ENTROPY MINIMIZATION AND SCHRÖDINGER PROCESSES IN INFINITE DIMENSIONS

By Hans Föllmer and Nina Gantert

Humboldt-Universität and Technische Universität Berlin

Schrödinger processes are defined as mixtures of Brownian bridges which preserve the Markov property. In finite dimensions, they can be characterized as h-transforms in the sense of Doob for some space–time harmonic function h of Brownian motion, and also as solutions to a large deviation problem introduced by Schrödinger which involves minimization of relative entropy with given marginals. As a basic case study in infinite dimensions, we investigate these different aspects for Schrödinger processes of infinite-dimensional Brownian motion. The results and examples concerning entropy minimization with given marginals are of independent interest.

1. Introduction. Let P denote the distribution of a Brownian motion X_t, $0 \leq t \leq 1$, with state space $S = R^d$ and initial distribution μ_0. The conditional distribution with respect to an initial value $X_0 = x$ and a terminal value $X_1 = y$ is given by the Brownian bridge P_x^y from x to y, and so we can write

$$(1.1) \qquad P = \int P_x^y \mu(\, dx, \, dy),$$

where μ denotes the joint distribution of (X_0, X_1) under P. A measure Q on $C[0, 1]$ has these same conditional distributions with respect to X_0 and X_1 if and only if it is of the form

$$(1.2) \qquad Q = \int P_x^y \nu(\, dx, \, dy)$$

for some probability measure ν on $S \times S$. Such a mixture of Brownian bridges will be called a *Schrödinger process* if it preserves the Markov property of P, that is, if X_t, $0 \leq t \leq 1$, is again a Markov process under Q.

Schrödinger (1931) initiated the study of these processes by considering the following problem of large deviations. Let ν_0 and ν_1 be distributions on S, and look for the most likely behavior of a large collection of independent Brownian motions governed by P under the constraint that the empirical distributions at times $t = 0$ and $t = 1$ are close to ν_0 and to ν_1. In the limit of an infinite particle system, this behavior can be described by independent

---

motions, each governed by a process Q of the form (1.2) where the representing measure ν minimizes the relative entropy $H(\nu|\mu)$ under the constraint that the marginals of ν are given by ν_0 and ν_1; for a precise statement and a proof based on Sanov's theorem, see Föllmer (1988) or Dawson, Gorostiza and Wakolbinger (1990). The density of the entropy minimizing measure $\nu \approx \mu$ admits a factorization of the form

$$(1.3) \qquad \frac{d\nu}{d\mu}(x, y) = f(x)g(y),$$

where f and g are nonnegative measurable functions on S. Such a measurable factorization implies that Q has the structure of an *h-path process* in the sense of Doob (1984), where h is a space–time harmonic function of Brownian motion. In particular, we obtain the Markov property of Q, and so Q is indeed a Schrödinger process as defined above.

For a measure $\nu \approx \mu$ with $H(\nu|\mu) < \infty$, the Markov property of the associated process $Q \approx P$ is, in fact, equivalent both to the factorization (1.3) of the density and to the property that ν minimizes relative entropy subject to given marginals. Some of these implications are also valid if Brownian motion on R^d is replaced by a general Markov process, but not all of them. The difficult part is to derive the factorization (1.3), either from the Markov property of Q or from the minimization of relative entropy. Positive results have been proved only under strong regularity conditions on the underlying Markov process. These conditions were motivated by finite-dimensional diffusion theory. In this paper, our main purpose is to throw some light on the structure of Schrödinger processes in infinite dimensions. At the same time, we clarify some questions concerning entropy minimization subject to given marginals; these results may be of independent interest.

In Section 2 we recall some basic facts about Schrödinger processes. Section 3 can be read without regard to our discussion of Schrödinger processes. Here we discuss the structure of the unique measure $\nu \approx \mu$ on a product space $S \times S$ which minimizes $H(\cdot|\mu)$ under the constraint that the marginals are fixed. In general, the minimizing measure can be characterized by the fact that the density can be approximated by functions of the form (1.3). However, the passage from this approximation to a factorization of the density involves a delicate closure property of the space of product functions under almost sure convergence. An argument of Rüschendorf and Thomsen (1993) yields the existence and essential uniqueness of the factorization (1.3) under the strong regularity condition

$$(1.4) \qquad \mu \ll \mu_0 \otimes \mu_1.$$

In view of infinite dimensions, this condition is much too restrictive. On the other hand, the first counterexample in Section 5 shows that, without some kind of regularity assumption, there may not exist any measurable factorization of the form (1.3), even if we assume that the density is bounded both from above and away from 0.

As a basic case study in infinite dimensions, we consider the distribution P of a Brownian motion with state space $S = R^\infty$ and some initial distribution μ_0 on S. Alternatively, we can choose the state space $S = C[0, 1]$ and view P as the distribution of a Brownian sheet. The standard regularity assumptions used in the literature on Schrödinger processes no longer hold for P. On the other hand, P exhibits as much regularity as one can hope for in the context of infinite-dimensional stochastic analysis. In Section 4 we show that the Markov property of a Schrödinger process $Q \approx P$ does imply a factorization of the form (1.3) if the density of Q is bounded from above. It follows that Q is an h-process in the sense of Doob, that is,

$$(1.5) \qquad \left.\frac{dQ}{dP}\right|_{\mathscr{F}_t} = \frac{d\nu_0}{d\mu_0}(X_0)\frac{h(X_t, t)}{h(X_0, 0)},$$

where h is some space–time harmonic function of infinite-dimensional Brownian motion. The proof combines regularity properties of the finite-dimensional projections of P with a passage to the limit where we use martingale arguments and the special product structure of P. These methods also allow us to clarify the role of entropy minimization in this infinite-dimensional context, beyond the restriction (1.4). For a process Q defined in terms of $\nu \approx \mu$ as in (1.2), we show that minimization of the relative entropy, the Markov property and the factorization of the density are all equivalent if the density of ν is bounded from above. On the other hand, the second counterexample in Section 5 shows that some strong restriction on the density is really needed. We construct a Schrödinger process Q of the Brownian sheet which minimizes the relative entropy subject to given marginals and whose density belongs to L^p for any $p < \infty$, but which does not admit a measurable factorization of the form (1.3).

Throughout this paper we limit the discussion to the case $\nu \approx \mu$, where μ is a given reference measure. For Brownian motion in finite dimensions, the Markov property of a process Q of the form (1.2) with a general measure ν can be characterized by a splitting property of ν which generalizes the factorization (1.3); see Jamison (1974). In the case of infinite-dimensional Brownian motion, the problem of characterizing a general Schrödinger process by some splitting property of its representing measure is still open. A satisfactory solution should involve the spatial Gibbs structure of ν viewed as a random field on $(R^2)^\infty$, and some partial results in this direction have been obtained by Brockhaus (1995). The problem disappears if the initial distribution is concentrated on one point. In that case, the structure of Schrödinger processes is described by the parabolic Martin boundary on a Wiener space; see Föllmer (1991).

2. Preliminaries. Let $(X_t)_{0 \le t \le 1}$ be a Markov process with Polish state space S and transition probabilities $P_{s,t}(x, dy)$, defined on a probability space $(\Omega, \mathscr{F}, P)$ with filtration $(\mathscr{F}_t)_{0 \le t \le 1}$. We denote by

$$(2.1) \qquad \mu_t = P \circ X_t^{-1}, \qquad 0 \le t \le 1,$$

the marginal distributions on S and by

$$(2.2) \qquad \mu = P \circ (X_0, X_1)^{-1}$$

the joint distribution of (X_0, X_1) under P.

For a probability measure $\nu \approx \mu$ on $S \times S$ with density

$$(2.3) \qquad \varphi = \frac{d\nu}{d\mu} > 0 \quad \mu\text{-a.s.,}$$

consider the associated probability measure $Q \approx P$ on $(\Omega, \mathcal{F})$ defined by

$$(2.4) \qquad \frac{dQ}{dP} = \varphi(X_0, X_1)$$

and denote by

$$(2.5) \qquad \nu_t = Q \circ X_t^{-1}, \qquad 0 \le t \le 1,$$

the induced marginal distributions. The measure Q has the same conditional distribution given the initial value X_0 and the terminal value X_1 as the underlying measure P. For any $t \in [0, 1]$ the density of Q with respect to P on $\mathcal{F}_t$ is given by

$$(2.6) \qquad \frac{dQ}{dP}\bigg|_{\mathcal{F}_t} = \varphi_{0, t}(X_0, X_t),$$

where $\varphi_{0, t}$ is defined by

$$(2.7) \qquad \varphi_{0, t}(x, y) = \int \varphi(x, z) P_{t, 1}(y, dz).$$

This implies that, for $0 \le s \le t \le 1$, the conditional distribution of X_t with respect to $\mathcal{F}_s$ under Q is given by

$$(2.8) \qquad Q[X_t \in dy | \mathcal{F}_s] = (\varphi_{0, s}(X_0, X_s))^{-1} P_{s, t}(X_s, dy) \varphi_{0, t}(X_0, y).$$

Thus, a prediction given $\mathcal{F}_s$ will, in general, involve both the present state X_s and the initial state X_0, and so we cannot expect, in general, that (X_t) is again a Markov process under Q.

