence of an optimal strategy, we need to show the existence of the decomposition of \( V_t = E^Q[H(S_T)|\mathcal{F}_t] \). We can construct the decomposition in the following way.

**Theorem 2.9.** Let \( M \) be the martingale part of \( S_t \) and \( \zeta^H_s \) be as in Theorem 5. Then, \( V_t = E^Q[H(S_T)|\mathcal{F}_t] \) has a decomposition

\[ V_t = V_0 + \int_0^t \zeta^H_s dS_s + L_t, \]

where \( L_t \) is a square integrable \( \mathbb{P} \)-martingale such that \( (L, M)_t = 0 \) under \( \mathbb{P} \). In other words, there exists an optimal strategy.

**Proof.** The theorem follows from the next two lemmas (Lemma 2.10 and Lemma 2.11). \( \square \)

**Lemma 2.10.** \( L_t = V_t - \int_0^t \zeta^H_s dS_s \) is a square integrable \( \mathbb{P} \)-martingale.

**Proof.** Since \( H_t \) and \( \int_0^t \zeta^H_s dS_s \) are \( \mathbb{Q} \)-local martingales, \( L_t \) is a \( \mathbb{Q} \)-local martingale.

Let \( L_t = M'_t + A'_t \) be the canonical decomposition of \( L_t \) under \( \mathbb{P} \) where \( M'_t \) is the local martingale part and \( A'_t \) is the predictable part. By the Girsanov-Meyer theorem,

\[ M'_t - \int_0^t \frac{1}{Z_s} d\langle Z, M'_\rangle_s = L_t, \quad (2.27) \]
where $Z_t$ is the density process as in Theorem 2. By the uniqueness of the decomposition, there is only one $M'_t$ which satisfies (2.27). On the other hand,

$$
\int_0^t \frac{1}{Z_{s^-}} d\langle Z, L \rangle_t
= \int_0^t \frac{1}{Z_{s^-}} (d\langle Z, V \rangle_s - \xi^H_s d\langle Z, S \rangle_s)
= \int_0^t \left( \frac{h(S_{s^-})}{f(S_{s^-})^2 + g(S_{s^-})^2 \lambda(S_{s^-}) \sigma^2} \right)^2 (d\langle S, V \rangle_s - \xi^H_s d\langle S, S \rangle_s)
= 0,
$$

which implies $M'_t = L$, i.e. $L$ is a $\mathbb{P}$ local martingale. To show that $L$ is a square integrable $\mathbb{P}$ martingale, it is enough to show that $E\langle L, L \rangle_T < \infty$. Notice that

$$
\langle L, L \rangle_T = \langle V - \int_0^T \xi^H_s dS_s, V - \int_0^T \xi^H_s dS_s \rangle_T
= \langle V, V \rangle_T - 2 \int_0^T \xi^H_s d\langle S, V \rangle_s + \int_0^T (\xi^H_s)^2 d\langle S, S \rangle_s
= \langle V, V \rangle_T - \int_0^T (\xi^H_s)^2 d\langle S, S \rangle_s
= \langle V, V \rangle_T - \int_0^T (\xi^H_s)^2 (f(S_{s^-})^2 + g(S_{s^-})^2 \sigma^2 \lambda(S_{s^-})) ds,
$$

and

$$
E\langle L, L \rangle_T = EV^2_T - \int_0^T E((\xi^H_s)^2 (f(S_{s^-})^2 + g(S_{s^-})^2 \sigma^2 \lambda(S_{s^-}))) ds < \infty.
$$

\[ \square \]

**Lemma 2.11.** $\langle L, M \rangle_t = 0$ under $\mathbb{P}$.

**Proof.**

$$
\langle L, M \rangle_t = \langle V - \int_0^t \xi^H_s dS_s, \int_0^t (f(S_{s^-})dB_s + g(S_{s^-})dR_s) \rangle_t
= \langle V, \int_0^t (f(S_{s^-})dB_s + g(S_{s^-})dR_s) \rangle_t
- \int_0^t \xi^H_s f(S_{s^-}) d\langle S, B \rangle_s - \int_0^t \xi^H_s g(S_{s^-}) d\langle S, R \rangle_s
= \int_0^t f(S_{s^-}) d\langle V, B \rangle_s + \int_0^t g(S_{s^-}) d\langle V, R \rangle_s
- \int_0^t \xi^H_s f(S_{s^-}) d\langle S, B \rangle_s - \int_0^t \xi^H_s g(S_{s^-}) d\langle S, R \rangle_s
= \langle S, V \rangle_t - \int_0^t \xi^H_s f(S_{s^-}) d\langle S, B \rangle_s - \int_0^t \xi^H_s g(S_{s^-}) d\langle S, R \rangle_s. \tag{2.29}
$$
Notice that
\[ f(S_s -)d(S,B)_s + g(S_s -)d(S,R)_s = f(S_s -)^2 ds + g(S_s -)^2 \sigma^2 \lambda(S_s -)ds \]
\[ = d\langle S,S \rangle_s. \]

Therefore,
\[ \int_0^t \xi^H_t(f(S_s -)d(S,B)_s + g(S_s -)d(S,R)_s) = \int_0^t d\langle S,V \rangle_s = \langle S,V \rangle_t \quad (2.30) \]

Now, from (2.29) and (2.30), we get
\[ \langle L,M \rangle_t = \langle S,V \rangle_t - \langle S,V \rangle_t = 0. \]

\[ \Box \]

3. Application to Liquidity Modelling

We discuss a possible application of this type of price process model to the topic of liquidity. In recent liquidity research (see for example [1], [5], [6], [19], [20], [30]) one uses the notion of a supply curve, which is a mathematical model of a limit book for the purchase of relatively liquid stocks. This means as demand increases very rapidly, one must “climb” the ladder of the limit order book to fill the orders, as demand temporarily exceeds ready supply. Therefore one sees small jumps upward during this period, the size of the jump depending on the price and the structure of the order book. In empirical/modelling work by Marcel Blais in his PhD thesis [2], for liquid stocks the supply curve takes the form:

\[ S(t, x) = \begin{cases} M^+_t x + b(t), & x \geq 0 \\ M^-_t x + a(t), & x < 0 \end{cases}, \]

where \( t \) is time, \( x \) is the size of an initiated trade (\( x > 0 \) is a buy, and \( x < 0 \) is a sell, both of \( x \) shares of stock), and \( S(t, x) \) is the price paid or received per share if the order is of size \( x \). \( M^+_t \) and \( M^-_t \) represent the slopes of what turned out to be, as shown by fitting the data to cubic splines, a piecewise linear supply curve. Actually, in many cases for liquid stocks, it turned out that \( M^+_t = M^-_t \); We then simplify notation by writing \( M_t \); that is, the supply curve is linear, with a random and time varying slope (although Blais’ study indicates that the variation of \( M_t \) is small, and continuous in \( t \)). Also, often \( a(0) = b(0) \), which means that the bid-ask spread is negligible. In this case, it might be reasonable to represent the price process as the solution to the following SDE:

\[ dS_t = \sigma(S_t)dB_t + b(S_{t-}, M_s)d(N_t - \int_0^t \lambda(S_{s-}, M_s)ds) + \mu(S_t)dt, \quad (3.1) \]

which is a variant of equation (1.1). We see in equation (3.1) that the intensity of jumps changes with the price, which reflects the climb up the order book, and the typical structure that as one climbs, the offered prices become more sparse and more largely spaced; and the coefficient \( b \) depending on \( M \) represents the slope of the supply curve affecting the size of the small jump changes in price, with a steeper slope representing an increased inelasticity, and thus one would expect \( b(s, m) \) to be increasing in \( m \).
Indeed, one can imagine other applications of this type of modelling, when changes in small jumps are sensitive to price, both for their frequency and their size, a phenomenon which happens often in financial modelling.

4. Conclusion

This paper introduces a method to find an explicit form of the local risk minimization strategy in general class of models with jumps, extending the results of Frey [14] and Chan [7]. It gives flexibility of models, by changing an intensity function $\lambda$ or a jump distribution $\nu(dx)$. It can also be applied to the deterministic intensity case, simply choosing $\lambda$ independent of the price process. The case when the price process is driven by a Lévy process, which was studied by Chan [7], also follows using a Lévy decomposition and a specific choice of jumps. The most original aspect of our model is an instantaneous feedback of the current price. It would be ideal if we allow an intensity depending on the whole path of the stock history, which is $\lambda(S_s, 0 \leq s \leq t)$ instead of $\lambda(S_t)$, but we defer this to future research.

One problem of $\lambda(S_s, 0 \leq s \leq t)$ is the loss of the Markov property, which plays a key role in our approach. This result may be extended to more general jump diffusion cases. But in general this method may fail since one cannot in general guarantee the existence of the decomposition (2.22) and failure of the martingale representation theorem. Our model also allows asymmetric return distributions, which is another advantage. The stochastic representation of the Feynman-Kac type integro-differential equation naturally suggests to us a Monte Carlo method to obtain a solution, but this too will be addressed in the future.

References

28. Philip Protter: School of OR&IE, Cornell University, Ithaca, NY 14853, USA
E-mail address: pep@cornell.edu
ABSTRACT WIENER SPACE, REVISITED

DANIEL W. STROOCK

Abstract. This note contains some of my ruminations about L. Gross’s theory of abstract Wiener space. None of the ideas introduced or conclusions drawn here is new. Instead, this is only my interpretation of a couple of the beautiful ideas and conclusions which appeared in Gross’s seminal 1965 article [1].

1. The Basic Idea

Consider a separable, real Hilbert space \( H \). When \( H \) is finite dimensional, the standard, Gaussian measure \( W_H \) for \( H \) is the Borel measure given by

\[
W_H(dh) = \left(2\pi\right)^{-\frac{\dim(H)}{2}} e^{-\frac{|h|_H^2}{2}} \lambda_H(dh),
\]

(1.1)

where \( \lambda_H \) denotes the Lebesgue measure (i.e., the translation invariant measure which assigns measure 1 to a unit cube in \( H \)). When \( H \) is infinite dimensional, the \( W_H \) is also given by (1.1), only it fails to exist. The reason it fails to exist is well-known: if it did, then, for any orthonormal basis \( \{h_m : m \geq 0\} \), the random variables \( h \in H \mapsto X_m(h) = \langle h, h_m \rangle_H \) would be independent, standard normal random variables and therefore, by the strong law of large numbers, \( \|h\|^2 = \sum_{m=0}^{\infty} X_m(h)^2 \) would be infinite for \( W_H \)-almost every \( h \).

Put another way, \( H \) is simply too small to accommodate \( W_H \). The idea introduced by Gross was to overcome this problem by completing \( H \) with respect to a more forgiving norm than \( \| \cdot \|_H \) in such a way that the resulting Banach space would be large enough to house \( W_H \). To make this precise, he defined the triple \((H, \Theta, W_H)\) to be an abstract Wiener space if \( \Theta \) is a separable Banach space into which \( H \) is continuously embedded as a dense subset and \( W_H \) is the Borel measure on \( \Theta \) which has the “right” Fourier transform, the one which (1.1) predicts it should. That is, because \( H \) is continuously embedded as a dense subspace of \( \Theta \), its dual space \( \Theta^* \) can be continuously embedded as a dense subspace of \( H \). Namely, given \( \lambda \in \Theta^* \), one can use the Riesz representation theorem for Hilbert space to determine \( h_\lambda \in H \) by the relation \( \langle h, h_\lambda \rangle_H = \langle h, \lambda \rangle \), \( h \in H \). Then, because (1.1) predicts that \( h \sim (h, h_\lambda)_H \) should be a centered normal with variance \( \|h_\lambda\|^2_H \), having the “right” Fourier transform means that

\[
\hat{W}_H(\lambda) = \int_{\Theta} e^{i\langle \theta, \lambda \rangle} W_H(d\theta) = e^{-\frac{|h_\lambda|_H^2}{2}}.
\]

(1.2)

\[\text{2000 Mathematics Subject Classification. Primary 60G15, 60G17.}\]
\[\text{Key words and phrases. Abstract Wiener space.}\]
Putting aside for moment the problem of constructing $\Theta$, it is important to observe that, except in finite dimensions, there are a myriad ways of choosing $\Theta$ for the same $H$, and there seems to be no canonical choice. My own way of thinking about this situation is to interpret $H$ as a scaffold onto which one has to put a coating before it is habitable. How thick to make the coat is a matter of taste.

2. Wiener Series

Of course, the reason why Gross chose the term “abstract Wiener space” is that N. Wiener’s construction of Brownian motion can be viewed as the original case in which a satisfactory $\Theta$ was found for a particular $H$. In Wiener’s case, $H$ is the Hilbert space of absolutely continuous $h : [0, 1] \to \mathbb{R}$ such that $h(0) = 0$ and $\dot{h} \in L^2([0, 1]; \mathbb{R})$ with the norm $\|h\|_H = \|\dot{h}\|_{L^2([0, 1]; \mathbb{R})}$. One way to describe how Wiener went about one of his three constructions is to say that he chose an orthonormal basis $\{h_m : m \geq 0\}$ for this $H$ and then considered the random series

$$\sum_{m=0}^{\infty} X_m h_m, \quad (2.1)$$

where $\{X_m : m \geq 0\}$ are independent, standard normal random variables. In terms to Gross’s idea, this is an entirely natural idea. Indeed, if $W_H$ lived on $H$, then

$$h = \sum_{m=0}^{\infty} (h, h_m)_H h_m,$$

and the random variables $h \sim (h, h_m)_H$ would be independent, standard normals.

Starting from (2.1), the problem of constructing a $\Theta$ for Gross’s triple becomes that of finding a Banach space in which the series in (2.1) converges almost surely. To see that this is in fact exactly the same problem, note that if the series is almost surely convergent in $\Theta$, then we can take $W_H = F^* \gamma^N$, where $\gamma$ is the standard normal distribution on $\mathbb{R}$ and $F : \mathbb{R}^N \to \Theta$ is defined by

$$F(x) = \begin{cases} \sum_{m=0}^{\infty} x_m h_m & \text{when the series converges}, \\ 0 & \text{otherwise}, \end{cases}$$

for $x = (x_0, \ldots, x_m, \ldots) \in \mathbb{R}^N$. Checking that this $W_H$ has the right Fourier transform is trivial. Namely, if $\lambda \in \Theta^*$, then, because the convergence is in $\Theta$,

$$\mathcal{W}_H(\lambda) = \lim_{n \to \infty} \mathbb{E}^\lambda \left[\exp \left(\sqrt{-1} \sum_{m=0}^{n} (h_m, h_\lambda)_H X_m\right)\right]$$

$$= \exp \left(-\frac{1}{2} \sum_{m=0}^{\infty} (h_m, h_\lambda)_H^2\right)$$

$$= e^{-\|h_\lambda\|_H^2}.$$

Conversely, if $(H, \Theta, W_H)$ is an abstract Wiener space, then the series in (2.1) must be almost surely convergent in $\Theta$. One way to show this is to take the following steps. First, using separability, check that the Borel field $\mathcal{B}_\Theta$ for $\Theta$ is the
smallest $\sigma$-algebra with respect to which all the maps $\theta \mapsto (\theta, \lambda)$ are measurable. Second, introduce the Paley–Wiener map $h \in H \mapsto I(h) \in L^2(\mathcal{W}_H; \mathbb{R})$ which is obtained by extending to all of $H$ the isometry given by $\|I(h_\lambda)\|_\theta = (\theta, \lambda)$ for $\lambda \in \Theta^*$. Third, note that the span of $\{I(h_m) : m \geq 0\}$ is $L^2(\mathcal{W}_H; \mathbb{R})$-dense in $\{I(h) : h \in H\}$, and use this together with the first step to check that the Borel field $\mathcal{B}_\Theta$ is contained in the $\mathcal{W}_H$-completion of $\sigma(\bigcup_0^\infty \mathcal{F}_n)$, where $\mathcal{F}_n$ is the $\sigma$-algebra generated by $\{I(h_m) : 0 \leq m \leq n\}$. Fourth, use the fundamental property of Gaussian families to see that $\theta \mapsto S_n(\theta) \equiv n \sum_{m=0}^n [I(h_m)](\theta)h_m \tag{2.2}$ is a $\mathcal{W}_H$-conditional expectation value of $\theta$ given $\mathcal{F}_n$. Finally, apply the Banach space version of the Marcinkiewitz convergence theorem (Doob’s martingale convergence theorem for martingales of the form $\mathbb{E}[X | \mathcal{F}_n]$) to conclude that $\theta = \lim_{n \to \infty} S_n(\theta)$ for $\mathcal{W}_H$-almost every $\theta$.

3. Making Wiener’s Series Converge

If one takes a cavalier attitude toward the space at which one will arrive, the remarks in Section 2 make it easy to construct a $\Theta$. For instance, refer to (2.1) and take $\Theta$ to be the completion of $H$ with respect to the Hilbert norm

$$
\|h\|_\Theta = \sqrt{\sum_{m=0}^\infty (1 + m)^{-2}(h, h_m)^2_H}.
$$

Because

$$
\sum_{m=0}^\infty \frac{X_m^2}{(1 + m)^2} < \infty \quad \text{almost surely},
$$

it is trivial to check that the series in (2.1) is almost surely convergent in $\Theta$. Of course, the problem with this approach is that it ignores all subtle cancellation properties and therefore leads to less than optimal results. For example, consider Wiener’s case, and take, as he did,

$$
h_0(t) = t \quad \text{and} \quad h_m(t) = \frac{2^\frac{1}{2} \sin(m\pi t)}{m\pi} \text{ for } m \geq 1.
$$

It is then easy to identify the $\Theta$ to which the above procedure leads as $L^2([0,1]; \mathbb{R})$, which is not the pathspace in which one wants Brownian paths to find themselves. One might hope to improve matters by taking

$$
\|h\|_\Theta = \sum_{m=0}^\infty (1 + m)^{2(1-\alpha)}(h, h_m)^2_H.
$$

As long as $\alpha \in (0, \frac{1}{2})$, there is no doubt that the Wiener series converges in the corresponding $\Theta$. Moreover, in Wiener’s case with his choice of basis, one can identify to resulting $\Theta$ as the Sobolev space of $L^2$-functions whose $\alpha$th order derivative is in $L^2$. Unfortunately, this is again not sufficient to get Brownian paths to be continuous, since Sobolev’s embedding theory for functions on $[0,1]$ requires square integrable $\alpha$th order derivatives for some $\alpha > \frac{1}{2}$. As we now know,
there are many ways to circumvent this difficulty. One way is to abandon Sobolev in favor of his student Besov. That is, for $p \in (1, \infty)$ and $\beta > 0$, define

$$\|h\|_{p,\beta} = \left(\int\int_{[0,1]^2} \left(\frac{|h(t) - h(s)|}{|t - s|^\beta}\right)^p dsdt\right)^{\frac{1}{p}}.$$

By Doob’s inequality for Banach space valued martingales

$$\mathbb{E}^P \left[ \sup_{n \geq 0} \|S_n\|_{p,\beta}^p \right]^{\frac{1}{p}} \leq \frac{p}{p - 1} \sup_{n \geq 0} \mathbb{E}^P \left[ \|S_n\|_{p,\beta}^p \right]^{\frac{1}{p}},$$

where $S_n$ denotes the nth partial sum of the series in (2.1). At the same time, because the $S_n(t) - S_n(s)$ is a centered Gaussian, if $h_{s,t}(\tau) = t \wedge \tau - s \wedge \tau$, then

$$\mathbb{E}^P \left[ |S_n(t) - S_n(s)|^p \right] = K_p \mathbb{E}^P \left[ |S_n(t) - S_n(s)|^2 \right]^\frac{p}{2} \leq K_p \left( \sum_{m=0}^\infty (h_{s,t}, h_m)^2_H \right)^\frac{p}{2},$$

and

$$\sum_{m=0}^\infty (h_{s,t}, h_m)^2_H = \|h_{s,t}\|_H^2 = |t - s|.$$

Thus,

$$\mathbb{E}^P \left[ \|S_n\|_{p,\beta}^p \right] \leq K_p \int_{[0,1]^2} |t - s|^{p(\frac{1}{2} - \beta)} dsdt \equiv K_{p,\beta}.$$  

Since $K_{p,\beta} < \infty$ whenever $\beta < \frac{1}{2} + \frac{1}{p}$, it follows that, for each $\beta \in \left(0, \frac{1}{2}\right)$ and $p \in (1, \infty)$,

$$\mathbb{E}^P \left[ \sup_{n \geq 0} \|S_n\|_{p,\beta}^p \right] < \infty.$$

Knowing, as we already do, that $\{S_n : n \geq 0\}$ converges in $L^2([0,1]; \mathbb{R})$ almost surely, it is now elementary to check that, for each $\beta \in \left(0, \frac{1}{2}\right)$ and all $p \in (0, \infty)$, $\{S_n : t \geq 0\}$ is almost surely convergent in the Besov space $B_{p,\beta}$ obtained by completing $H$ with respect to $\| \cdot \|_{p,\beta}$. Finally, Besov’s embedding theorem says that the space of $\alpha$-Hölder continuous functions is continuously embedded in $B_{p,\beta}$ whenever $\alpha < \beta - \frac{2}{p}$. Hence, this procedure proves that, for the classical Wiener case, one can take $\Theta$ to be any one of the $\alpha$-Hölder spaces as long as $\alpha < \frac{1}{2}$.

Of course, the preceding is not the most elementary route to the almost sure convergence result just derived. Because we know that such a result holds for all choices of bases once it holds for any one of them, it makes sense to look for a basis which makes the derivation particularly simple. Such a basis was found by P. Lévy, who was not thinking in terms of orthonormal bases, and by Z. Ciesielski, who was. The basis which they took was the Haar basis, the derivative of whose elements are the $L^2([0,1]; \mathbb{R})$-orthonormalization of the indicator functions of dyadic intervals.
That is, \( h_0 = 1 \) and, if \( m = 2^\ell + k \) for some \( \ell \in \mathbb{N} \) and \( 0 \leq k < 2^\ell \),
\[
\hat{h}_m(t) = 2^\ell \begin{cases} 
1 & \text{when } 2^\ell t \in \left[k, \frac{2k+1}{2}\right), \\
-1 & \text{if } 2^\ell t \in \left[\frac{2k+1}{2}, k+1\right), \\
0 & \text{otherwise}.
\end{cases}
\]

The advantage of this basis is that one can easily check that if \( \| \cdot \|_u \) is the uniform norm on \( C([0,1]; \mathbb{R}) \), then the associated \( S_n \)’s satisfy
\[
\sup_{n > N} \| S_n - S_N \|_u \leq \sum_{\ell=0}^{\infty} 2^{-\frac{\ell}{2}} \max_{0 \leq k < 2^\ell} |X_{2^\ell+k}| \quad \text{for } N \geq 2^\ell.
\]
Hence, since
\[
\max_{0 \leq k < 2^\ell} |X_{2^\ell+k}| \leq \left( \sum_{2^{\ell} \leq m < 2^{\ell+1}} |X_m|^4 \right)^{\frac{1}{4}},
\]
it is clear that
\[
\lim_{N \to \infty} \mathbb{P} \left( \sup_{n > N} \| S_n - S_N \|_u \right) = 0.
\]

4. Some Properties of Abstract Wiener Spaces

In this concluding section I will discuss a few properties about abstract Wiener spaces, most of which are elementary applications of the Wiener series representation discussed in Section 2.

I begin with the renowned Cameron–Martin formula, the one which says that if \( h \in H \) and
\[
R_h(\theta) = \exp(\langle I(h) \rangle (\theta) - \frac{1}{2} \| h \|^2_H),
\]
then the distribution of \( \theta \to \theta + h \) under \( \mathcal{W}_H \) is absolutely continuous with respect to \( \mathcal{W}_H \) and that \( R_h \) is the corresponding Radon–Nikodym derivative. From the standpoint of Wiener series, this observation comes down to the fact that if one translates the standard Gauss distribution \( \gamma \) on \( \mathbb{R} \) by \( a \in \mathbb{R} \), then the translated measure is absolutely continuous with respect to \( \gamma \) and has Radon–Nikodym derivative \( \exp(ax - \frac{1}{2}a^2) \). To see why this implies Cameron and Martin’s result, assume that \( h \neq 0 \), set \( h_0 = \frac{h}{\| h \|_H} \), choose \( \{ h_m : m \geq 1 \} \) so that \( \{ h_m : m \geq 0 \} \) is an orthonormal basis in \( H \), and observe that
\[
\int_{\mathbb{R}} \Phi(\theta + h) \mathcal{W}_H(d\theta) = \int_{\mathbb{R}^n} \Phi(S_\infty(x) + h) \gamma^n(dx),
\]
where \( S_\infty(x) \) is the \( \gamma^n \)-almost sure limit of the partial sums in (2.1). Now apply the one-dimensional result with \( a = \| h \|_H \) to the 0th coordinate, and check that
\[
\int_{\mathbb{R}^n} e^{\| h \|^2_{H} - \frac{1}{2} \| h \|^2_H} \Phi(S_\infty(x)) \gamma^n(dx) = \mathbb{E}^{\mathcal{W}_H}[R_h \Phi].
\]

The next observation is that if \( (H, \Theta, \mathcal{W}_H) \) is an abstract Wiener space, then the support of \( \mathcal{W}_H \) is the whole of \( \Theta \). To see this, first observe that, because \( H \) is dense in \( \Theta \), one need only check that balls centered at elements of \( H \) have positive \( \mathcal{W}_H \)-measure. Second, use the Cameron–Martin formula to check that, for any \( \in H \), \( \mathcal{W}_H(B_{\Theta}(h, r)) > 0 \) if \( \mathcal{W}_H(B_{\Theta}(0, r)) > 0 \). Hence, all that remains is to
show that $W_H(\|\theta\|_\Theta < r) > 0$ for all $r > 0$. To this end, choose an orthonormal basis $\{h_m : m \geq 0\}$ for $H$, and remember that (cf. (2.2)), $W_H$-almost surely, $S_n(\theta) \rightarrow \theta$ in $\Theta$. Moreover, as an application of the fundamental property of Gaussian families, one can easily check that $\theta$ is $W_H$-independent of $\theta - S_n(\theta)$. Hence,

$$W_H(\|\theta\|_\Theta < r) \geq W_H(\|\theta - S_n(\theta)\|_\Theta < \frac{r}{2}) W_H(\|S_n(\theta)\|_\Theta < \frac{r}{2}).$$

By taking $n$ large enough, the first factor on the right can be made positive. At the same time,

$$\|S_n(\theta)\|_\Theta \leq C \left( \sum_{m=0}^{n} \|I(h_m)\|_\Theta \right)^\frac{1}{2},$$

where $C < \infty$ is the bound on the map taking $H$ into $\Theta$. Hence, the second factor dominates the $\gamma^{n+1}$-measure of the ball $B_{R_{n+1}}(0, \frac{r}{2^2})$, which is positive for all $n$’s.