DEFINITION 2.9.   The process Q will be called a *Schrödinger process* if $(X_t)_{0 \le t \le 1}$ has the Markov property under Q.

Let us introduce a *sufficient* condition for Q to be a Schrödinger process. Suppose that the density φ of ν with respect to μ admits a factorization of the form

$$(2.10) \qquad \varphi(x, y) = f(x) g(y),$$

where f and g are nonnegative measurable functions on S. Note that (2.3) implies $f > 0$ μ_0-a.s. and $g > 0$ μ_1-a.s. The nonnegative measurable function h on $S \times [0, 1]$ defined by

$$(2.11) \qquad h(x, t) = \int g(y) P_{t, 1}(x, dy)$$

is *space–time harmonic* in the sense that

$$(2.12) \qquad h(x, s) = \int h(y, t) P_{s, t}(x, dy)$$

for $0 \leq s \leq t \leq 1$. In terms of h, the density of Q with respect to P on F_t can be expressed as

$$(2.13) \qquad \varphi_{0, t} = f(X_0) h(X_t, t) = \frac{d\nu_0}{d\mu_0}(X_0) \frac{h(X_t, t)}{h(X_0, 0)}.$$

This shows that Q has the structure of an *h-path process* in the sense of Doob (1984), page 566; cf. Jamison (1975). In particular, $(X_t)_{0 \leq t \leq 1}$ is a Markov process under Q with transition probabilities

$$(2.14) \qquad Q_{s, t}(x, dy) = h(x, s)^{-1} P_{s, t}(x, dy) h(y, t).$$

We have thus established the well-known fact [see, e.g., Jamison (1974)] that a measurable factorization of the density implies the Markov property of Q:

PROPOSITION 2.15. *Let $Q \approx P$ be a process of the form* (2.4). *If the density φ admits a measurable factorization of the form* (2.10), *then Q is a Schrödinger process.*

Let us now turn to the question to which extent the converse holds. As shown by Jamison (1974), the Markov property of Q does imply a factorization (2.10) if the process $(X_t)_{0 \leq t \leq 1}$ has strictly positive transition densities

$$(2.16) \qquad p_{s, t}(x, y) > 0, \qquad x, y \in S, 0 \leq s < t \leq 1,$$

with respect to some fixed reference measure λ on S; this will also follow from Proposition 2.25 below. On the other hand, it is easy to construct degenerate examples of Schrödinger processes such that the factorization (2.10) fails to hold.

EXAMPLE 2.17. Let $S = R^2$ and $P = \mu = \mu_0 \otimes \mu_1$ on $\Omega = S \times S$ with

$$(2.18) \qquad \mu_0 = \tfrac{1}{2}(\delta_{(0, 1)} + \delta_{(0, -1)}), \qquad \mu_1 = \tfrac{1}{2}(\delta_{(1, 0)} + \delta_{(-1, 0)})$$

and define $X_t = X_0(1 - t) + tX_1$, $0 \leq t \leq 1$. For a given $\gamma \in (0, 1)$, take any $0 < \alpha, \beta < 1$ such that $\gamma = \tfrac{1}{2}(\alpha + \beta)$, and define $\nu \approx \mu$ in terms of the density

$$(2.19) \qquad \varphi(x, y) = \begin{cases} 2\alpha, & \text{if } x = (0, +1), y = (+1, 0), \\ 2(1 - \alpha), & \text{if } x = (0, +1), y = (-1, 0), \\ 2\beta, & \text{if } x = (0, -1), y = (+1, 0), \\ 2(1 - \beta), & \text{if } x = (0, -1), y = (-1, 0). \end{cases}$$

The marginals are given by

$$(2.20) \qquad \nu_0 = \mu_0, \qquad \nu_1 = \gamma\delta_{(1, 0)} + (1 - \gamma)\delta_{(-1, 0)}.$$

For any such choice of α, β, the associated process $Q = \nu$ is a Schrödinger process; the Markov property follows from the fact that the full trajectory can be reconstructed from the value X_t at any time $t \in (0, 1)$. However, it is easy to check that a factorization of the density is only possible for $\alpha = \beta = \gamma$. The example also shows that a Schrödinger process may not be uniquely determined by its marginal distributions ν_0 and ν_1.

The following *weak factorization* for a Schrödinger process $Q \approx P$ holds in full generality, and this will be the starting point for our analysis of Schrödinger processes of infinite-dimensional Brownian motion. For a process $Q \approx P$ given by (2.4), we denote by φ_t the density of ν_t with respect to μ_t, by $\varphi_{t,1}$ the density of the joint distribution of (X_t, X_1) under Q with respect to the joint distribution under P and by

$$(2.21) \qquad \psi_{t,1}(x, y) = \varphi_{t,1}(x, y)\varphi_t(x)^{-1}$$

the conditional density of X_1 with respect to X_t.

LEMMA 2.22.   *A process $Q \approx P$ of the form* (2.4) *is a Schrödinger process if and only if the density φ admits for each $t \in (0, 1)$ a factorization*

$$(2.23) \qquad \varphi(X_0, X_1) = \tilde{\varphi}_{0,t}(X_0, X_t)\tilde{\psi}_{t,1}(X_t, X_1), \quad \text{P-a.s.,}$$

where $\tilde{\varphi}_{0,t}$ and $\tilde{\psi}_{t,1}$ are nonnegative measurable functions on $S \times S$. In this case, we can use in particular the functions $\varphi_{0,t}$ and $\psi_{t,1}$ defined above.

PROOF.   The Markov property of Q clearly implies (2.23) with $\tilde{\varphi}_{0,t} = \varphi_{0,t}$ and $\tilde{\psi}_{t,1} = \psi_{t,1}$. Conversely, any measurable factorization of the form (2.23) implies that the conditional distribution (2.8) of Q only involves the present state X_s, and this amounts to the Markov property of Q. □

It is easy to check that condition (2.16) implies that the conditional distribution $\mu[\cdot | X_t = z]$ of (X_0, X_1) given $X_t = z$ satisfies

$$(2.24) \qquad \mu \ll \mu[\cdot | X_t = z] \quad \text{for } \mu_t\text{-a.a. } z \in S.$$

Under the regularity condition (2.24) for the underlying Markov process P, Proposition 2.15 does admit a converse; that is, the Markov property of a process $Q \approx P$ of the form (2.4) implies a factorization of the density.

PROPOSITION 2.25.   *If condition* (2.24) *holds for some $t \in (0, 1)$, then any Schrödinger process $Q \approx P$ admits a measurable factorization* (2.10).

PROOF.   Fit $t \in (0, 1)$. The Markov property of Q implies the weak factorization (2.23). We can conclude that, for μ_t-a.a. $z \in S$,

$$(2.26) \qquad \varphi(x, y) = \varphi_{0,t}(x, z)\psi_{t,1}(z, y)$$

for $\mu[\cdot | X_t = z]$-a.a. (x, y). Now choose $z \in S$ such that conditions (2.24) and (2.26) are both satisfied. This implies that the strong factorization (2.10) holds μ-a.s., where the measurable functions f and g on S are defined by

$$(2.27) \qquad f(x) = \varphi_{0,t}(x, z), \qquad g(y) = \psi_{t,1}(z, y). \qquad \square$$

REMARK 2.28. (a) A measurable factorization of the form (2.10) implies that the functions f and g solve Schrödinger's system of equations

(2.29)
$$\frac{d\nu_0}{d\mu_0}(x) = f(x)\int g(y) P_{0,1}(x, dy),$$
$$\frac{d\nu_1}{d\mu_1}(y) = g(y)\int f(x)\hat{P}_{1,0}(y, dx),$$

where $\hat{P}_{1,0}(y, \cdot)$ denotes the conditional distribution of X_0 under P, given $X_1 = y$; cf. Schrödinger (1931, 1932). Conversely, any pair of nonnegative measurable functions f and g solving (2.29) induces via (2.10) a probability measure $\nu \approx \mu$ on $S \times S$ with given marginals ν_0 and ν_1. The existence and uniqueness of a pair of solutions of (2.29) with given marginals $\nu_i \approx \mu_i$, $i = 0, 1$, have been discussed by Bernstein (1932), Fortet (1940), Beurling (1960) and Jamison (1974) under strong regularity conditions such as (2.16). For a discussion of (2.29) from a physical point of view, we refer to Schrödinger (1931, 1932) and, for example, Nagasawa (1993) and Aebi (1996) and the references to the recent literature on stochastic mechanics given there. For connections to the stochastic calculus of variations, see also, for example, Thieullen (1993) and Thieullen and Zambrini (1995). The classical DAD problem can be formulated in terms of Schrödinger's system (2.29) on a finite state space. From this point of view, various generalizations are discussed in Borwein, Lewis and Nussbaum (1994).