The next property of an abstract Wiener space is one about which I do not feel completely comfortable. The property is the converse to the Cameron–Martin formula: the translate of $W_H$ by a $\varphi \in \Theta \setminus H$ is singular to $W_H$. To explain this property, let $\varphi \in \Theta$ be given, and use $T_{\varphi}W_H$ to denote the translate of Wiener measure by $\varphi$. Next, for $\lambda \in \Theta^*$ with $\|h_\lambda\|_H = 1$, let $\mathcal{F}_\lambda$ be the $\sigma$-algebra generated by $\theta \sim (\theta, \lambda)$ and set

$$r_\lambda(\theta) = \exp(\langle \varphi, \lambda \rangle \theta - \frac{1}{2} \|\varphi, \lambda\|^2).$$

Then, proceeding as in the proof of the Cameron–Martin formula, one can easily check that $r_\lambda$ is the Radon–Nikodym derivative of $T_{\varphi}W_H \restriction \mathcal{F}_\lambda$ with respect to $W_H \restriction \mathcal{F}_\lambda$. Hence, if $T_{\varphi}W_H$ is not singular to $W_H$ and if $R$ is the Radon–Nikodym derivative of its absolutely continuous part, then

$$r_\lambda \geq \mathbb{E}^{W_H} [R \mid \mathcal{F}_\lambda] \geq \mathbb{E}^{W_H} [R^\frac{1}{2} \mid \mathcal{F}_\lambda]^2,$$

and so

$$\exp \left(-\frac{\langle \varphi, \lambda \rangle^2}{8}\right) = \mathbb{E}^{W_H} [r_\lambda^\frac{1}{2}] \geq \alpha \equiv \mathbb{E}^{W_H} [R^\frac{1}{2}] \in (0, 1].$$

Since this means that $\|\varphi, \lambda\| \leq \sqrt{-8 \log \alpha} \|h_\lambda\|_H$ for all $\lambda \in \Theta^*$, it follows that $\varphi$ must be in $H$. In conjunction with the result of Cameron and Martin, this means that $T_{\varphi}W_H \ll W_H$ or $T_{\varphi}W_H \perp W_H$ according to whether $\varphi$ is or is not an element of $H$. So far so good. What bothers me is that there is another approach to this problem. Namely, choose $\{\lambda_m : m \geq 0\} \subseteq \Theta^*$ so that $\{h_{\lambda_m} : m \geq 0\}$ forms an orthonormal basis in $H$. Then, as an application of the Wiener series representation of $W_H$ and Kakutani’s theorem about absolute continuity of product measures, one can show that $T_{\varphi}W_H \ll W_H$ or $T_{\varphi}W_H \perp W_H$ according to whether $\sum_{m=0}^{\infty} \langle \varphi, \lambda_m \rangle^2$ converges or diverges. Combining these two, we arrive at the conclusion that, for any $\varphi \in \Theta$, $\varphi \in H$ if and only if $\sum_{m=0}^{\infty} \langle \varphi, \lambda_m \rangle^2$ converges, diverges. Although this latter conclusion seems reasonable, I do not think that the analogous statement holds for every separable Hilbert space $H$ which is continuously embedded as a dense subset of a Banach space $\Theta$, but it does hold if $H$ and $\Theta$ are components of an abstract Wiener triple.

I close with a remark which seems potentially useful, even though I have not found any particular use for it. It is based on the fact that, up to isometry, all
infinite dimensional, separable, real Hilbert space are the same. If one investigates what implication this fact has for abstract Wiener spaces, one finds that it leads to a rigid relationship between the family of abstract Wiener spaces associated with different Hilbert spaces. To be precise, let $F$ be a linear isometry from $H^1$ onto $H^2$, and let $\Theta^1$ be a Banach space for which $(H^1, \Theta^1, W_{H^1})$ is an abstract Wiener space. Then there is a Banach space $\Theta^2$ and a linear isometry $\tilde{F}$ from $\Theta^1$ onto $\Theta^2$ such that $\tilde{F} \upharpoonright H = F$ and $(H^2, \Theta^2, \tilde{F}, W_{H^2})$ is an abstract Wiener space. Like many such abstract results, this one is easier to prove than to state. To prove it, define $\|h^2\|_{\Theta^2} = \|F^{-1}h^2\|_{\Theta^1}$, and let $\Theta^2$ be the completion of $H^2$ with respect to $\| \cdot \|_{\Theta^2}$. Trivially, $F$ is an isometry from $H^1$ onto $H^2$ when $H^1$ is given the norm $\| \cdot \|_{\Theta^1}$ and $H^2$ the norm $\| \cdot \|_{\Theta^2}$. Hence, $F$ admits a unique extension $\tilde{F}$ as an isometry from $\Theta^1$ onto $\Theta^2$. Moreover, if $\mu = \tilde{F}_* W_{H^1}$, then

$$\hat{\mu}(\lambda^2) = \hat{W}_{H^1}(\tilde{F}^\top \lambda^2) = \exp(-\frac{1}{2} \|h^1_{\tilde{F}^\top \lambda^2}\|_{H^1}^2),$$

where $\tilde{F}^\top$ is the adjoint map from the dual space of $\Theta^2$ to the dual of $\Theta^1$. Finally, it is easy to check that $h^1_{\tilde{F}^\top \lambda^2} = F^{-1}h^2_{\lambda^2}$, which, since $F$ is an isometry, completes the proof that $(H^2, \Theta^2, \tilde{F}, W_{H^2})$ is an abstract Wiener space.

The reason for my thinking that this remark might be useful is that it allows one to choose one Hilbert space, for instance, the classical one, and use the family of abstract Wiener spaces associated with this one to calibrate the abstract Wiener spaces for any Hilbert space. For example, if one believes that $C_0 \equiv \{ \theta \in C([0, 1]; \mathbb{R}) : \theta(0) = 0 \}$ is better than $L^2([0, 1]; \mathbb{R})$ in the classical case, then one ought to believe that for any Hilbert space the $\Theta$ corresponding to $C_0$ is better than the one corresponding to $L^2([0, 1]; \mathbb{R})$. Unfortunately, except for inclusion properties, the meaning of better here is not entirely clear and may not be of any significance. In this connection, I have been wondering about the following question, to which one of my readers may already know the answer. Clearly, the family of abstract Wiener spaces for a given Hilbert $H$ space forms a net under inclusion. Is it true that the limit of this net is $H$ itself and, if so, in what sense?

Acknowledgment. I am grateful to Linan Chen for several conversations during which many of these ideas became clearer to me.

References


Daniel W. Stroock: Department of Mathematics 2-272, M.I.T., Cambridge, MA 02139, USA

E-mail address: dws@math.mit.edu
LARGE-TIME BEHAVIOR OF NON-SYMMETRIC FOKKER-PLANCK TYPE EQUATIONS

ANTON ARNOLD, ERIC CARLEN, AND QIANGCHANG JU

Abstract. Large time asymptotics of the solutions to non-symmetric Fokker-Planck type equations are studied by extending the entropy method to this case. We present a modified Bakry-Emery criterion that yields convergence of the solution to the steady state in relative entropy with an explicit exponential rate. In parallel it also implies a logarithmic Sobolev inequality w.r.t. the steady state measure. Explicit examples illustrate that skew-symmetric perturbations in the Fokker Planck operator can “help” to improve the constant in such a logarithmic Sobolev inequality.

1. Introduction

In this paper we consider the large-time behavior of the Cauchy problem for linear Fokker-Planck type equations (advection-diffusion equations) for probability densities $\rho(x,t)$:

$$
\rho_t = \mathcal{L}\rho := \text{div}(D[\nabla \rho + \rho(\nabla \phi + F)]), \quad x \in \mathbb{R}^n, \quad t > 0,
$$

(1.1a)

$$
\rho(t = 0) = \rho_I \in L^1_+(\mathbb{R}^n),
$$

(1.1b)

with the confinement potential $\phi = \phi(x)$ satisfying $e^{-\phi} \in L^1(\mathbb{R}^n)$, and the symmetric, locally uniformly positive definite diffusion matrix $D = D(x) = (d_{ij}(x))$.

Due to the divergence form we obviously have the conservation property

$$
\int_{\mathbb{R}^n} \rho(x,t) dx = \int_{\mathbb{R}^n} \rho_I(x) dx,
$$

(1.2)

and without restriction of generality we shall always assume

$$
\int_{\mathbb{R}^n} \rho_I(x) dx = \int_{\mathbb{R}^n} e^{-\phi(x)} dx = 1.
$$

Now suppose that the given vector field $F(x)$ and the scalar potential $\phi(x)$ satisfy

$$
\text{div}(D e^{-\phi} F) = 0 \quad \text{on} \quad \mathbb{R}^n.
$$

(1.3)

Then the unique normalized steady state of (1.1a) is $\rho_{\infty} = e^{-\phi} \in L^1_+(\mathbb{R}^n)$. Because of (1.3), $\mathcal{L}_{SSP} := \text{div}(D\rho F)$ is the skew-symmetric part of the operator $\mathcal{L}$ in $L^2(\mathbb{R}^n; \rho_{\infty}^{-1}(dx))$ acting on $\rho$, and this skew-symmetric part annihilates the steady state $\rho_{\infty}$. Hence, the steady state is independent of $F$.

2000 Mathematics Subject Classification. Primary 35K57, 35B40, 26D10; Secondary 37A35, 47D07, 60J60.

Key words and phrases. Entropy method, Fokker-Planck equation, large-time convergence, logarithmic Sobolev inequality, Bakry-Emery condition.
In the sequel we shall assume that the data $\phi$, $D$, $F$, and $\rho_I$ are sufficiently regular (for example, $\phi \in W^{2,\infty}(\mathbb{R}^n; \mathbb{R})$, $d^{ij} \in W^{2,\infty}(\mathbb{R}^n; \mathbb{R})$, $i, j = 1, \ldots, n$, and $F = (F_i(x)) \in W^{1,\infty}(\mathbb{R}^n; \mathbb{R}^n)$) such that (1.1) has a unique solution $\rho \in C([0, \infty), L^1(\mathbb{R}^n))$, and $\rho \in C([0, \infty), L^r(\mathbb{R}^n; \rho_\infty^{-1}(dx)))$ if $\rho_I \in L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))$.

We remark that (by a simple minimum principle) $\rho_I(x) \geq 0$ implies $\rho(x,t) > 0$ for all $x \in \mathbb{R}^n, t > 0$.

Simple examples of (1.1a) include the (symmetric) Fokker-Planck equation [18, 22]

$$\rho_t = \text{div}(\nabla \rho + x \rho), \quad x \in \mathbb{R}^n, \quad t > 0,$$

where $\phi(x) = |x|^2/2 + \text{const}, D = I$ (I being the identity matrix), and $F = 0$.

As $t \to \infty$ its solution converges with an exponential rate towards the Gaussian steady state

$$\rho_\infty(x) = (2\pi)^{-n/2}e^{-|x|^2/2}.$$

An important example for a non-symmetric equation is the quantum-kinetic Wigner-Fokker-Planck equation (cf. [3, 19]) with the quadratic confinement potential $V(y) = |y|^2/2$:

$$w_t + v \cdot \nabla_y w - y \cdot \nabla_v w = D_{pp}\Delta_y w + 2\gamma \text{div}_y(v w) + 2D_{pq}\text{div}_y(\nabla_v w) + D_{qq}\Delta_y v, \quad y, v \in \mathbb{R}^n, \quad t > 0.$$

(1.4) can indeed be recast in the form of (1.1a) (see [19] for details). It describes the evolution of the Wigner function $w(y, v, t)$ with the position variable $y$ and velocity $v$, and $(y, v)$ plays here the role of $x$ in (1.1a). Under simple (and physically necessary) assumptions on the r.h.s. of (1.4), $w(t)$ also converges exponentially to the unique steady state $w_\infty$.