(b) Condition (2.24) implies

(2.30)
$$\mu \ll \mu_0 \otimes \mu_1.$$

To see this, note that the Markov property of P implies that the conditional distribution in (2.24) can be chosen as a product measure on $S \times S$. However, if $\mu \ll \nu_0 \otimes \nu_1$ for any product measure, then it is easy to check that μ is also absolutely continuous with respect to the product of its own marginals.

(c) As shown by Example 2.17 where $\mu = \mu_0 \otimes \mu_1$, condition (2.30) is not strong enough to guarantee that a Schrödinger process $Q \approx P$ admits a factorization of its density.

Regularity conditions such as (2.16) and (2.24) are motivated by finite-dimensional diffusion theory. In the context of infinite-dimensional stochastic analysis, they would be much too strong. As a case study in infinite dimensions, we are going to look in Section 4 at the factorization problem for infinite-dimensional Brownian motion. Before we turn to this special setting, let us first clarify the relation between the factorization problem and Schrödinger's original problem of minimizing relative entropy under the constraint that the marginals at times 0 and 1 are fixed.

3. Minimizing relative entropy with given marginals. Let μ be a probability measure on $S \times S$, where S denotes our Polish state space. We fix two probability measures ν_0 and ν_1 on S and denote by $M(\nu_0, \nu_1)$ the convex set of probability measures ν on $S \times S$ whose marginals coincide with ν_0

and ν_1. Let us assume that there exists a measure $\nu \in M(\nu_0, \nu_1)$ such that

(3.1) $\nu \approx \mu$ and $H(\nu|\mu) < \infty$,

where

(3.2) $H(\nu|\mu) = \begin{cases} \int \log \dfrac{d\nu}{d\mu} \, d\nu, & \text{if } \nu \ll \mu, \\ \infty, & \text{otherwise} \end{cases}$

denotes the *relative entropy* of a probability measure ν with respect to μ. As shown by Csiszár (1975), this assumption implies the existence and uniqueness of a measure $\nu^* \in M(\nu_0, \nu_1)$ which minimizes the relative entropy with respect to μ under the constraint that the marginals are given by ν_0 and ν_1:

(3.3) $H(\nu^*|\mu) = \min\{H(\nu|\mu)|\nu \in M(\nu_0, \nu_1)\}$.

Moreover, we have

(3.4) $\nu^* \approx \mu$.

Let φ^* denote the density of ν^* with respect to μ. The proof of Theorem 3.1 in Csiszár (1975) shows that $\log \varphi^*$ belongs to the closure in $L^1(\nu^*)$ of the space of sums

(3.5) $S(\nu_0, \nu_1) = \{a \oplus b | a \in L^1(\nu_0), \, b \in L^1(\nu_1)\}$,

where $a \oplus b$ denotes the measurable function on $S \times S$ defined by $(a \oplus b)(x, y) = a(x) + b(y)$. The minimizing measure ν^* can actually be characterized by such an approximation of the density.

PROPOSITION 3.6.   *Consider a measure $\nu \in M(\nu_0, \nu_1)$ which satisfies condition* (3.1), *and let φ denote the density of ν with respect to μ. The following properties are equivalent*:

(i) *ν coincides with the minimizing measure ν^*.*
(ii) *There are nonnegative measurable functions f_n and g_n on S, $n = 1, 2, \ldots$, such that*

(3.7) $\int f_n(x) g_n(y) \mu(dx, dy) = 1$,

(3.8) $\log f_n \in L^1(\nu_0), \qquad \log g_n \in L^1(\nu_1)$

and

(3.9) $\lim_n (\log f_n(x) + \log g_n(y)) = \log \varphi(x, y)$ *in* $L^1(\nu)$.

PROOF.   (i) For $\nu = \nu^*$ the result of Csiszár (1975) page 153, shows that $\log \varphi^*$ belongs to the closure of $S(\nu_0, \nu_1)$ in $L^1(\nu^*)$. In addition, the argument preceding (3.48) below shows that the approximating functions $\varphi_n(x, y) = f_n(x)g_n(y)$ can be chosen as densities of probability measures ν_n which minimize $H(\cdot|\mu)$ under finitely many constraints of the form $\int h_i \, d\nu_n = c_i$. Thus, it is no loss of generality to assume the normalization (3.7).

(ii) Suppose that $\log \varphi$ admits an approximation of the form (3.9) and define the probability measures ν_n in terms of the densities

$$(3.10) \qquad \frac{d\nu_n}{d\mu}(x, y) = f_n(x) g_n(y), \qquad n = 1, 2, \ldots .$$

Condition (3.8) implies that the product is strictly positive ν-a.s., hence μ-a.s., and so we have $\nu_n \approx \mu$. Let us now take any $\tilde{\nu} \in M(\nu_0, \nu_1)$. We want to show that $H(\tilde{\nu}|\mu) \geq H(\nu|\mu)$, and so we can assume $H(\tilde{\nu}|\mu) < \infty$. This implies $\tilde{\nu} \ll \mu \approx \nu_n$ and

$$(3.11) \qquad \frac{d\tilde{\nu}}{d\mu} = \frac{d\tilde{\nu}}{d\nu_n} \frac{d\nu_n}{d\mu}$$

μ-a.s., hence $\tilde{\nu}$-a.s. Taking logarithms and integrating with respect to $\tilde{\nu}$, we obtain

$$(3.12) \qquad H(\tilde{\nu}|\mu) = H(\tilde{\nu}|\nu_n) + \int \log(f_n g_n) \, d\tilde{\nu}.$$

Due to condition (3.8) we can separate the integral and write

$$H(\tilde{\nu}|\mu) = H(\tilde{\nu}|\nu_n) + \int \log f_n \, d\nu_0 + \int \log g_n \, d\nu_1$$

$$(3.13)$$

$$= H(\tilde{\nu}|\nu_n) + \int \log(f_n g_n) \, d\nu,$$

since both $\tilde{\nu}$ and ν belong to $M(\nu_0, \nu_1)$. However, $H(\tilde{\nu}|\nu_n) \geq 0$, and so we can use (3.9) to conclude

$$H(\tilde{\nu}|\mu) \geq \lim_n \int \log(f_n g_n) \, d\nu$$

$$(3.14)$$

$$= \int \log \varphi \, d\nu = H(\nu|\mu).$$

This shows that ν coincides with the minimizing measure ν^* defined by (3.3). $\qquad\square$

Let us now show that our criterion for $\nu = \nu^*$ is satisfied if the density of ν admits a measurable factorization. Note that we do not assume the integrability conditions $\log f \in L^1(\nu_0)$ and $\log g \in L^1(\nu_1)$ which are used, for example, in Corollary (3.1) of Csiszár (1975); see also (3.52) below.

COROLLARY 3.15. *Consider a measure $\nu \in M(\nu_0, \nu_1)$ which satisfies condition* (3.1). *If the density φ admits a factorization*

$$(3.16) \qquad \varphi(x, y) = f(x) g(y), \quad \nu\text{-a.s.},$$

with nonnegative measurable functions f and g on S, then ν coincides with the minimizing measure ν^*.

PROOF.   Define $f_n = c_n((f \wedge n) \vee 1/n)$, $g_n = (g \wedge n) \vee 1/n$, $\varphi_n(x, y) = f_n(x) g_n(y)$, where c_n is a normalizing factor such that $\int \varphi_n \, d\mu = 1$. The two separate integrability conditions in (3.8) are clearly satisfied. Since $\varphi \log \varphi$ is bounded from below, the finite entropy condition $H(\nu|\mu) < \infty$ is equivalent to

$$(3.17) \qquad\qquad \varphi \log \varphi \in L^1(\mu) \quad \text{or} \quad \log \varphi \in L^1(\nu).$$

Observe that $\varphi_n/c_n > 1$ implies $f > 1/n$ and $g > 1/n$, thus $\varphi_n/c_n \leq \varphi$. Similarly, $\varphi_n/c_n < 1$ implies $f < n$ and $g < n$, thus $\varphi_n/c_n \geq \varphi$. Therefore, we have $|\log(\varphi_n/c_n)| \leq |\log \varphi|$. Since $\lim_n c_n = 1$ and $\lim \varphi_n = \varphi$ ν-a.s., we can apply Lebesgue's theorem to get

$$(3.18) \qquad\qquad \lim_n \log \varphi_n = \log \varphi \quad \text{in } L^1(\nu).$$

Thus, we have checked our criterion in (3.6) for $\nu = \nu^*$. $\square$

It is sometimes taken for granted that Corollary 3.15 admits a converse, since the approximation (3.9) seems to suggest that the density φ^* of the minimizing measure ν^* admits a measurable factorization of the form (3.16) ν^*-a.s., hence μ-a.s.; see, for example, the "if" part of Corollaries (3.1) and (3.2) in Csiszár (1975). However, this is a delicate point. In fact, our first counterexample in Section 5 will show that the converse may fail to hold, even if we assume that the density φ is both bounded and bounded away from 0.