Other examples of non-symmetric Fokker-Planck equations appear in the modelling of polymeric fluid flows, where $\rho(x,t), x \in \mathbb{R}^n$ describes the distribution of polymeric chains of length and orientation given by $x$. In a given homogeneous shear flow $u(x) = \mathbf{F} \cdot x$ the (scaled) evolution equation reads (cf. [14] for details)

$$\rho_t = \frac{1}{2}\text{div}(\nabla \rho + \rho(\nabla \phi - 2\mathbf{F} \cdot x)).$$

In this paper we are interested in the possibly exponential decay rate of $\rho(t)$ towards $\rho_\infty$ in relative entropy, i.e.

$$\epsilon(\rho|\rho_\infty) := \int_{\mathbb{R}^n} \rho \ln \frac{\rho}{\rho_\infty} dx \geq 0.$$  \hspace{1cm} (1.6)

This exponential convergence is closely related to the hypercontractivity of the semigroup generated by $L$ and to the validity of a logarithmic Sobolev inequality (LSI) w.r.t. the steady state measure $\rho_\infty$ (cf. [12, 13, 4]). In the case $D(x) \equiv I$ this inequality would read, if it holds,

$$\int_{\mathbb{R}^n} f^2 \ln f^2 \rho_\infty dx - \left( \int_{\mathbb{R}^n} f^2 \rho_\infty dx \right) \ln \left( \int_{\mathbb{R}^n} f^2 \rho_\infty dx \right) \leq C \int_{\mathbb{R}^n} |\nabla f|^2 \rho_\infty dx$$

for some fixed $C < \infty$ and all $f \in L^2(\mathbb{R}^n; \rho_\infty(dx))$. Notice that $\phi$ enters the inequality through $\rho_\infty$, but that $F$ does not. The question to be addressed here
is whether it is ever advantageous to consider a non-reversible evolution (i.e., one with \( F \neq 0 \)) when attempting to establish the validity of (1.7) through the entropy method of [5, 7, 4]. Perhaps surprisingly, the answer is yes.

To be more specific we shall now briefly outline this idea for the simplest case when \( D = I \), following the preliminary note [2]: Consider the symmetric part (in \( L^2(\mathbb{R}^n; \rho_\infty^{-1}dx) \)) of the Fokker-Planck operator in (1.1a), i.e.

\[
\mathcal{L}_S \rho := \text{div} \left( e^{-\phi} \nabla \frac{\rho}{e^{-\phi}} \right)
\]

and assume that the potential \( \phi(x) \) is uniformly convex, i.e. it satisfies a Bakry-Emery condition (BEC):

\[
(\text{A1}) \quad \exists \lambda_1 > 0 \text{ such that } \frac{\partial^2 \phi}{\partial x^2} = \left( \frac{\partial^2 \phi(x)}{\partial x_i \partial x_j} \right)_{i,j=1,...,n} \geq \lambda_1 I \quad \forall x \in \mathbb{R}^n
\]

(in the sense of positive definite matrices). Then it is well known that \( \rho_\infty(t) \), the solution of \( \rho_t = \mathcal{L}_S \rho \) converges to \( \rho_\infty \) in relative entropy with an exponential rate of (at least) \( 2\lambda_1 \), and the LSI (1.7) holds with \( C = 2/\lambda_1 \) (cf. [5, 7, 4]). Moreover, (1.3) implies that also \( \rho(t) \), the solution to the non-symmetric Fokker-Planck equation (1.1) converges to \( \rho_\infty \) in relative entropy with rate of (at least) \( 2\lambda_1 \) (cf. [4]).

On the other hand, consider now \( \mathcal{L}_{SS} \) with \( F(x) \neq 0 \) as a skew-symmetric perturbation of \( \mathcal{L}_S \) and assume that \((\phi, F)\) satisfy a generalized Bakry-Emery condition (GBEC):

\[
(\text{A2}) \quad \exists \lambda_2 > 0 \text{ such that } \frac{\partial^2 \phi}{\partial x^2} - \frac{1}{2} \left( \frac{\partial F}{\partial x} + \left( \frac{\partial F}{\partial x} \right)^\top \right) \geq \lambda_2 I \quad \forall x \in \mathbb{R}^n,
\]

where \( \left( \frac{\partial F}{\partial x} \right)_{i,j} = \frac{\partial F_i}{\partial x_j} \) denotes the Jacobian of \( F \). As we shall show, the relative entropy of \( \rho(t) \) then decays exponentially with rate (at least) \( 2\lambda_2 \), and the LSI (1.7) then holds with \( C = 2/\lambda_2 \). And in some cases the perturbation \( F \) gives rise to a ‘better’ constant \( \lambda_2 > \lambda_1 \), hence ‘improving’ (1.7).

The goal of this paper is threefold: to understand the large-time behavior of non-symmetric Fokker-Planck equations (with applications like (1.4), (1.5)), to analyze skew-symmetric perturbations \( \mathcal{L}_{SS} \) in order to possibly improve the LSI. And, finally, our analysis furnishes a proof of the entropy method for symmetric Fokker-Planck equations with full diffusion matrices \( D \) (which was not included in [4]).

This paper is organized as follows. We begin Section 2 by introducing the class of entropies with which we work. We then proceed to calculate the second derivative of the entropy in the presence of the skew-symmetric term, and derive the generalized Bakry–Emery condition. To get an estimate on the initial entropy in terms of the initial entropy production by the Bakry-Emery method, we need to know that the final entropy is zero. For this we need a theorem asserting that the entropy vanishes in the large time limit. We prove this in Theorem 2.5 for regular initial data. This part of the proof is somewhat technically involved, but once we have it, even for regular densities, the rest is straightforward: We then use the
results obtained up to this point to obtain a LSI for regular densities. Once this
is extended by simple closure, the fact that the entropy vanishes in the large time
limit, exponentially fast, then follows easily for general initial data.

In Section 3 we discuss several examples in which the skew-symmetric term
plays a crucial role in establishing the LSI.

2. Entropy Method for Non-symmetric Fokker-Planck Equations

2.1. Admissible relative entropies. As a generalization of the logarithmic en-
tropy (1.6) we now introduce the relative entropies that we shall use in the sequel.

Definition 2.1. Let $\psi \in C[0, \infty) \cap C^4(0, \infty)$ satisfy the conditions

$$\psi(1) = \psi'(1) = 0,$$  (2.1a)
$$\psi'' > 0, \text{ on } \mathbb{R}^+,$$  (2.1b)
$$(\psi''')^2 \leq \frac{1}{2} \psi'' \psi'''' \text{ on } \mathbb{R}^+.$$  (2.1c)

Let $\rho_1, \rho_2 \in L^1_+(\mathbb{R}^n)$ with $\int \rho_1 \, dx = \int \rho_2 \, dx = 1$ and $\rho_1/\rho_2 \in \mathbb{R}_{0^+}^\times \rho_2(dx)-a.e.$ Then

$$e_\psi(\rho_1|\rho_2) := \int_{\mathbb{R}^n} \psi\left(\frac{\rho_1}{\rho_2}\right) \rho_2(dx) \geq 0$$  (2.2)

is called an admissible relative entropy (of $\rho_1$ with respect to $\rho_2$) with generating
function $\psi$.

Our class of generating functions $\psi$ coincides with those considered in [5] (up
to the normalizations (2.1a)). The most typical examples of admissible relative
entropies are the physical relative entropy (1.6) generated by

$$\psi_1(\sigma) := \sigma \ln \sigma - \sigma + 1, \quad \sigma \geq 0,$$

and the $p$-entropies (or Tsallis relative entropies [23]) generated by

$$\psi_p(\sigma) := \frac{\sigma^p - p\sigma + 1}{p - 1}, \quad \sigma \geq 0; \quad 1 < p \leq 2.$$  (2.3)

For $p = 2$ we have $e_{\psi_2}(\rho_1|\rho_2) = ||\rho_1 - \rho_2||^2_{L^2(\mathbb{R}^n, \rho_2^{-1}(dx))}$.

The well-known Csiszár-Kullback inequality [10, 16, 24, 4] shows that the relative
entropies (2.2) are a ‘measure’ for the distance between two normalized
$L^1_+(\mathbb{R}^n)$-functions $\rho_1, \rho_2$:

$$\frac{1}{2} ||\rho_1 - \rho_2||^2_{L^1(\mathbb{R}^n)} \leq \frac{1}{\eta_2} e_\psi(\rho_1|\rho_2),$$  (2.4)

with the notation $\eta_2 := \psi''(1)$.

We remark that for each admissible relative entropy $e_\psi$, there exists a quadratic
superentropy $e_\varphi$ such that

$$0 \leq \psi(\sigma) \leq \eta_2(\sigma - 1)^2 =: \varphi(\sigma), \quad \sigma > 0,$$

and hence $e_\psi(\rho_1|\rho_2) \leq e_\varphi(\rho_1|\rho_2)$ (cf. Lemma 2.6 in [4]).
2.2. Generalized Bakry-Emery condition and Ricci tensor. As in (A1) and (A2), a Bakry-Emery condition (BEC) on the coefficients \((\phi, D, F)\) of the Fokker-Planck operator \(\tilde{L}\) will be our main assumption for deriving exponential decay of the relative entropy. For the subsequent analysis it is convenient to rewrite (1.1). We set

\[
\mu := \rho / \rho_{\infty},
\]

which satisfies the IVP

\[
\begin{align*}
\mu_t &= \tilde{L}_\mu := \rho^{-1}_\infty \text{div}(\rho_\infty D \nabla \mu) + F^T D \nabla \mu \\
&= \text{div}(D \nabla \mu) - (\nabla \phi - F)^T D \nabla \mu, \quad x \in \mathbb{R}^n, \quad t > 0, \\
\mu_I &= \rho_I / \rho_{\infty} \in L^1(\mathbb{R}^n, \rho_\infty(dx)).
\end{align*}
\]

Condition (A1) is a special case of the well-known Bakry-Emery condition for logarithmic Sobolev-inequalities \([5, 6, 7, 4]\). In order to extend the approach of Bakry and Emery to non-symmetric Fokker-Planck equations we shall now introduce a new generalized Bakry-Emery condition (GBEC). For understanding the BEC in the case of general (symmetric and uniformly positive definite) diffusion matrices \(D(x)\) we shall need some notions from basic differential geometry (see, e.g. \([8]\), §7, 8). Therefore we consider the Riemannian manifold \(\mathcal{M} = (\mathbb{R}^n; D^{-1})\), with \(D(x)^{-1} := (d_{ij}(x))\) as covariant metric tensor.

The Ricci tensor of a symmetric Fokker-Planck operator was defined in \([7]\) (cf. also \([4]\)). Here we shall extend this definition to non-symmetric Fokker-Planck operators that involve a vector field \(F = (F_1, \ldots, F_n)^T\). The Fokker–Planck operator in (2.5) can be decomposed as

\[
\tilde{L} = \Delta^D + \mathcal{X},
\]

where

\[
\Delta^D \mu := (\det D)^{1/2} \text{div} \left[ (\det D)^{-1/2} D \nabla \mu \right]
\]

is the Laplace–Beltrami operator on \(\mathcal{M}\) (cf. \([9]\), §1). And

\[
\mathcal{X} := \sum_{i=1}^n X^i \frac{\partial}{\partial x_i}
\]

is a vector field (or, equivalently, a directional derivative) on \(\mathcal{M}\), with the components

\[
X^i(x) := -\sum_{j=1}^n d^{ij} \left[ \frac{\partial}{\partial x_j} \left( \phi(x) - \frac{1}{2} \ln \det D(x) \right) - F_j \right].
\]

The Christoffel symbols are defined as the elements of the 3-tensor:

\[
\Gamma^l_{ij} := \frac{1}{2} \sum_{k=1}^n d^{kl} \left( \frac{\partial d_{jk}}{\partial x_i} + \frac{\partial d_{ki}}{\partial x_j} - \frac{\partial d_{ij}}{\partial x_k} \right).
\]

The Riemann curvature tensor of \(\mathcal{M}\) then reads

\[
R^l_{ikj} := \frac{\partial}{\partial x_i} \Gamma^l_{jk} - \frac{\partial}{\partial x_j} \Gamma^l_{ik} + \sum_{m=1}^n \Gamma^l_{im} \Gamma^m_{jk} - \sum_{m=1}^n \Gamma^l_{jm} \Gamma^m_{ik}.
\]
and the (symmetric) Ricci-tensor of $\mathcal{M}$ is (cf. [20], §C.3)

$$\rho_{ij} := \sum_{k=1}^{n} R_{ik} R_{kj}. \quad (2.9)$$

The covariant derivative of a vector field $\mathcal{X} = (X^1, \ldots, X^n)$ is given by

$$\nabla_i X^j := \frac{\partial X^j}{\partial x^i} + \sum_{k=1}^{n} \Gamma^j_{ik} X^k. \quad (2.10)$$

We define the symmetric covariant derivative (2-tensor) of $\mathcal{X}$ as

$$\left(\nabla^S \mathcal{X}\right)_{ij} := \frac{1}{2} \sum_{l=1}^{n} \left( d_{jl} \nabla_i X^l + d_{il} \nabla_j X^l \right). \quad (2.11)$$

We now define the Ricci tensor of a non-symmetric Fokker-Planck operator as

$$\text{Ric}^{ij}(x) := \sum_{k,l=1}^{n} d^{ik} d^{jl} \left[ \rho_{kl} - \left(\nabla^S \mathcal{X}\right)_{kl}(x) \right]. \quad (2.12)$$

with the components of $\mathcal{X}$ defined in (2.6) (cf. [7, 4] for the symmetric counterpart).

Our GBEC for a general symmetric, positive definite diffusion matrix now reads:

\[ (A3) \exists \lambda_3 > 0 \text{ such that } \text{Ric}(x) \geq \lambda_3 D(x) \quad \forall x \in \mathbb{R}^n \]

(in the sense of positive definite matrices). From the explicit form of $\text{Ric}$ (see (2.13) below) one easily sees that (A3) reduces to (A2) for $D(x) \equiv I$. And in the case of a scalar diffusion (i.e. $D(x) = D(x) I$) it reads:

\[ (A4) \exists \lambda_3 > 0 \text{ such that } \]

\[ \left( \frac{1}{2} - \frac{n}{4} \right) \frac{1}{D} \nabla D \nabla D \left( D \frac{\partial D}{\partial x^2} \right) \left( \nabla \phi - F \right) - \frac{1}{2} \left( \nabla D \nabla \phi - F \right) \]

\[ + \frac{D}{2} \left( \frac{\partial D}{\partial x} \left( \nabla \phi - F \right) + \frac{2}{\partial x} \left( \nabla D \nabla \phi - F \right) \right) \]

\[ \geq \lambda_3 I \]

\[ \forall x \in \mathbb{R}^n. \]

With tedious calculations (see the Appendix), the explicit form of the GBEC reads:

\[ \exists \lambda_3 > 0 \text{ such that } \]

\[ U^T \left[ \frac{1}{2} \text{Tr} \left( D \frac{\partial^2}{\partial x^2} D + \frac{1}{2} \left( \nabla^T D \nabla \right) D - D \left( \frac{\partial^2}{\partial x^2} D \right) + D \left( \frac{\partial^2 \phi}{\partial x^2} \right) D \right) \right] U \]

\[ - \frac{1}{2} \left( \frac{\partial D}{\partial x} + \frac{\partial D}{\partial x} \right) \left( D \left( \nabla \phi - F \right) \right) \]

\[ + \frac{1}{2} \left( U^T \nabla D \left( \nabla \phi - F \right) \right) \]

\[ - \frac{1}{2} \text{Tr} \left( E^T + D E D^{-1} - N(D) D^{-1} \right)^2 \]

\[ \geq \lambda_3 U^T D U \]
∀x ∈ \mathbb{R}^n and any vector field \( U : \mathbb{R}^n \to \mathbb{R}^n \). Here we used the matrix \( E = (e^j_i) := (\partial_i d^{ik}) U_k \) (\( i \) is the ‘first index’ in \( e^j_i \)). And \( \mathcal{N} := U^\top D \) is a scalar differential operator, which acts elementwise when applied to a matrix. The expression \( \frac{\partial^2}{\partial x^2} D \) denotes the (formal) matrix product between the Hessian operator and the matrix \( D \), i.e. \( \partial_{ij} (d^{ik}) \).

### 2.3. Exponential decay of the entropy dissipation and the relative entropy.

In this section, we shall first obtain the exponential decay of the entropy dissipation. As in [4], we consider the entropy dissipation

\[
I_\psi(\rho(t)|\rho_\infty) := \frac{d}{dt} \psi(\rho(t)|\rho_\infty)
\]

and the entropy dissipation rate

\[
R_\psi(\rho(t)|\rho_\infty) := \frac{d}{dt} I_\psi(\rho(t)|\rho_\infty).
\]

Eq. (2.14) is referred to as entropy equation. To facilitate the computations we rewrite (1.1a) in the following form:

\[
\rho_t = \text{div}(D(\rho_\infty U + \rho F))
\]

with the notation \( U = \nabla(\frac{\rho}{\rho_\infty}) \). Differentiating the relative entropy \( e_\psi(\rho(t)|\rho_\infty) \) w.r.t. time gives

\[
I_\psi(\rho(t)|\rho_\infty) = \int_{\mathbb{R}^n} \psi\left(\frac{\rho}{\rho_\infty}\right) \rho_t \, dx.
\]

By using (2.16) we obtain after an integration by parts

\[
I_\psi(\rho(t)|\rho_\infty) = - \int_{\mathbb{R}^n} \psi''\left(\frac{\rho}{\rho_\infty}\right) U^\top D U \rho_\infty \, dx + T
\]

where \( T := \int_{\mathbb{R}^n} \psi\left(\frac{\rho}{\rho_\infty}\right) \text{div}(DF\rho) \, dx \). In terms of (1.3), we get

\[
\text{div}(DF\rho) = U^\top DF \rho_\infty,
\]

from which we have

\[
T = - \int_{\mathbb{R}^n} \psi\left(\frac{\rho}{\rho_\infty}\right) U^\top D F \rho_\infty \, dx
\]

\[
= \int_{\mathbb{R}^n} \nabla^\top \psi\left(\frac{\rho}{\rho_\infty}\right) D F \rho_\infty \, dx
\]

\[
= 0
\]

by again using (1.3). Therefore

\[
I_\psi(\rho(t)|\rho_\infty) = - \int_{\mathbb{R}^n} \psi''\left(\frac{\rho}{\rho_\infty}\right) U^\top D U \rho_\infty \, dx \leq 0,
\]

due to the strict convexity of \( \psi \) and the positivity of \( D \).
Now, we compute (2.15):

\[
R_\psi(\rho(t) | \rho_\infty) = - \int_{\mathbb{R}^n} \psi'''(\frac{\rho}{\rho_\infty}) \rho_t U^\top D U \, dx
- 2 \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) U^\top D U_t \rho_\infty \, dx,
\]

\[
= R_1 + R_2, \tag{2.21}
\]

where

\[
R_1 := - \int_{\mathbb{R}^n} \psi'''(\frac{\rho}{\rho_\infty}) \rho_t U^\top D U \, dx
\]

and

\[
R_2 := -2 \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) U^\top D U_t \rho_\infty \, dx.
\]

With (2.16) we get

\[
R_1 = - \int_{\mathbb{R}^n} \psi'''(\frac{\rho}{\rho_\infty}) \text{div}(D U \rho_\infty) U^\top D U \, dx
- \int_{\mathbb{R}^n} \psi'''(\frac{\rho}{\rho_\infty}) \text{div}(D F \rho) U^\top D U \, dx
= \tilde{R}_1 + T_3, \tag{2.22}
\]

where

\[
\tilde{R}_1 := - \int_{\mathbb{R}^n} \psi'''(\frac{\rho}{\rho_\infty}) \text{div}(D U \rho_\infty) U^\top D U \, dx
\]

and

\[
T_3 := - \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) \text{div}(D F \rho) U^\top D U \, dx.
\]

Using (2.19) and an integration by parts lead to

\[
T_3 = - \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) \nabla^\top \left( \frac{\rho}{\rho_\infty} \right) D F (U^\top D U) \rho_\infty \, dx
= - \int_{\mathbb{R}^n} \nabla^\top \psi''(\frac{\rho}{\rho_\infty}) D F (U^\top D U) \rho_\infty \, dx
= \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) \text{div}(D F (U^\top D U) \rho_\infty) \, dx
= \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) \nabla^\top (U^\top D U) D F \rho_\infty \, dx.
\]

From (2.16) and (2.19), it follows that

\[
U_t = \nabla \left( \frac{1}{\rho_\infty} \text{div}(D U \rho_\infty) + U^\top D F \right).
\]

Then

\[
R_2 = -2 \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) U^\top D \nabla [e^{\phi} \text{div}(D e^{\phi} U)] \rho_\infty \, dx
- 2 \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) U^\top D \nabla (U^\top D F) \rho_\infty \, dx
= \tilde{R}_2 + T_4, \tag{2.23}
\]
where
\[ \tilde{R}_2 := -2 \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) U^\top D \nabla \left[ e^{\phi} \text{div}(De^{-\phi}U) \right] \rho_\infty \, dx \]
and
\[ T_4 := -2 \int_{\mathbb{R}^n} \psi''(\frac{\rho}{\rho_\infty}) U^\top D \nabla (U^\top DF) \rho_\infty \, dx. \]

Clearly, the computations which lead to (2.20) and (2.21) are formal. However, they can easily be justified for initial data \( \rho_I \in L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx)) \subset L^1(\mathbb{R}^n) \) and for entropy generators without singularity at \( \sigma = 0 \) by taking into account the semigroup property of the evolution in \( L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx)) \), and the fact that \( \rho > 0 \) on \( \mathbb{R}^n \), \( t > 0 \). General admissible entropies can easily be dealt with by a local cut-off at \( \sigma = 0 \).

**Remark 2.2.** Following the approach from [4] we have to give a meaning to \( I_\psi(\rho|\rho_\infty) \) even when \( \rho \) becomes zero (which may be the case at the initial state). For positive and differentiable functions \( \mu = \mu(x) \) we have
\[ \psi''(\mu)(\nabla \mu)^\top D \nabla \mu = \left( \nabla \int_1^\mu \sqrt{\psi''(s)} \, ds \right)^\top D \nabla \int_1^\mu \sqrt{\psi''(s)} \, ds. \quad (2.24) \]

Hence, we set for \( \rho \geq 0 \)
\[ I_\psi(\rho|\rho_\infty) := -\int_{\mathbb{R}^n} (\nabla w)^\top D \nabla w \rho_\infty \, dx, \quad w := h_\psi \left( \frac{\rho}{\rho_\infty} \right), \quad (2.25) \]
if \( w \in H^1_{\text{loc}}(\mathbb{R}^n) \) with
\[ h_\psi(\mu) := \int_1^\mu \sqrt{\psi''(s)} \, ds, \quad \mu > 0. \quad (2.26) \]

As shown is [4], \( h_\psi \) is Hölder continuous with exponent 1/2 locally at \( \mu = 0 \).

We now return to proving the exponential decay of \( e_\psi(\rho(t)|\rho_\infty) \) under additional assumptions on \( \phi, F, \) and \( D \). At first we shall derive an exponential decay rate for the entropy dissipation \( I_\psi \) by using the special form of the entropy dissipation rate (2.21).

**Lemma 2.3.** Let the initial condition \( \rho_I \in L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx)) \) satisfy \( |I_\psi(\rho_I|\rho_\infty)| < \infty \) for the admissible entropy \( e_\psi \). Assume that the coefficients \( \phi(x), F(x), \) and \( D(x) \) of (1.1a) satisfy the condition (A3). Then the entropy dissipation converges to 0 exponentially:
\[ |I_\psi(\rho(t)|\rho_\infty)| \leq e^{-2\lambda_3 t} |I_\psi(\rho_I|\rho_\infty)|, \quad t > 0. \quad (2.27) \]
Proof. An integration by parts (which can be justified as mentioned above) yields
\[
\tilde{R}_1 = \int_{\mathbb{R}^n} \psi'^V (e^\phi \rho)(U^\top D U)^2 e^{-\phi} \, dx \\
+ 2 \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} (DU)^\top \frac{\partial U}{\partial x} \frac{\partial U}{\partial x} \, dx \\
+ \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} \left( U_i U_j \frac{\partial^2 U_i}{\partial x_k} d^k U_l \right) \, dx \\
= \int_{\mathbb{R}^n} \psi'^V (e^\phi \rho)(U^\top D U)^2 e^{-\phi} \, dx \\
+ \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} U^\top \left[ D \frac{\partial U}{\partial x} + \frac{\partial (DU)}{\partial x} D \right] U \, dx.
\]

Here and in the sequel we use the Einstein summation convention for double indices. Also we abbreviate \( \frac{\partial}{\partial x} \) by \( \partial \).