Thus, additional regularity properties are needed in order to obtain a closure property of the space of sums. The following proposition is a result in this direction. The existence part is due to Rüschendorf and Thomsen (1993). For $\nu \approx \nu_0 \otimes \nu_1$ it reduces to Lemma 2.5 in Donsker and Varadhan (1975). Part (i) of Proposition 3.19 also follows from the results of Borwein and Lewis (1992). We include a proof because we want to illustrate at which point the construction of a measurable factorization breaks down in the context of our counterexamples in Section 5.

PROPOSITION 3.19.   *Let ν be a probability distribution on $S \times S$ with marginals ν_0 and ν_1. Let $(a_n), (b_n)$ be sequences of real-valued measurable functions on S such that*

$$(3.20) \qquad\qquad \lim_n (a_n(x) + b_n(y)) = c(x, y), \quad \nu\text{-a.s.},$$

where c is a finite measurable function on $S \times S$.

(i) *There are real-valued functions a and b on S such that the decomposition*
$$(3.21) \qquad\qquad c(x, y) = a(x) + b(y)$$
holds ν-a.s.

(ii) *If ν satisfies the regularity condition*

$$(3.22) \qquad\qquad \nu \ll \nu_0 \otimes \nu_1,$$

then c admits a measurable decomposition, that is, the functions a and b in (3.21) can be chosen to be measurable.

(iii) *Under condition (3.22) the decomposition is essentially unique up to additive constants. More precisely, there is a countable number of disjoint measurable rectangles $A_i \times B_i$ with $\mu[\cup_i(A_i \times B_i)] = 1$ such that $a(\cdot)$ is uniquely determined up to an additive constant ν_0-a.s. on each A_i and $b(\cdot)$ is uniquely determined up to an additive constant ν_1-a.s. on each B_i.*

PROOF. The existence proof is a modification of the argument in Rüschendorf and Thomsen (1993).

(i) We introduce the measurable set

$$(3.23) \qquad A = \left\{ (x, y) \in S \times S | \exists \lim_n (a_n(x) + b_n(y)) \in R^1 \right\}$$

and its measurable sections $A_x = \{y | (x, y) \in A\}$ for $x \in S$. Let us show that, for any $x, x' \in S$, we have either

$$(3.24) \qquad A_x = A_{x'} \quad \text{or} \quad A_x \cap A_{x'} = \varnothing.$$

Suppose that there is some $z \in A_x \cap A_{x'}$ and take $y \in A_x$. In order to check that $y \in A_{x'}$, note that

$$(3.25) \quad \begin{aligned} a_n(x') &+ b_n(y) \\ &= (a_n(x') + b_n(z)) - (b_n(z) + a_n(x)) + (a_n(x) + b_n(y)) \end{aligned}$$

does converge to a finite limit.

(ii) Let $\nu(x, \cdot)$ denote the conditional distribution of the second coordinate y under ν, given the first coordinate x. The set

$$(3.26) \qquad S_0 = \left\{ x \in S | \nu(x, A_x) = 1 \right\}$$

is measurable and satisfies $\nu_0(S_0) = 1$, and for any $x \in S_0$ we have $A_x \neq \varnothing$. In the same way, we define a measurable set $S_1 \subseteq S$ with $\nu_1(S_1) = 1$ such that the measurable sections $A^y = \{x \in S | (x, y) \in A\}$ satisfy $A^y \neq \varnothing$ for any $y \in S_1$. Due to (3.24) we can define an equivalence relation $\sim$ on S_0 by

$$(3.27) \qquad x \sim x' \quad \text{iff} \quad A_x = A_{x'}.$$

Let $E(x) \subseteq S_0$ be the equivalence class of $x \in S_0$. For any choice of $y \in A_x$, we can write

$$(3.28) \qquad E(x) = \left\{ x' \in S_0 | \exists \lim(a_n(x') + b_n(y)) \in R^1 \right\},$$

and so $E(x)$ is a measurable subset of S_0.

(iii) Using the axiom of choice, we choose a system of representatives containing exactly one member of each equivalence class. Define the functions $\tilde{a}_n$ by

$$(3.29) \qquad \tilde{a}_n(x) = a_n(x) - a_n(\bar{x}),$$

where $\bar{x}$ is the representative of $E(x)$. For $x \in S_0$ and $y \in A_x$, we have

$$(3.30) \qquad \tilde{a}_n(x) = (a_n(x) + b_n(y)) - (a_n(\bar{x}) + b_n(y)).$$

The first term on the right-hand side converges since $y \in A_x$, the second term converges since $y \in A_{\bar{x}} = A_x$, and so there exists a finite limit

$$(3.31) \qquad\qquad a(x) = \lim_n \tilde{a}_n(x), \qquad x \in S_0.$$

For $y \in S_1$ there is exactly one representative $\bar{x}(y)$ such that $y \in A_{\bar{x}(y)}$. Let us define the functions $\tilde{b}_n$ on S_1 by

$$(3.32) \qquad\qquad \tilde{b}_n(y) = b_n(y) + a_n(\bar{x}(y))$$

and note, as in (3.31), that there exists a finite limit

$$(3.33) \qquad\qquad b(y) = \lim_n \tilde{b}_n(y), \qquad y \in S_1.$$

For any $(x, y) \in A \cap (S_0 \times S_1)$, we can use (3.31) to write

$$(3.34) \qquad \lim(a_n(x) + b_n(y)) = \lim_n \tilde{a}_n(x) + \lim_n (b_n(y) + a_n(\bar{x})).$$

Since $y \in A_{\bar{x}}$ we have $\bar{x}(y) = \bar{x}$, and so we can rewrite (3.34) as

$$(3.35) \quad \lim(a_n(x) + b_n(y)) = \lim_n \tilde{a}_n(x) + \lim_n \tilde{b}_n(y) = a(x) + b(y).$$

We have

$$(3.36) \qquad\qquad \nu(S_0 \times S_1) = \nu(S \times S_1) = \nu_1(S_1) = 1,$$

hence $\nu(A \cap (S_0 \times S_1)) = 1$, and so we can conclude from (3.20) and (3.35) that the decomposition (3.22) holds ν-a.s.

(iv) So far there is no reason why the functions a and b should be measurable. In fact, the counterexamples in Section 5 will show that without additional assumptions there may not exist any measurable version of the decomposition (3.21). The problem is that the definition of the functions

$$(3.37) \qquad\qquad \tilde{a}_n = a_n - a_n \circ T, \qquad \tilde{b}_n = (b_n + a_n \circ T)$$

uses the selection map $T: S \to S$ defined by $T(x) = \bar{x}$, and T may not be measurable. Since each equivalence class $E(x)$ is measurable, the problem will disappear as soon as there is only a countable number of equivalence classes since T is clearly measurable in this case.

Now note that assumption (3.22) implies $\nu(x, \cdot) \ll \nu_1$ for ν_0-a.a. $x \in S$, hence $\nu_1(A_x) > 0$ for ν_0-a.a. $x \in S_0$. This allows us to replace S_0 by the measurable set $\tilde{S}_0 = \{x \in S_0 | \nu_1(A_x) > 0\}$ in our construction above. However, since there is at most a countable number of sets A_x with $x \in \tilde{S}_0$, there is only a countable number of equivalence classes in $\tilde{S}_0$. This implies the measurability of $\tilde{a}_n(\cdot)$ and $\tilde{b}_n(\cdot)$, hence of $a(\cdot)$ and $b(\cdot)$.

(v) The uniqueness statement follows if we apply the preceding construction to the set $\tilde{A} = \{(x, y) \in S \times S | a(x) - \tilde{a}(x) = b(y) - \tilde{b}(y)\}$. We define A_i as a level set of the form $\{a(\cdot) - \tilde{a}(\cdot) = \alpha_i\}$, and B_i as the corresponding level set $\{b(\cdot) - \tilde{b}(\cdot) = \alpha_i\}$. As in (iv), condition (3.22) implies that it is enough to consider at most a countable number of such level sets. $\square$

REMARK 3.38. (a) If ν is equivalent to the product of its marginals, then the construction of a measurable decomposition (3.21) simplifies and reduces to the proof in Donsker and Varadhan (1975), page 403. In fact, $\nu \approx \nu_0 \otimes \nu_1$ implies $\nu_1 \ll \nu(x, \cdot)$, hence $\nu_1(A_x) = 1$ for ν-a.a. $x \in S_0$, and this allows us to replace S_0 by one single equivalence class. In particular, the uniqueness part simply states that the functions $a(\cdot)$ and $b(\cdot)$ are a.s. unique up to one additive constant. Note that this simplified construction only requires the assumption $\nu_0 \otimes \nu_1 \ll \nu$.

(b) The multiplicative analogue of Proposition 3.19 follows by taking logarithms: under the condition (3.22) an approximation

$$(3.39) \qquad \lim_n \left(f_n(x) g_n(y) \right) = \varphi(x, y) > 0, \quad \nu\text{-a.s.,}$$

with nonnegative measurable functions f_n and g_n implies a measurable factorization of the form

$$(3.40) \qquad \varphi(x, y) = f(x) g(y),$$

with nonnegative measurable functions f and g. If we have even $\nu \approx \nu_0 \otimes \nu_1$, then it follows from part (a) that both f and g are uniquely determined a.s. up to a multiplicative constant.