Motivated by the scalar diffusion case (cf. [4]), we introduce
\[
S_1 := \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} \left[ 2(DU)^\top \frac{\partial U}{\partial x} (DU) + U_i U_j \frac{\partial^2 U_i}{\partial x_k} d^k U_l \right] \, dx \\
= \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} U^\top \left[ D \frac{\partial U}{\partial x} + D \frac{\partial (DU)}{\partial x} \right] U \, dx \quad \text{ (2.28)}
\]

In a cumbersome calculation we obtain
\[
S_2 := \tilde{R}_2 - S_1 \\
= 2 \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} \left[ U^\top \left( D \frac{\partial (D\nabla \phi)}{\partial x} \right)^\top - D \frac{\partial^2}{\partial x^2} (D) \\
- \frac{1}{2} (\nabla \phi^\top D \nabla \phi) D + \frac{1}{2} \text{Tr}(D \frac{\partial^2}{\partial x^2} D) + \frac{1}{2} (\nabla^\top D \nabla) D \right] U \\
- \frac{1}{4} \text{Tr} \left( D^\top D + \frac{1}{2} \text{Tr}(D \frac{\partial^2}{\partial x^2} D) + \frac{1}{2} (\nabla^\top D \nabla) D \right)^2 \right] \, dx \\
= 2 \int_{\mathbb{R}^n} \psi'''(e^\phi \rho) e^{-\phi} \left[ D \frac{\partial U}{\partial x} + \frac{1}{2} E^\top + \frac{1}{2} D E D^{-1} - \frac{1}{2} N(D) D^{-1} \right] \, dx
\]
\]

Next we rewrite \( T_3 \) and \( T_4 \) as
\[
T_3 = \int_{\mathbb{R}^n} \psi''(e^\phi \rho) e^{-\phi} \left[ F^\top D \frac{\partial U}{\partial x} DU + F^\top D \left( \frac{\partial (DU)}{\partial x} \right)^\top U \right] \, dx,
\]
and

\[ T_4 = -2 \int_{\mathbb{G}^n} \psi''(e^\phi \rho) e^{-\phi} \left[ U^\top \frac{\partial F}{\partial x} D U + U^\top D \left( \frac{\partial (DU)}{\partial x} \right)^\top F \right] dx, \]

where we used \( \frac{\partial U}{\partial x} = \left( \frac{\partial U}{\partial x} \right)^\top \), since \( U \) is a gradient.

Using the fact that \( \frac{\partial}{\partial x} \left( D U \right) = E + \frac{\partial U}{\partial x} D \), we have

\[ T_3 + T_4 = 2 \int_{\mathbb{G}^n} \psi''(e^\phi \rho) e^{-\phi} \left[ -U^\top \frac{\partial F}{\partial x} D U + U^\top \left( \frac{1}{2} E^\top D - DE \right) F \right] dx \]

Then

\[ T_1 + T_3 + T_4 = 2 \int_{\mathbb{G}^n} \psi''(e^\phi \rho) e^{-\phi} \left\{ U^\top \left[ \frac{1}{2} \text{Tr} \left( D \frac{\partial^2}{\partial x^2} D \right) + \frac{1}{2} \left( (\nabla^2 \phi) \nabla \right) D \right] 
- \frac{1}{2} \left( \frac{\partial^2}{\partial x^2} D \right) + \frac{1}{2} D \left( \frac{\partial F}{\partial x} \right) + \frac{1}{2} D \left( \frac{\partial F}{\partial x} \right)^\top \right) D 
- \frac{1}{2} \left( (\nabla \phi - F)^\top D \nabla \right) U 
+ \frac{1}{2} \left( U^\top \nabla \phi - F \right)^\top E^\top D U 
- \frac{1}{4} \text{Tr} \left( E^\top + \nabla \phi - F \right)^\top E^\top D U 
\right\} dx. \]

Condition (A3) (or (2.13)) leads to the estimates:

\[ T_1 + T_3 + T_4 \geq 2 \lambda_3 \int_{\mathbb{G}^n} \psi''(e^\phi \rho) e^{-\phi} U^\top D U dx. \]

All in all we have by using (2.28)

\[ R_1 + R_2 = (\tilde{R}_1 + S_1 + T_2) + (T_1 + T_3 + T_4) \]

\[ \geq \int_{\mathbb{G}^n} \left\{ \psi'(e^\phi \rho) (U^\top D U)^2 + 2 \psi''(e^\phi \rho) U^\top \left( D \frac{\partial U}{\partial x} + \frac{\partial (DU)}{\partial x} \right) U \right. 
+ 2 \psi''(e^\phi \rho) e^{-\phi} \text{Tr} \left\{ D \frac{\partial U}{\partial x} + \frac{1}{2} E^\top + \frac{1}{2} D E D^{-1} - \frac{1}{2} N(D) \right\} dx 
\right\} e^{-\phi} \]

The first integral can be written as

\[ \int_{\mathbb{G}^n} \text{Tr}(X Y) e^{-\phi} dx, \]
where $X$ and $Y$ are the $2 \times 2$-matrices

$$X = \begin{pmatrix} 2\psi''(e^\rho) & 2\psi'''(e^\rho) \\ 2\psi''(e^\rho) & \psi IV(e^\rho) \end{pmatrix}$$

and, resp.,

$$Y = \begin{pmatrix} \frac{1}{2}U^\top \left(D \frac{\partial U}{\partial x} D + \frac{\partial (DU)}{\partial x} D \right) U \\ \frac{1}{2}U^\top \left(D \frac{\partial U}{\partial x} D + \frac{\partial (DU)}{\partial x} D \right) U \right) \right) \end{pmatrix},$$

with

$$\alpha = \text{Tr} \left[ D \frac{\partial U}{\partial x} + \frac{1}{2} E^\top + \frac{1}{2} D E D^{-1} - \frac{1}{2} N(D) D^{-1} \right]^2.$$  

$X$ is non-negative definite since $\psi$ generates an admissible entropy (cf. Definition 2.1). Next, we will show that $Y$ is also non-negative definite. To this end, we introduce the symmetric matrices $Z$ and $W$ as follows:

$$Z := \sqrt{D} \frac{\partial U}{\partial x} \sqrt{D} + \frac{1}{2} \sqrt{D}^{-1} E^\top \sqrt{D} + \frac{1}{2} \sqrt{D} E \sqrt{D}^{-1} - \frac{1}{2} \sqrt{D}^{-1} N(D) \sqrt{D}^{-1}$$

and

$$W := \sqrt{D} (U \otimes U) \sqrt{D}.$$  

Using the cyclicity of the trace, we prove

$$\alpha = \text{Tr} Z^2,$$

$$\text{Tr} W^2 = \text{Tr} \left[ \sqrt{D} (U \otimes U) D (U \otimes U) \sqrt{D} \right]$$

and

$$\begin{aligned} \text{Tr}(W Z) &= \text{Tr}(Z W) \\ &= \text{Tr} \left[ \left( D \frac{\partial U}{\partial x} D + \frac{1}{2} E^\top D + \frac{1}{2} D E - \frac{1}{2} N(D) \right) U \otimes U \right] \\ &= U^\top \left[ \left( D \frac{\partial U}{\partial x} D + \frac{1}{2} E^\top D + \frac{1}{2} D E - \frac{1}{2} N(D) \right) U \right] \\ &= U^\top \left[ \left( D \frac{\partial U}{\partial x} D + \frac{1}{2} E^\top D \right) U \right] \\ &= \frac{1}{2} U^\top \left[ \left( D \frac{\partial U}{\partial x} D + \frac{\partial (DU)}{\partial x} D \right) U \right]. \end{aligned}$$

Then, it follows that

$$Y = \begin{pmatrix} \text{Tr} Z^2 & \text{Tr}(Z W) \\ \text{Tr}(W Z) & \text{Tr} W^2 \end{pmatrix}. $$
Since
\[
\begin{pmatrix}
Z^2 & ZW \\
ZW & W^2
\end{pmatrix} \geq 0
\]
by the positivity of partial traces (we include a proof in Lemma 2.4 below for completeness), \(Y\) is nonnegative. Thus
\[
\int_{\mathbb{R}^n} \text{Tr}(XY)e^{-\phi} dx \geq 0
\]
and we have for the entropy dissipation rate (2.21):
\[
R_{\psi}(\rho(t)|\rho_\infty) \geq 2\lambda_3 \int_{\mathbb{R}^n} \psi''(e^\phi \rho)e^{-\phi} U^\top D U dx = -2\lambda_3 I_\psi(\rho(t)|\rho_\infty).
\]

The assertion now follows from
\[
\frac{d}{dt} |I_\psi(\rho(t)|\rho_\infty)| \leq -2\lambda_3 |I_\psi(\rho(t)|\rho_\infty)|.
\]
\[\tag{2.29}\]
□

Lemma 2.4. Let \(P = P^\top \geq 0\), and
\[
P = \begin{pmatrix}
P_{11} & P_{12} \\
P_{21} & P_{22}
\end{pmatrix},
\]
where \(P_{ij}, i, j = 1, 2\) are \(n \times n\) matrices. Then
\[
Q := \begin{pmatrix}
\text{Tr}P_{11} & \text{Tr}P_{12} \\
\text{Tr}P_{21} & \text{Tr}P_{22}
\end{pmatrix} \geq 0.
\]

Proof. Let \(I_j := (I_{kl})_{2 \times 2n}, j = 1, \ldots, n\), where \(I_{1j} = I_{2,n+j} = 1\); the other elements are 0. Then we have
\[
I_j P I_j^\top = \begin{pmatrix}
P_{jj} & P_{j,n+j} \\
P_{n+j,j} & P_{n+j,n+j}
\end{pmatrix} \geq 0.
\]
Hence, \(Q = \sum_{j=1}^n I_j P I_j^\top \geq 0\). □

Next, we shall derive the exponential decay of the relative entropy. For this purpose, we first show the convergence of \(\rho(t)\) to \(\rho_\infty\) in relative entropy (without a rate, for the moment). We remark that the analogous result for the symmetric Fokker-Planck equation was obtained in [4], §2.1 using spectral theory. Specifically, \(\sigma(L_S) \subset \mathbb{R}_0^+\) when considering \(L_S\) in \(L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))\). Hence, \(\|\rho(t)\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))}\) is monotonically decaying also for the non-symmetric Fokker-Planck equation (1.1). And we have the apriori estimate
\[
\|\rho(t)\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))} \leq \|\rho_1\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))}, \quad t \geq 0.
\]
\[\tag{2.30}\]

In contrast, we shall derive it here from the decay of the entropy dissipation:

Theorem 2.5. Let \(\rho_1/\sqrt{\rho_\infty} \in L^2(\mathbb{R}^n), |I_{\psi_1}(\rho_1|\rho_\infty)| < \infty\), and let the coefficients \(\phi(x), F(x),\) and \(D(x)\) satisfy condition (A3). Then
(a) \(e_{\psi_1}(\rho(t)|\rho_\infty) \to 0\) as \(t \to \infty\) for \(1 \leq p < 2\).
(b) If, additionally, \(e_{\psi_{2+\epsilon}}(\rho_1|\rho_\infty) < \infty\) for some \(\epsilon > 0\), then \(e_{\psi_2}(\rho(t)|\rho_\infty) \to 0\).
(c) Let \( e_\phi \) be any admissible relative entropy, and \( e_\phi \) its quadratic superentropy with \( |I_\phi(\rho|\rho_\infty)| < \infty \). Then \( e_\phi(\rho(t)|\rho_\infty) \to 0 \).

**Proof.** First we establish this result for the (logarithmic) physical relative entropy \( e(\rho|\rho_\infty) \). Its entropy dissipation satisfies

\[
|I(\rho(t)|\rho_\infty)| = \int_{\mathbb{R}^n} \frac{\rho_\infty^2}{\rho} \left( \nabla^\top \frac{\rho}{\rho_\infty} D \nabla \frac{\rho}{\rho_\infty} \right) dx
\]

\[
= 4 \int_{\mathbb{R}^n} \left( \nabla^\top \frac{\rho}{\rho_\infty} D \nabla \sqrt{\frac{\rho}{\rho_\infty}} \right) dx.
\]

Since \( D(x) \) is locally uniformly strictly positive definite, \( \rho_\infty > 0 \) and \( \rho_\infty \in L^\infty_{\text{loc}}(\mathbb{R}^n) \), Lemma 2.3 implies that

\[
\nabla \sqrt{\frac{\rho_k}{\rho_\infty}} - c_k(\Omega) \xrightarrow{k \to \infty} 0 \quad \text{in} \quad L^2(\Omega),
\]

(2.31)

with the notation

\[
\rho_k := \rho(t_k) \quad \text{and} \quad c_k(\Omega) := \int_\Omega \sqrt{\rho_k/\rho_\infty} dx/\text{vol}(\Omega).
\]

For any \( \Omega \) fixed, we have

\[
\left\| \sqrt{\frac{\rho_k}{\rho_\infty}} \right\|^2_{L^2(\Omega)} = \int_\Omega \frac{\rho_k}{\rho_\infty} dx \leq C(\Omega) \int_{\mathbb{R}^n} \rho_k dx = C(\Omega).
\]

Thus, \( c_k(\Omega) \) is uniformly bounded with respect to \( k \) for \( \Omega \) fixed. Since \( \rho_\infty \in L^\infty(\Omega) \), (2.31) implies that

\[
\sqrt{\rho_k} - c_k(\Omega) \sqrt{\rho_\infty} \xrightarrow{k \to \infty} 0 \quad \text{in} \quad L^2(\Omega).
\]

Because

\[
\|\rho_k - c_k^2(\Omega)\rho_\infty\|_{L^1(\Omega)} \leq \|\sqrt{\rho_k} - c_k(\Omega)\sqrt{\rho_\infty}\|_{L^2(\Omega)} (\|\sqrt{\rho_k}\|_{L^2(\Omega)} + c_k(\Omega)\|\sqrt{\rho_\infty}\|_{L^2(\Omega)}),
\]

we have

\[
\rho_k - c_k^2(\Omega)\rho_\infty \xrightarrow{k \to \infty} 0 \quad \text{in} \quad L^1(\Omega).
\]

(2.32)

Due to the uniform boundedness of \( c_k(\Omega) \), there exists a subsequence (still denoted by \( \{c_k(\Omega)\} \)) such that

\[
c_k(\Omega) \xrightarrow{k \to \infty} c(\Omega).
\]

Now we choose the domain sequence \( \Omega_N := B_N(0) \subset \mathbb{R}^n \). And take the diagonal subsequence of all \( \{c_k(\Omega_N)\} \) such that for any \( N \) fixed, we have

\[
c_k(\Omega_N) \xrightarrow{k \to \infty} c_N(\Omega_N).
\]

(2.33)
In view of (2.32) and (2.33), we obtain
\[ \rho_k \xrightarrow{k \to \infty} c_N^2(\Omega_N)\rho_\infty \quad \text{in} \quad L^1(\Omega_N). \] (2.34)

Since \( \rho_\infty > 0 \) in \( \mathbb{R}^n \), we conclude that \( c_N(\Omega_N) = c \) for all \( N \). Using (2.30) and the Hölder inequality we have
\[ \int_{\Omega_N} \rho_k \, dx \leq \|\rho_k\|_{L^2(\mathbb{R}^n;\rho_\infty(dx))} \left( \int_{\Omega_N} e^{-\phi} \, dx \right)^{1/2} \xrightarrow{N \to \infty} 0, \quad \text{uniformly in} \ k \in \mathbb{N}. \] (2.35)

Thus
\[ \rho_k \xrightarrow{k \to \infty} c^2 \rho_\infty \quad \text{in} \quad L^1(\Omega_N). \] (2.36)

Due to (1.2), we deduce that \( c = 1 \) and hence
\[ \rho_k \xrightarrow{k \to \infty} \rho_\infty \quad \text{in} \quad L^1(\mathbb{R}^n). \] (2.37)

Therefore
\[ \mu_k := \frac{\rho_k}{\rho_\infty} \rightarrow 1 \quad \text{in measure} \]

(in the measure space \((\mathbb{R}^n, \rho_\infty(dx))\)). The three assertions of the Lemma will now be discussed separately.

**Part (a):**
In order to apply Vitali’s convergence theorem we rewrite
\[ e_{\psi_p}(\rho_k|\rho_\infty) = \frac{1}{p-1} \left[ \|\mu_k\|^p_{L^p(\mathbb{R}^n;\rho_\infty(dx))} - 1 \right], \quad 1 < p < 2. \]

Proceeding as for (2.35) we obtain \( \forall \Omega \subset \mathbb{R}^n: \)
\[ \left| \int_{\Omega} \mu_k^p \rho_\infty \, dx \right| \leq \|\mu_k\|^p_{L^p(\mathbb{R}^n;\rho_\infty(dx))} \left( \int_{\Omega} \rho_\infty \, dx \right)^{1-p/2}. \] (2.37)

And this yields both the uniform integrability of \( \{\mu_k^p\} \) and the uniform decay of its ‘tails’. Thus, Vitali’s theorem yields
\[ \mu_k \xrightarrow{k \to \infty} 1 \quad \text{in} \quad L^p(\mathbb{R}^n;\rho_\infty(dx)), \]

and hence \( e_{\psi_p}(\rho_k|\rho_\infty) \rightarrow 0. \)

For the logarithmic entropy the result follows from \( \psi_1(\sigma) \leq \psi_p(\sigma), \ \sigma \geq 0. \)

**Part (b):**
From (2.20) we obtain the apriori estimate for the \((2 + \varepsilon)-\)entropy:
\[ \frac{1}{1 + \varepsilon} \left[ \|\mu(t)\|_{L^{2+\varepsilon}(\mathbb{R}^n;\rho_\infty(dx))}^2 - 1 \right] \]
\[ = e_{\psi_{2+\varepsilon}}(\rho(t)|\rho_\infty) \leq e_{\psi_{2+\varepsilon}}(\rho_\infty|\rho_\infty), \quad t \geq 0. \]

Now, estimating \( \int_{\Omega} \mu_k^2 \rho_\infty \, dx \) analogously to (2.37) proves the assertion.

**Part (c):**
Here we consider the decay of the quadratic superentropy \( e_\varphi \) that satisfies
\[ 0 \leq e_\varphi(\rho(t)|\rho_\infty) \leq e_\varphi(\rho(t)|\rho_\infty) := \eta_2 \|\rho(t) - \rho_\infty\|^2_{L^2(\mathbb{R}^n;\rho_\infty(dx))}. \] (2.38)
From (2.20) its entropy dissipation satisfies
\[ |I_\psi(\rho(t)|\rho_\infty)| = \eta_2 \int_{\mathbb{R}^n} \left( \nabla^\top \frac{\rho(t)}{\rho_\infty} \mathbf{D} \frac{\rho(t)}{\rho_\infty} \right) \rho_\infty \, dx. \]

A similar analysis as before yields that for any bounded Lipschitz domain \( \Omega \subset \mathbb{R}^n \) and an arbitrary sequence \( t_k \to \infty \), it holds:
\[ \frac{\rho_k}{\rho_\infty} - d_k(\Omega) \xrightarrow{k \to \infty} 0 \quad \text{in} \quad L^2(\Omega) \]
with \( d_k(\Omega) := \int_\Omega \frac{\rho_k}{\rho_\infty} \, dx / \text{vol}(\Omega) \). Since
\[ \int_\Omega \left[ \frac{\rho_k}{\rho_\infty} - 1 \right]^2 \rho_\infty \, dx \leq C(\Omega) \]
(because of (2.30)), \( d_k(\Omega) \) is also uniformly bounded with respect to \( k \) for \( \Omega \) fixed. Now we can take the diagonal subsequence of all \( d_k(\Omega_N) \) such that for any \( N \) fixed, we have
\[ d_k(\Omega_N) \xrightarrow{k \to \infty} d_N \quad \text{(2.39)} \]
and
\[ \frac{\rho_k}{\rho_\infty} \xrightarrow{k \to \infty} d_N(\Omega_N) \quad \text{in} \quad L^2(\Omega_N). \quad \text{(2.40)} \]
From the previous analysis we know that \( d_N(\Omega_N) = c_N^2(\Omega_N) \) and \( d_N = 1 \) for all \( N \). Since \( \rho_\infty > 0 \), (2.40) implies
\[ \int_{\Omega_N} \left( \frac{\rho_k}{\rho_\infty} - 1 \right)^2 \rho_\infty \, dx \xrightarrow{k \to \infty} 0. \quad \text{(2.41)} \]
The monotone decay of \( \|\rho(t)\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))} \) and (2.30) imply that there exists a subsequence (still denoted by \( \{\rho_k\} \)) with
\[ \rho_k \to \tilde{\rho} \quad \text{in} \quad L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx)) \quad \text{and} \quad \|\rho_k\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))} \searrow \tau \geq 0. \]
(2.36) then implies \( \tilde{\rho} = \rho_\infty \) and the weak lower semicontinuity of the norm yields
\[ \tau \leq \|\rho_\infty\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))} = \left( \int \rho_\infty \, dx \right)^{1/2} = 1. \]
Indeed, we have \( \tau = 1 \), since
\[ \forall \varepsilon > 0 : \quad \exists N = N(\varepsilon) \quad \text{with} \quad \|\rho_\infty\|_{L^2(\Omega_N; \rho_\infty^{-1}(dx))} \geq 1 - \varepsilon. \]
The strong convergence (2.41) then implies
\[ \|\rho_k\|_{L^2(\Omega_N; \rho_\infty^{-1}(dx))} \to \|\rho_\infty\|_{L^2(\Omega_N; \rho_\infty^{-1}(dx))} \geq 1 - \varepsilon, \]
and hence
\[ \tau = \lim_{k \to \infty} \|\rho_k\|_{L^2(\Omega_N; \rho_\infty^{-1}(dx))} \geq 1 - \varepsilon. \]
Weak convergence of \( \rho_k \) and convergence of its norms then imply
\[ 0 \leq e_\psi(\rho(t)|\rho_\infty) \leq e_\psi(\rho(t)|\rho_\infty) = \eta_2 \|\rho(t) - \rho_\infty\|_{L^2(\mathbb{R}^n; \rho_\infty^{-1}(dx))}^2 \xrightarrow{k \to \infty} 0. \]
This proves the assertion. \( \square \)
In the above theorem, the assumptions \( \rho_I / \sqrt{\rho_\infty} \in L^2(\mathbb{R}^n) \) and \( |I_{\psi_1}(\rho_I|\rho_\infty)| < \infty \) are perhaps unnaturally restrictive. However, once we have proved a logarithmic Sobolev inequality for \( \rho_I \) smooth with compact support, simple closure yields us the inequality in full generality, and then the conclusion of the Theorem follows immediately without this assumption. Thus, nothing is lost in making this assumption. We then obtain:

**Theorem 2.6.** Let \( e_\psi \) be an admissible relative entropy and \( e_\psi(\rho_I|\rho_\infty) < \infty \). Let the coefficients \( \phi(x), F(x), \) and \( D(x) \) satisfy condition \( (A3) \). Then the relative entropy converges to \( \theta \) exponentially:

\[
e_\psi(\rho(t)|\rho_\infty) \leq e^{-2\lambda t}e_\psi(\rho_I|\rho_\infty), \quad t > 0. \tag{2.42}
\]

Moreover, the convex Sobolev inequality \( (LSI \ for \ \psi = \psi_1) \)

\[
\int_{\mathbb{R}^n} \psi \left( \frac{\rho}{\rho_\infty} \right) \rho_\infty(dx) \leq \frac{1}{2\lambda_3} \int_{\mathbb{R}^n} \nabla^\top h_\psi \left( \frac{\rho}{\rho_\infty} \right) D\nabla h_\psi \left( \frac{\rho}{\rho_\infty} \right) \rho_\infty(dx) \tag{2.43}
\]

with \( h_\psi \) from (2.26) holds

\[
\forall \rho \in L^1_1(\mathbb{R}^n) \quad \text{with} \quad \int_{\mathbb{R}^n} \rho dx = 1. \tag{2.44}
\]

This inequality, of course, does not require our usual normalization \( \int \rho(x)dx = 1 \). Note that \( L^1_1(\mathbb{R}^n) \) in (2.44) can be replaced by \( L^1(\mathbb{R}^n) \) if \( \psi \) is quadratic.

**Proof.** We proceed in two steps and first derive (2.42) for

\[
\rho_I \in S := \{ \rho \in L^2_n(\mathbb{R}^n, \rho_\infty^{-1}(dx)) \mid |I_{\psi_1}(\rho|\rho_\infty)| + |I_\psi(\rho|\rho_\infty)| < \infty \}.
\]

From the Theorem 2.5(c) and Lemma 2.3 we then know that \( e_\psi(\rho(t)|\rho_\infty) \to 0 \) and \( I_{\psi}(\rho(t)|\rho_\infty) \to 0 \) as \( t \to \infty \). Hence, integrating (2.29) (which also holds under condition \( (A3) \)) over \( (t, \infty) \) gives

\[
I_\psi(t) = \frac{d}{dt}e_\psi(t) \leq -2\lambda_3e_\psi(t), \quad t \geq 0, \tag{2.45}
\]

which proves the exponential entropy decay for sufficiently regular initial data.

In explicit terms (2.45) just is the convex Sobolev inequality (2.43) for all sufficiently regular \( \rho_I \). Then, by simple closure (cf. the proof of Corollary 2.18 in [4]), we obtain this inequality in full generality. Once we have this, we no longer need Theorem 2.5 to prove \( e_\psi(\rho(t)|\rho_\infty) \to 0 \), and we obtain the full result. \( \square \)

The desired \( L^1 \)-convergence of \( \rho(t) \) to \( \rho_\infty \) is now a direct consequence of Theorem 2.6 and the Csiszár-Kullback inequality (2.4):

**Corollary 2.7.** Let \( e_\psi \) be an admissible relative entropy and \( e_\psi(\rho_I|\rho_\infty) < \infty \). Let the coefficients \( \phi(x), F(x), \) and \( D(x) \) satisfy condition \( (A3) \). Then the solution of (1.1) satisfies

\[
\|\rho(t) - \rho_\infty\|_{L^1(\mathbb{R}^n)} \leq e^{-\lambda t} \sqrt{\frac{2}{\eta_2} e_\psi(\rho_I|\rho_\infty)}, \quad t > 0, \tag{2.46}
\]

with the notation \( \eta_2 = \psi''(1) \).
3. Examples

In this section we shall construct examples to illustrate how the non-symmetric perturbation $\text{div}(D\rho F)$ can help to “improve” the constant in the LSI (1.7). For simplicity of the presentation we confine ourselves here to the case $D(x) \equiv I$.