Let us now connect the results of this section with our discussion of the Markov property in Section 2 and with Schrödinger's original problem. Consider a probability measure $Q \approx P$ and denote by ν the joint distribution of (X_0, X_1) under Q. The relative entropy of Q with respect to P is given by

$$(3.41) \qquad H(Q|P) = H(\nu|\mu) + \int H(Q_x^y|P_x^y)\nu(dx, dy),$$

where Q_x^y resp. P_x^y denotes a conditional distribution of Q, resp. P, given $X_0 = x$ and $X_1 = y$. Since relative entropy is nonnegative, minimization of $H(\cdot|P)$ under a constraint on the marginals at times 0 and 1 proceeds in two steps. In the first place, Q must be of the form (2.4) so that the second term in (3.41) vanishes. Thus, the problem is reduced to the minimization of $H(\cdot|\mu)$ under the constraint that the marginals are given by ν_0 and ν_1. Consider a measure $\nu \in M(\nu_0, \nu_1)$ which satisfies condition (3.1). Denote by φ the density of ν with respect to μ and by $Q \approx P$ the associated process defined by (2.4). We are going to compare the following properties of the measure ν:

(i) ν *admits a measurable factorization* (3.16) *of the density*;
(ii) ν *coincides with the minimizing measure* ν^* *defined by* (3.3);
(iii) Q *is a Schrödinger process.*

We recall the regularity condition

$$(3.42) \qquad \mu \ll \mu[\cdot|X_t = z] \quad \text{for } \mu_t\text{-a.a. } z \in S$$

for the underlying process P which was used in Proposition 2.25.

THEOREM 3.43.   *Assume that $\nu \in M(\nu_0, \nu_1)$ satisfies condition* (3.1). *Then we have* (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii). *Under the additional condition* $\mu \ll \mu_0 \otimes \mu_1$, *we have* (ii) $\Rightarrow$ (i), *and under the stronger condition* (3.42), *all three properties are equivalent.*

PROOF.   We have already shown that (i) $\Rightarrow$ (ii) holds in general, and that (iii) $\Rightarrow$ (i) holds under condition (3.42); see Corollary 3.15 and Proposition 2.25. The implication (i) $\Rightarrow$ (iii) holds even without the finite entropy condition; see (2.16). Recall from (2.30) that (3.42) implies $\mu \ll \mu_0 \otimes \mu_1$, and note that this is equivalent to condition (3.22) since $\nu \approx \mu$ and $\nu_0 \otimes \nu_1 \approx \mu_0 \otimes \mu_1$. Combining the multiplicative version (3.40) of Proposition 3.19 with a.s. convergence along a subsequence in the approximation (3.9), we see that (ii) $\Rightarrow$ (i) under condition $\mu \ll \mu_0 \otimes \mu_1$. It only remains to show (ii) $\Rightarrow$ (iii).

Since S is a Polish space, the set $M(\nu_0, \nu_1)$ can be described as

$$(3.44) \qquad M(\nu_0, \nu_1) = \left\{ \nu \,\middle|\, \int \psi_k \, d\nu = c_k, \, k = 1, 2, \dots \right\},$$

where $h_k \in C_b(S)$, $\psi_k = h_k(X_0)$ if k is even and $\psi_k = h_k(X_1)$ if k is odd. Let

$$(3.45) \qquad A_n = \left\{ \nu \,\middle|\, \int \psi_k \, d\nu = c_k, \, 1 \le k \le n \right\}.$$

For each $n \ge 1$ there is a unique measure ν_n^* such that

$$(3.46) \qquad H(\nu_n^* | \mu) = \min_{\nu \in A_n} H(\nu | \mu),$$

and its density φ_n^* with respect to μ is of the form

$$(3.47) \qquad \varphi_n^* = Z_n^{-1} \exp\left(\sum_{i=1}^{n} \lambda_{n,k} \psi_k \right) = f_n(X_0) g_n(X_1),$$

with constants Z_n, $\lambda_{n,k}$, $1 \le k \le n$, and measurable nonnegative functions f_n, g_n on S; cf. Csiszár (1975). As shown in Föllmer (1988), page 163, we have

$$(3.48) \qquad \lim_n H(\nu^* | \nu_n^*) = 0.$$

Define the processes Q_n^* corresponding to ν_n^* by

$$(3.49) \qquad \frac{dQ_n^*}{dP} = \varphi_n^*(X_0, X_1).$$

Using the factorization (3.47) and Proposition 2.15, we see that each measure Q_n^* has the Markov property. Due to (3.49) we have

$$(3.50) \qquad \lim_n H(Q^* | Q_n^*) = \lim_n H(\nu^* | \nu_n^*) = 0,$$

and this implies convergence in total variation

$$(3.51) \qquad \lim_n \| Q^* - Q_n^* \| = 0.$$

However, it is easy to check that the Markov property is preserved under convergence in total variation, and this implies that Q^* is a Schrödinger process. $\square$

REMARK 3.52. (a) A measure $\nu \in M(\nu_0, \nu_1)$ is uniquely determined by property (ii), hence also by property (i). Under condition (3.42) we can also conclude that ν is uniquely determined by (iii). In other words, (3.42) implies the uniqueness of a Schrödinger process $Q \approx P$ with finite relative entropy $H(Q|P)$ and given marginals ν_0 and ν_1. Note that the uniqueness of Schrödinger processes with given marginals does not hold in general, as shown by Example 2.17.

(b) The implication (ii) $\Rightarrow$ (i) is stated as Theorem 10.5 in Nagasawa (1993) but not fully proved. Note that we do not claim that there is a measurable factorization (3.16) such that the integrability conditions $\log f \in L^1(\nu_0)$ and $\log g \in L^1(\nu_0)$ are satisfied. In fact, such a factorization may not exist. Consider, for example, one-dimensional Brownian motion with some initial distribution μ_0. In this case we have $\mu \approx \mu_0 \otimes \mu_1$. Thus, the factorization of the function $\varphi(x, y) = \exp(y - x - 1/2)$ into $f(x) = \exp(-x)$ and $g(y) = \exp(y - 1/2)$ is μ-a.s. unique up to a multiplicative constant; cf. Remark 3.38. Since $\log \varphi$ has distribution $N(-\frac{1}{2}, 1)$ under μ, the measure ν defined by the density φ satisfies

$$(3.53) \qquad H(\nu|\mu) = \int \varphi \log \varphi \, d\mu = \tfrac{1}{2} < \infty.$$

However, since $\nu_0 = \mu_0$, the function $\log f(x) = -x$ cannot belong to $L^1(\nu)$ as soon as we choose the distribution μ_0 such that it does not have a finite mean. Another example can be found in Rüschendorf and Thomsen (1993).

In the next section we will consider an infinite-dimensional Brownian motion P on $S = R^\infty$. In this case, neither the regularity assumption (3.42) nor the condition $\mu \ll \mu_0 \otimes \mu_1$ in Theorem 3.43 can be expected to hold. Nevertheless, we are going to show that a Schrödinger process $Q \approx P$ with bounded density φ does admit a measurable factorization. This will involve different methods which exploit the special structure of P.

4. A factorization of Schrödinger processes of the Brownian sheet.
Let P be the distribution of an infinite collection of independent Brownian motions with initial distributions μ_0^i, $i = 1, 2, \ldots$. If P^i denotes the Wiener measure on $C[0, 1]$ with initial distribution μ_0^i, then P is defined as the product measure

$$(4.1) \qquad P = \prod_{i=1}^{\infty} P^i$$

on the canonical path space

$$(4.2) \qquad \Omega = \prod_{i=1}^{\infty} C[0, 1] = C([0, 1], S),$$

with infinite-dimensional state space $S = R^\infty$. We denote by

$$X_t = \left(X_t^i \right)_{i=1, 2, \ldots}, \qquad 0 \leq t \leq 1,$$

the coordinate process on Ω, by $\mu_t = P \circ X_t^{-1}$, $0 \le t \le 1$, the marginal distributions on S and by $\mu = P \circ (X_0, X_1)^{-1}$ the joint distribution of (X_0, X_1) under P.

REMARK 4.3. As in the Lévy–Ciesielski construction of Brownian motion as a superposition of the Schauder functions $\{e_i,\ i \ge 1\}$ with independent $N(0, 1)$-coefficients, we could pass to the representation

$$\tilde{X}_t(\cdot, \tau) = \sum_{i=1}^{\infty} X_t^i e_i(\tau), \qquad 0 \le \tau \le 1,$$

with state space $\tilde{S} = C[0, 1]$ and to the corresponding distribution $\tilde{P}$ on

$$\tilde{\Omega} = C([0, 1], \tilde{S});$$

cf. for example, Gantert (1994). Thus, our infinite-dimensional Brownian motion could be viewed as a $C[0, 1]$-valued diffusion or, equivalently, as a *Brownian sheet* $(X_t(\cdot, \tau))$ indexed by the unit square. For our purpose, it will be convenient to use the explicit product structure (4.1).