Assume that $\phi(x)$ is smooth on $\mathbb{R}^n$ and satisfies

(i) $\nabla \phi(0) = 0$; $\frac{\partial^2 \phi}{\partial x^2}(x) > 0$, $\forall x \neq 0$;

(ii) $\frac{\partial^2 \phi}{\partial x^2} \geq \lambda I > 0$ on $\mathbb{R}^n \setminus B_\delta(0)$ for some (small) $\delta > 0$;

(iii) $\frac{\partial^2 \phi}{\partial x^2}(0) =$

$$
\begin{pmatrix}
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & \frac{\partial^2 \phi}{\partial x^2}(0)
\end{pmatrix},
$$

where $\frac{\partial^2 \phi}{\partial x^2}(0) > 0$. Clearly, this confinement potential $\phi(x)$ satisfies the BEC (A1) only with the convexity constant $\lambda_1 = 0$. Let $\rho_\infty = e^{-\phi(x)}$ be normalized on $\mathbb{R}^n$.

Our goal is to find a vector field $F = (F_1(x), \ldots, F_n(x))^\top$ with $\text{div}(\rho_\infty F) = 0$ such that the generalized Bakry-Emery condition (GBEC) holds, i.e.

$$
\exists \lambda_2 > 0 \text{ such that } \frac{\partial^2 \phi}{\partial x^2} - \frac{1}{2} \left( \frac{\partial F}{\partial x} + \left( \frac{\partial F}{\partial x} \right)^\top \right) \geq \lambda_2 I \quad \forall x \in \mathbb{R}^n.
$$

More precisely, we shall construct $F \in \text{Lip}(\mathbb{R}^n)$ with $\text{supp} F \subset [-L, L]^n$ and $L > 0$ sufficiently small, such that

$$
\frac{\partial F}{\partial x}(0) =
\begin{pmatrix}
-1 & 0 & \cdots & 0 & 0 \\
0 & -1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & -1 & 0 \\
* & * & \cdots & * & n - 1
\end{pmatrix},
$$

where the derivatives $\partial_j F_n(0)$ are yet unspecified (we use the abbreviation $\partial_j := \frac{\partial}{\partial x_j}$). The first principal minors of

$$
G := \frac{\partial^2 \phi}{\partial x^2}(0) - \frac{\varepsilon}{2} \left( \frac{\partial F}{\partial x} + \left( \frac{\partial F}{\partial x} \right)^\top \right)(0)
$$

are $\varepsilon, \ldots, \varepsilon^{n-1}$, and its determinant is of the form

$$
\frac{\partial^2 \phi}{\partial x^2}(0) \varepsilon^{n-1} + \mathcal{O}(\varepsilon^n).
$$

Then, for some $\varepsilon > 0$ sufficiently small, it holds: $\det G > 0$ and $(\phi, \varepsilon F)$ clearly satisfies the GBEC (A2). We remark that $F$ could be chosen as smooth as desired, by using easy modifications of the strategy below.

Now we shall construct two vector fields $F$ and $J = (J_1(x), \ldots, J_{n-1}(x))^\top$ that satisfy
\[
\rho_\infty F = \begin{pmatrix}
\partial_n J_1 \\
\partial_n J_2 \\
\vdots \\
-\partial_1 J_1 - \partial_2 J_2 - \cdots - \partial_{n-1} J_{n-1}
\end{pmatrix}
\]

and hence, \( \text{div}(\rho_\infty F) = 0 \).

For \( j = 1, \ldots, n-1 \) we put
\[
-F_j(x_1, \ldots, x_n) := \begin{cases}
0, & -L \leq x_n \leq -L/2; \\
\frac{f(x_j) \cos(\frac{\pi x_n}{L}) \prod_{k \neq j} \cos^2\left(\frac{x_k \pi}{L}\right)}{\int_{-L/2}^{L/2} \sin^2\left(\frac{2\pi x_k}{L}\right) \rho_\infty(x_1, \ldots, x_n) \, dx_n}, & -L/2 \leq x_n \leq L/2; \\
g_j(x_1, \ldots, x_{n-1}) \sin\left(\frac{\pi x_n}{L}\right), & L/2 \leq x_n \leq L,
\end{cases}
\]
with \( L > 0 \) to be chosen later. \( f(s) \) is a smooth function on \( \mathbb{R} \) with support in \([-L, L]\) and it satisfies

(i) \( f(\pm L) = f'(\pm L) = f''(\pm L) = 0 \);

(ii) \( f \geq 0, f'(0) = 1, f''(0) = 0 \).

Further,
\[
g_j(x_1, \ldots, x_{n-1}) := \frac{\int_{-L}^{L} F_j(x_1, \ldots, x_{n-1}, \tilde{x}_n) \rho_\infty(x_1, \ldots, x_{n-1}, \tilde{x}_n) \, d\tilde{x}_n}{|x_n| \leq L; |x_n| > L},
\]
which implies
\[
\int_{-L}^{L} F_j \rho_\infty \, dx_n = 0, \quad \forall (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}, \quad \forall j = 1, \ldots, n-1. \quad (3.2)
\]

Next we define for \( j = 1, \ldots, n-1 \):
\[
J_j(x_1, \ldots, x_n) := \begin{cases}
\int_{-L}^{L} F_j(x_1, \ldots, x_{n-1}, \tilde{x}_n) \rho_\infty(x_1, \ldots, x_{n-1}, \tilde{x}_n) \, d\tilde{x}_n, & |x_n| \leq L; \\
0, & |x_n| > L.
\end{cases}
\]
Due to (3.2) we have \( J_j \in \text{Lip}(\mathbb{R}^n) \). Finally we put
\[
F_n := -\rho_\infty \sum_{k=1}^{n-1} \partial_k J_k.
\]

In order to verify (3.1), one easily finds for \( j = 1, \ldots, n-1 \):
\[
\partial_j F_j(0) = -f'(0) = -1; \\
\partial_k F_j(0) = 0, \quad k \neq j.
\]
In order to analyze $\partial_n F_n$ we use $\nabla \phi(0) = 0$ in $\text{div}(\rho_\infty F) = 0$ and obtain

$$\rho_\infty(0) \sum_{j=1}^n \partial_j F_j(0) = 0.$$  

Hence

$$\partial_n F_n(0) = -\sum_{j=1}^{n-1} \partial_j F_j(0) = n - 1.$$  

Thus, $\frac{\partial F}{\partial x}(0)$ is of form (3.1) and $(\phi, \varepsilon F)$ satisfies the GBEC (A2) for some (small) $\varepsilon > 0$.

**Appendix: Calculation of the Ricci Tensor**

The definition (2.12) of the Ricci tensor gives (using the Einstein summation convention)

$$U^\top \text{Ric}(x) U = U_i \text{Ric}^{ij} U_j = U_i d^k_i d^l_j \rho_{kl} U_j - U_i d^k_i d^l_j (\nabla^S X)_{kl} U_j =: W_1 + W_2.$$  

Using the definitions (2.7)-(2.9), after a long computation, we have

$$W_1 = U_i d^k_i d^l_j R_{kpjl} U_j$$  

$$= U_i d^k_i d^l_j \left( \partial_p \Gamma^p_{lk} - \partial_l \Gamma^p_{kp} + \Gamma^p_{pm} \Gamma^m_{lk} - \Gamma^p_{lm} \Gamma^m_{pk} \right) U_j$$  

$$= \frac{1}{2} U_i d^y_j d^l_j \partial_p d^m_j U_j - \frac{1}{2} U_i \left( d^k_i d^y_j d^l_j + d^l_j \partial_p d^y_i \right) U_j$$  

$$+ \frac{1}{2} U_i \partial_q d^l_j d^y_i \partial_p d^y_j U_j - \frac{1}{4} U_i \left( d^k_i d^y_j d^l_j \partial d^m_j \partial d^p_j + 2 d^p_j d^l_j \partial_p d^y_j \right) U_j$$  

$$+ 2 d^p_j d^l_j \partial_p d^y_i \partial_m d^k_i d^y_j U_j - 2 d^k_i d^y_j \partial_p d^y_j \partial_d^m_i d^y_j U_j$$  

$$+ \frac{1}{2} U_i \left( d^k_i d^y_j \partial d^p_j \partial d^m_i \partial_d^p_j + \frac{1}{4} d^k_i d^y_j \partial d^p_j \partial d^m_i \partial_d^p_j \right) U_j$$  

$$= U^\top \left[ \frac{1}{2} \text{Tr} \left( D \frac{\partial^2}{\partial x^2} \right) D + \frac{1}{2} (\nabla^T D \nabla) D - D \left( \frac{\partial^2}{\partial x^2} D \right) \right] U$$  

$$- \frac{1}{4} \text{Tr} \left( E^\top + D E D^{-1} - N(D) D^{-1} \right)^2$$  

$$+ \frac{1}{4} U^\top D \left( \frac{\partial (D \nabla \ln(\det D))}{\partial x} + \left( \frac{\partial (D \nabla \ln(\det D))}{\partial x} \right)^\top \right) U$$  

$$- \frac{1}{4} U^\top \left[ (\nabla \ln(\det D))^\top D \nabla \right] D U,$$
where we have used formulas such as
\[ d^{qp} \partial_k d_{qp} = - \partial_k \ln(\det D), \]
\[ d^{ql} \partial_k d_{lm} = - d_{lm} \partial_k d_{ql}. \]

Next we compute \( W_2 \), which involves \( \phi(x) \) and \( F(x) \). We use (2.10) and (2.11) to obtain
\[ W_2 = - \frac{1}{2} U_i d^{ik} \left[ d_{im} \nabla_k X^m + d_{km} \nabla_i X^m \right] d^{jl} U_j \]
\[ = - \frac{1}{2} U_i d^{ik} \left[ d_{im} \left( \partial_k X^m + \Gamma^m_{kp} X^p \right) + d_{km} \left( \partial_i X^m + \Gamma^m_{ip} X^p \right) \right] d^{jl} U_j \]
\[ = - \frac{1}{2} U_i d^{ik} \left( d_{im} \partial_k X^m + d_{km} \partial_i X^m \right) d^{jl} U_j \]
\[ = - \frac{1}{2} U_i d^{ik} \left( d_{im} \Gamma^m_{kp} X^p + d_{km} \Gamma^m_{ip} X^p \right) d^{jl} U_j \]
\[ =: V_1 + V_2. \]

From (2.6), we have
\[ V_1 = \frac{1}{2} U_i d^{ik} \left[ d_{im} \left( \partial_q \left( \phi \left( \frac{1}{2} \ln(\det D) \right) - F_q \right) \right) \right] \]
\[ + d_{km} \left( \partial_q \left( \phi \left( \frac{1}{2} \ln(\det D) \right) - F_q \right) \right) \right] d^{jl} U_j \]
\[ = \frac{1}{2} \left( U^\top D E (\nabla \phi - F) + (\nabla \phi - F)^\top E^\top D U \right) \]
\[ + U^\top D \left( \frac{\partial^2 \phi}{\partial x^2} - \frac{1}{2} \left( \frac{\partial F}{\partial x} + \left( \frac{\partial F}{\partial x} \right)^\top \right) \right) \]
\[ - \frac{1}{4} U^\top D \left( \frac{\partial \left( D \nabla \ln(\det D) \right)}{\partial x} + \left( \frac{\partial \left( D \nabla \ln(\det D) \right)}{\partial x} \right)^\top \right) U. \]

From (2.6) and (2.7), we have
\[ V_2 \]
\[ = - \frac{1}{2} U_i d^{pq} d^{kl} \left( \partial_q \left( \phi \left( \frac{1}{2} \ln(\det D) \right) - F_q \right) \right) U_j \]
\[ = - \frac{1}{2} \left[ \nabla \left( \phi \left( \frac{1}{2} \ln(\det D) \right) - F \right) \right]^\top D E U \]
\[ = - \frac{1}{2} U^\top \left[ \left( \nabla \phi - F \right)^\top D \nabla \right] U + \frac{1}{4} U^\top \left( \left( \nabla \ln(\det D) \right)^\top D \nabla \right) D U. \]

Hence, the GBEC (A3) can be written as
\[ W_1 + W_2 = W_1 + V_1 + V_2 \geq \lambda_3 U^\top D U, \]
which is exactly (2.13).

Acknowledgment. A. Arnold was partially supported by the ESF in the project “Global and geometrical aspects of nonlinear partial differential equations”, the Wissenschaftskolleg Differentialgleichungen of the FWF, and by the DFG under Grant no. AR 277/3-3. E. Carlen was partially supported by U.S. National Science
Foundation grant DMS 06-00037. Q. Ju was supported by the NNSFC grant no. 10701011, the Wittgenstein 2000 Award of Peter Markowich (funded by the Austrian FWF), and by the Austrian-Chinese Technical-Scientific Collaboration Agreement funded by the OAD. The authors acknowledges fruitful discussions with Peter Markowich and Cedric Villani.

References


**Anton Arnold:** Institut für Analysis und Scientific Computing, Technische Universität Wien, Wiedner Hauptstr. 8, A-1040 Wien, Austria  
*E-mail address:* anton.arnold@tuwien.ac.at

**Eric Carlen:** Department of Mathematics, Hill Center, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854, USA  
*E-mail address:* carlen@math.rutgers.edu

**Qiangchang Ju:** Institute of Applied Physics and Computational Mathematics, P.O. Box 8009-28, Beijing, 100088, China  
*E-mail address:* qiangchang_ju@yahoo.com