For $x = (x^i) \in S$ and $y = (y^i) \in S$, the infinite-dimensional Brownian bridge from x to y is defined as the product measure

$$(4.4) \qquad\qquad P_x^y = \prod_{i=1}^{\infty} P_{x^i}^{y^i}$$

on Ω. For a probability measure ν on $S \times S$, the corresponding process Q is given by

$$(4.5) \qquad\qquad Q = \int P_x^y \nu(dx, dy).$$

From now on we assume that ν is equivalent to the joint distribution μ of (X_0, X_1) under P with density

$$(4.6) \qquad\qquad \varphi = \frac{d\nu}{d\mu} > 0 \quad \mu\text{-a.s.}$$

Then Q is equivalent to P with density

$$(4.7) \qquad\qquad \frac{dQ}{dP} = \varphi(X_0, X_1).$$

Due to Proposition 2.15, Q is a Schrödinger process if the density φ has the product form (2.10). In our infinite-dimensional context, the question is to which extent the converse holds, that is, under which conditions the Markov property of Q implies a measurable factorization of the density. The following theorem provides an answer to this question. The second counterexample in Section 5, where Q is a Schrödinger process whose density is in L^p for any $p < \infty$ but does not admit a factorization, will show that some strong restriction on the density is really needed.

THEOREM **4.8.** *Suppose that $Q \approx P$ is a Schrödinger process with density $\varphi \in L^{\infty}(\mu)$. Then φ admits a factorization of the form*

$$(4.9) \qquad \varphi(x, y) = f(x) g(y) \quad \mu\text{-a.s.},$$

where f and g are positive measurable functions on the state space $S = R^{\infty}$. In particular, Q has the structure of an h-process for some space–time harmonic function of infinite-dimensional Brownian motion.

PROOF.   Let us fix a time $t \in (0, 1)$. Our starting point is the weak factorization

$$(4.10) \qquad \varphi(X_0, X_1) = \varphi_{0, t}(X_0, X_t) \psi_{t, 1}(X_t, X_1), \quad P\text{-a.s.},$$

which follows from the Markov property of Q; cf. (2.23). Using the regularity of the finite-dimensional projections of P, we are going to eliminate the dependence on the first m coordinates of X_t in (4.10). Then we will pass to the limit, $m \to \infty$. Using the product structure of P, we will combine martingale arguments with Kolmogorov's zero–one law in order to eliminate completely the dependence on the intermediate value X_t. This will lead to the factorization (4.9).

(i) For $m \geq 1$ we define the σ-fields

$$(4.11) \qquad A_m = \sigma\left(\left\{X_0, X_1, X_t^i, i > m\right\}\right)$$

and

$$(4.12) \qquad B_m = \sigma\left(\left\{X_t^i, i \leq m\right\}\right).$$

The conditional distribution of P with respect to B_m is of the form

$$(4.13) \qquad P\left[\cdot \,|X_t^i = z^i, i = 1, \ldots, m\right] = \prod_{i \leq m} P^i\left[\cdot \,|X_t^i = z^i\right] \otimes \prod_{i > m} P^i.$$

Since

$$(4.14) \qquad P^i\left[\cdot \,|X_t^i = z^i\right] \approx P^i \quad \text{on } \sigma\left(X_0^i, X_1^i\right),$$

we see that

$$(4.15) \qquad P^i\left[\cdot \,|X_t^i = z^i, i = 1, \ldots, m\right] \approx P \quad \text{on } A_m$$

for any choice of z^i, $i = 1, \ldots, m$. In fact, the equivalence (4.15) can be shown to hold even without the assumption that the initial distribution μ_0 is a product measure.

(ii) Let us use the notation

$$X_{(m)} = (x^1, \ldots, x^m), \qquad x^{(m)} = (x^{m+1}, x^{m+2}, \ldots)$$

for a given sequence $x \in R^{\infty}$. For a fixed $z \in R^{\infty}$, the equivalence (4.15) together with the weak factorization (4.10) implies the factorization

$$(4.16) \quad \varphi(X_0, X_1) = \varphi_{0, t}\left(X_0, \left(z_{(m)}, X_t^{(m)}\right)\right) \psi_{t, 1}\left(\left(z_{(m)}, X_t^{(m)}\right), X_1\right), \quad P\text{-a.s.}$$

We are now going to show that z can be chosen such that we can pass from (4.16) to the asymptotic factorization

(4.17)
$$\varphi(X_0, X_1) = \limsup_m \varphi_{0,t}\left(X_0, \left(z_{(m)}, X_t^{(m)}\right)\right)$$
$$\times \liminf_m \psi_{t,1}\left(\left(z_{(m)}, X_t^{(m)}\right), X_1\right), \quad P\text{-a.s.}$$

Note first that for any $z \in R^\infty$ the lim sup in (4.17) is finite since $\varphi_{0,t} \le \|\varphi\|_\infty$. However, for two sequences (a_n) and (b_n) of positive real numbers with constant product $a_n b_n = c > 0$ and $\sup a_n < \infty$, we can conclude that

(4.18)
$$c = \left(\limsup_n a_n\right)\left(\liminf_n b_n\right)$$

as soon as we know that $\liminf_n b_n < \infty$. In fact, for a subsequence (n_l) with $\liminf_n b_n = \lim_l b_{n_l} < \infty$, we have

(4.19)
$$c = \lim_l \left(a_{n_l} b_{n_l}\right) \le \left(\limsup_n a_n\right)\left(\liminf_n b_n\right),$$

and, for a subsequence (n_k) with $\lim_k a_{n_k} = \limsup_n a_n$, we get the reversed inequality. Thus, we have to show that there is a choice of $z \in R^\infty$ such that the lim inf in (4.17) is finite P-a.s. Since

(4.20)
$$\psi_{t,1} = \varphi_{t,1}(\varphi_t)^{-1},$$

we have

(4.21)
$$\psi_{t,1}\left(\left(z_{(m)}, X_t^{(m)}\right), X_1\right) \le \|\varphi\|_\infty \left(\varphi_t\left(z_{(m)}, X_t^{(m)}\right)\right)^{-1},$$

and so the problem is reduced to finding $z \in R^\infty$ such that

(4.22)
$$\limsup_m \varphi_t\left(z_{(m)}, X_t^{(m)}\right) > 0, \quad P\text{-a.s.}$$

(iii) Due to the product structure of P, we can apply Kolmogorov's zero–one and conclude that, for any choice of $z \in R^\infty$,

(4.23)
$$\limsup_m \varphi_t\left(z_{(m)}, X_t^{(m)}\right) = c(z), \quad P\text{-a.s.}$$

with some constant $c(z) \in [0, \|\varphi\|_\infty]$. By Fatou's lemma,

(4.24) $$c(z) = E\left[\limsup_m \varphi_t\left(z_{(m)}, X_t^{(m)}\right)\right] \ge \limsup_m E\left[\varphi_t\left(z_{(m)}, X_t^{(m)}\right)\right].$$

However, since P (resp., $\mu_t) = P \circ X_t^{-1}$ is a product measure, the function M_m on R^∞ defined by

(4.25) $$M_m(z) = E\left[\varphi_t\left(z_{(m)}, X_t^{(m)}\right)\right] = \int \varphi_t\left(z_{(m)}, x^{(m)}\right)\mu_t(dx)$$

is a version of the conditional expectation of φ_t with respect to μ_t, given the first m coordinates of $z \in R^\infty$. By martingale convergence,

(4.26)
$$\lim_m M_m(z) = \varphi_t(z), \quad \mu_t\text{-a.s.}$$

Since $\varphi_t(z) > 0$ μ_t-a.s., we see that (4.22) does hold for μ_t-a.a. $z \in R^\infty$. From now on we fix a sequence z such that (4.22) is satisfied. As we have seen in part (ii), this leads to the asymptotic factorization (4.17).

(iv) For $x = (x^i) \in R^\infty$ let P_x denote the product of Wiener measures P_{x^i} with initial point x^i, and let P^x denote the product of the measures P^{x^i}, where P^{x^i} denotes the conditional distribution of a Wiener measure with initial distribution μ_0^i, given that it assumes the value x^i at the terminal time $t = 1$. Using Kolmogorov's zero–one law for the product measure P_x, we obtain

$$(4.27) \qquad \limsup_m \varphi_{0,t}\left(x, \left(z_{(m)}, X_t^{(m)}\right)\right) = f(x), \quad P_x\text{-a.s.,}$$

where f denotes the measurable function on R^∞ defined by

$$(4.28) \qquad f(x) = \int \limsup_m \varphi_{0,t}\left(x, \left(z_{(m)}, X_t^{(m)}\right)\right) dP_x.$$

This implies

$$(4.29) \qquad \limsup_m \varphi_{0,t}\left(X_0, \left(z_{(m)}, X_t^{(m)}\right)\right) = f(X_0), \quad P\text{-a.s.}$$

In the same way, we can apply Kolmogorov's zero–one law for the measures P^x to the second factor in (4.17). This leads to

$$(4.30) \qquad \liminf_m \psi_{t,1}\left(\left(z_{(m)}, X_t^{(m)}\right), X_1\right) = g(X_1), \quad P\text{-a.s.,}$$

where g is the function defined by

$$(4.31) \qquad g(x) = \int \liminf_m \psi_{t,1}\left(\left(z_{(m)}, X_t^{(m)}\right), x\right) dP^x.$$

The factorization (4.17) together with (4.29) and (4.30) yields the desired factorization (4.9). As in Proposition 2.15 we can now construct a space–time harmonic function h such that Q is represented as an h-process of our infinite-dimensional Brownian motion. □

REMARK 4.32. (a) Instead of assuming that φ is bounded from above, we could also assume that φ is bounded away from 0. In that case φ^{-1} is bounded from above, and the process $\tilde{Q} \approx P$ defined by $\tilde{\varphi} = \varphi^{-1}(\int \varphi^{-1} d\mu)^{-1}$ has again the Markov property, since the weak factorization (2.23) for φ induces a weak factorization for $\tilde{\varphi}$. Thus, $\tilde{Q}$ is a Schrödinger process, and the factorization (4.9) for $\tilde{\varphi}$ implies a measurable factorization for φ.

(b) If φ is bounded both from above and away from 0, then the proof simplifies since (4.22) is clearly satisfied. In this case we could also drop the assumption that the initial distribution μ is a product measure. However, this requires a modification of part (iv) involving a passage to the limit $t \uparrow 1$ and the use of Blumenthal's zero–one law for the time-reversed processes P^x, since Kolmogorov's zero–one law for the spatial tail field may no longer hold for P^x.

Let us now relate Theorem 4.8 for infinite-dimensional Brownian motion P to our results in Section 3 on the minimization of relative entropy under fixed marginals. Note first that the regularity condition (3.42) clearly does not hold in our present context, not even if the initial distribution μ_0 is concentrated on one point. The weaker condition $\mu \ll \mu_0 \otimes \mu_1$ appearing in Theorem 3.43 only holds under very strong restrictions on the initial distribution μ_0. For example, it is satisfied if μ_0 is concentrated on the Cameron–Martin space $H \subset C[0, 1]$ (resp., on the sequence space $l^2 \subset R^\infty$). Nevertheless, we were able to show that a Schrödinger process $Q \approx P$ with bounded density φ admits a measurable factorization. Combining this result with the first part of Theorem 3.43, we see that, under the additional condition $\varphi \in L^\infty$, the three properties (i), (ii) and (iii) appearing in Theorem 3.43 are all equivalent in our present context:

THEOREM 4.33. *For infinite-dimensional Brownian motion P, consider a process $Q \approx P$ given by a probability measure $\nu \approx \mu$ on $S \times S$ as in* (4.5). *If the density φ of ν with respect to μ is bounded from above, then the Markov property* (iii) *of Q is equivalent both to a measurable factorization* (i) *of the density and to the property* (ii) *that ν minimizes relative entropy under the constraint that its marginals are fixed.*

In the next section we are going to construct a counterexample where $Q \approx P$ is a Schrödinger process of infinite-dimensional Brownian motion whose density belongs to L^p for any $p < \infty$, but where the implication (iii) $\Rightarrow$ (i) and even the implication (ii) $\Rightarrow$ (i) break down.

5. Two counterexamples. In this section we construct two examples of measures on a product space which minimize relative entropy under the constraint that the marginals are fixed, but which do not admit a measurable factorization of the density.

Our first example shows that even boundedness of the density, both from above and away from 0, does not guarantee that such a factorization holds. We take the space $S = \{0, 1\}^{\{1, \dots\}}$ of binary sequences and define a probability measure μ on $S \times S$ as follows. The marginal distribution μ_0 is defined as the Bernoulli measure with parameter $p = \frac{1}{2}$. For $x \in S$ and $k \geq 1$, we define $R_k(x) \in S$ as the sequence where the kth component of x has been removed. That is, $(R_k(x))^i = x^i$ for $i < k$ and $(R_k(x))^i = x^{i+1}$ for $i \geq k$. Let γ denote a strictly positive probability distribution on $\{1, 2, \dots\}$. For a given sequence $x = (x^1, x^2, \dots) \in S$, we define the transition probability $\mu(x, dy)$ as

(5.1) $$\mu(x, \cdot) = \sum_{k=1}^{\infty} \gamma_k \delta_{R_k(x)}.$$

Thus, the sequence $y \in S$ is obtained from $x \in S$ by removing, with probability γ_k, the kth component of x. Note that the marginal distribution μ_1 coincides with μ_0, and that $\mu(x, \cdot)$ is singular with respect to μ_1. In particular, μ does not satisfy the regularity condition $\mu \ll \mu_0 \otimes \mu_1$ appearing in Theorem 3.43.

We define

$$(5.2) \qquad f_n(x) = \exp\left(\sum_{k=1}^{n} x^k\right), \qquad g_n(y) = C^{-1}\exp\left(-\sum_{k=1}^{n-1} y^k\right),$$

where $C = \frac{1}{2}(e + 1)$. The functions φ_n on $S \times S$ defined by

$$(5.3) \qquad \varphi_n(x, y) = f_n(x)g_n(y)$$

satisfy

$$(5.4) \qquad \varphi_n(x, R_k(x)) = C^{-1}\exp(x^{\min(n, k)}) \in \{C^{-1}, eC^{-1}\}$$

and

$$(5.5) \qquad \int \varphi_n \, d\mu = \sum_{k=1}^{\infty} \gamma_k \int \varphi_n(x, R_k(x))\mu_0(dx) = 1.$$

Since $\mu(x, \cdot)$ is concentrated on the countable set $\{R_k(x)|k \geq 1\}$, the limit $\varphi := \lim_n \varphi_n$ exists μ-a.s. and satisfies

$$(5.6) \qquad \varphi(x, R_k(x)) = C^{-1}\exp(x^k) \in \{C^{-1}, eC^{-1}\}.$$

We have $\int \varphi \, d\mu = 1$ as in (5.5), and so φ is the density of a probability measure $\nu \approx \mu$. We denote by ν_0 and ν_1 the marginal distributions of ν.

PROPOSITION 5.7. *The measure ν has minimal relative entropy $H(\nu|\mu)$ under the constraint that the marginals are given by ν_0 and ν_1. Its density φ is bounded from above and away from 0, but it does not admit a measurable factorization.*

PROOF. (i) The boundedness of φ is clear since (5.4) shows that φ assumes only two positive values. Due to (5.4), we can apply Lebesgue's theorem to verify our criterion in Proposition 3.6. Thus, ν coincides with the measure ν^* which minimizes $H(\nu|\mu)$ under the constraint that the marginals are given by ν_0 and ν_1.

(ii) Note first that the product form (5.3) of φ_n does not directly imply a factorization of φ since $\lim_n f_n(x) = \infty$ μ_0-a.s. and $\lim_n g_n(y) = 0$ μ_1-a.s. Let us now assume that there are nonnegative measurable functions f and g on S such that

$$(5.8) \qquad f(x)g(y) = \varphi(x, y), \quad \mu\text{-a.s.,}$$

and let us show that this leads to a contradiction. Taking logarithms, we get

$$(5.9) \qquad \log f(x) + \log g(y) = \log \varphi(x, y), \quad \mu\text{-a.s.,}$$

hence

$$(5.10) \qquad \log f(x) + \log g(R_k(x)) = x^k - \log C, \quad \mu_0\text{-a.s.}$$

for each $k \geq 1$, due to (5.1), (5.6) and the assumption that γ is strictly positive. This implies, for each $k \geq 1$, the relation

$$(5.11) \quad \log f(0, \ldots, 0, x^k, x^{k+1}, \ldots) - \log f(0, \ldots, 0, 0, x^{k+1}, \ldots) = x^k$$
$$\mu_0\text{-a.s.,}$$

hence

(5.12) $\log f(x) = \log f(0, \dots, 0, x^{n+1}, x^{n+2}, \dots) + \sum_{k=1}^{n} x^k, \quad \mu_0\text{-a.s.}$

for each $n \geq 1$. Equation (5.12) is of the form

(5.13) $Z = Z_n + S_n, \quad P\text{-a.s.},$

where Z is finite P-a.s., Z_n and S_n are independent and S_n has binomial distribution $B(n, \frac{1}{2})$ under P. However, this leads to a contradiction. Choose a such that $P[|Z| \geq a] \leq \varepsilon$. Then

(5.14) $\varepsilon \geq P[Z \geq a] \geq P\left[Z_n \geq -\dfrac{n}{2} + a\right] P\left[S_n \geq \dfrac{n}{2}\right]$

by independence, and this implies

(5.15) $\limsup P\left[Z_n \geq -\dfrac{n}{2} + a\right] \leq 2\varepsilon.$

In the same way,

(5.16) $\varepsilon \geq P[Z \leq -a] \geq P\left[Z_n < -\dfrac{n}{2} + a\right] P\left[S_n \leq \dfrac{n}{2} - 2a\right]$

implies

(5.17) $\limsup P\left[Z_n < -\dfrac{n}{2} + a\right] \leq 2\varepsilon,$

whereas the sum of the first terms in (5.15) and (5.17) should clearly be greater than or equal to 1. $\square$

REMARK 5.18.   Let us illustrate how the construction of a measurable decomposition in the proof of Proposition 3.19 breaks down in our example. Due to (5.2) and (5.3), we have

(5.19) $c(x, y) = \log \varphi(x, y) = \lim_n (a_n(x) + b_n(y)), \quad \mu\text{-a.s.},$

where the function a_n on $S = R^\infty$ is defined by

(5.20) $a_n(x) = \sum_{i=1}^{n} x^i.$

In our case, the equivalence relation (3.27) on $S = S_0$ is given by

(5.21) $x \sim \tilde{x} \quad \text{iff } \exists\, m \geq 1 \text{ such that } x^i = \tilde{x}^i \text{ for all } i \geq m.$

The equivalence classes $E(x)$ are measurable with respect to the tail field

(5.22) $\mathcal{A}^* = \bigcap_m \sigma(\{x^i | i \geq m\});$

in fact, they are the atoms of $\mathcal{A}^*$. If there were a measurable choice of representatives $\tilde{x} = T(x)$, the map $T: S \to S$ would be measurable with respect to $\mathcal{A}^*$, hence constant μ_0-a.s. due to Kolmogorov's zero–one law. This

would imply that μ_0 is concentrated on a single atom, contradicting the fact that $\mu_0(E(x)) = 0$ for any atom $E(x)$.

Let us now return to the infinite-dimensional Brownian motion

$$(5.23) \qquad P = \prod_{i=1}^{\infty} P^i,$$

with state space $S = R^\infty$. For a process $Q \approx P$ of the form (4.5) with bounded density φ, we have seen in Theorem 4.33 that both the Markov property and the property of minimizing the relative entropy under given marginals imply a factorization of the density. Our next example will show that both implications may fail if boundedness is replaced by being in L^p for all $p < \infty$.

Let us specify the initial distributions

$$(5.24) \qquad \mu_0^i = \tfrac{1}{2}\big(\delta_0 + \delta_{1/\alpha_i}\big),$$

where $\alpha_1, \alpha_2, \ldots \in R^+$ are such that

$$(5.25) \qquad \sum_{i=1}^{\infty} \alpha_i^2 = 2.$$

The joint distribution of (X_0, X_1) under P is given by the measure

$$(5.26) \qquad \mu = \prod_{i=1}^{\infty} \tfrac{1}{2}\big(\delta_0 \otimes N(0,1) + \delta_{1/\alpha_i} \otimes N(1/\alpha_i, 1)\big)$$

on $S \times S$. The measure

$$(5.27) \qquad \nu = \prod_{i=1}^{\infty} \tfrac{1}{2}\big(\delta_0 \otimes N(\alpha_i, 1) + \delta_{1/\alpha_i} \otimes N(\alpha_i + 1/\alpha_i, 1)\big)$$

is equivalent to μ, and the density φ of ν with respect to μ is defined μ-a.s. by

$$(5.28) \qquad \begin{aligned} \varphi(x,y) &= \exp(-1) \prod_{i=1}^{\infty} \exp\big[\alpha_i(y^i - x^i)\big] \\ &= \exp(-1)\exp\left[\sum_{i=1}^{\infty} \alpha_i(y^i - x^i)\right]. \end{aligned}$$

Note that the sum in the exponent converges μ-a.s.: since the differences $y^i - x^i$, $i = 1, 2, \ldots$, are independent and identically distributed with $N(0,1)$ under μ, the partial sums form a martingale which is bounded in $L^2(\mu)$ due to (5.25), hence μ-a.s. convergent to a finite limit. We denote by ν_0 and ν_1 the marginals of ν.

Let $Q \approx P$ denote the process with density

$$(5.29) \qquad \frac{dQ}{dP} = \varphi(X_0, X_1).$$

Note that Q is of the form

$$(5.30) \qquad\qquad Q = \prod_{i=1}^{\infty} Q^i,$$

where Q^i is the distribution of a one-dimensional Brownian motion with constant drift α_i and initial distribution μ_0^i. We are now going to show that Q is a Schrödinger process with minimal relative entropy, but that Q does *not* have the structure of an *h*-transform of infinite-dimensional Brownian motion. In other words, it is not possible to construct a measurable factorization of the density.

PROPOSITION 5.31. *The process Q has the Markov property, it minimizes relative entropy $H(Q|P) = H(\nu|\mu)$ under the constraint that the marginals ν_0 and ν_1 are fixed, and the density φ belongs to L^p for any $p < \infty$. However, the density does not admit a measurable factorization of the form*

$$(5.32) \qquad\qquad \varphi(x, y) = f(x) g(y), \quad \mu\text{-}a.s.$$

PROOF.   (i) Let us write

$$(5.33) \qquad\qquad \varphi(x, y) = \lim_{n} \left[f_n(x) g_n(y) \right], \quad \mu\text{-a.s.},$$

where

$$(5.34) \qquad\qquad f_n(x) = c_n^{-1} \prod_{i=1}^{n} \exp(-\alpha_i x^i),$$

$$g_n(y) = \prod_{i=1}^{n} \exp(\alpha_i y^i), \qquad c_n = \exp\left(\tfrac{1}{2} \sum_{i=1}^{n} \alpha_i^2 \right).$$

The differences $y^i - x^i$, $i = 1, 2, \ldots$, are independent with distribution $N(\alpha_i, 1)$ under ν, and this implies that the partial sums

$$(5.35) \qquad \log\left[f_n(x) g_n(y) \right] = \sum_{i=1}^{n} \alpha_i(y^i - x^i) - \tfrac{1}{2} \sum_{i=1}^{n} \alpha_i^2$$

converge to $\log \varphi$ in $L^2(\nu)$. Due to Proposition 3.6, we can conclude that ν minimizes the relative entropy $H(\cdot|\mu)$ under the constraint that the marginals are fixed to be ν_0 and ν_1. Note also that (5.28) implies that

$$(5.36) \qquad\qquad \int \varphi^p \, d\mu = \exp[p(p-1)],$$

that is, the density φ belongs to $L^p(\mu)$ for any $p < \infty$.

(ii) The representation (5.30) shows that the coordinate process (X_t) has the Markov property under Q, and so Q is, in fact, a *Schrödinger process*. The Markov property also follows from the fact that ν minimizes the relative entropy $H(\cdot|\mu)$ under the given marginals; cf. Theorem 3.43. It is also implied by the weak factorization

$$(5.37) \qquad \varphi(X_0, X_1) = \varphi_{0,t}(X_0, X_t)\psi_{t,1}(X_t, X_1), \quad P\text{-a.s.}$$

for any $t \in (0, 1)$, where

$$\varphi_{0,t}(X_0, X_t) = \exp(-1) \prod_{i=1}^{\infty} \exp\left[\alpha_i(X_t^i - X_0^i)\right],$$

(5.38)

$$\psi_{t,1}(X_t, X_1) = \prod_{i=1}^{\infty} \exp\left[\alpha_i(X_1^i - X_t^i)\right];$$

cf. Lemma 2.22.

(iii) Let us first note that it is not possible to derive a factorization (5.32) directly from (5.33) since $\lim f_n(x) = 0$ and $\lim g_n(y) = \infty$ μ-a.s. In order to show that there is no alternative way of constructing a factorization, we proceed indirectly. Assume that there are nonnegative measurable functions f and g such that the factorization (5.32) holds μ-a.s. We have $f, g > 0$ μ-a.s. Taking logarithms, we get

$$\log f(x) + \log g(y) = \log \varphi(x, y)$$

(5.39)

$$= \sum_{i=1}^{\infty} \alpha_i(y^i - x^i) - 1, \quad \mu\text{-a.s.}$$

Let $\mu_{(n)}$ denote the conditional distribution of μ given that $x^i = 0$ for $i = 1, \ldots, n$. In our case, $\mu_{(n)}$ is absolutely continuous with respect to μ, and so (5.39) implies

$$\log f((0, \ldots, 0, x^{n+1}, \ldots)) + \log g(y)$$

(5.40)

$$= \log \varphi((0, \ldots, 0, x^{n+1}, \ldots), y)$$

$$= \sum_{i=1}^{n} \alpha_i y^i + \sum_{i=n+1}^{\infty} \alpha_i(y^i - x^i) - 1$$

$\mu_{(n)}$-a.s. However, $\mu_{(n)}$ and μ are equivalent on the σ-field generated by y and by the coordinates x^i, $i > n$, and so (5.40) holds also μ-a.s. Looking at the difference between (5.39) and (5.40), we see that for each $n \geq 1$ the relation

(5.41) $$\log f(x) = \log f(0, \ldots, 0, x^{n+1}, x^{n+2}, \ldots) - \sum_{i=1}^{n} \alpha_i x^i$$

holds μ-a.s. Equation (5.41) is of the form

(5.42) $$Z = Z_n + S_n, \quad P\text{-a.s.}$$

where Z is finite P-a.s., Z_n and S_n are independent and S_n has binomial distribution $B(n, \frac{1}{2})$ under P. As in the previous example, this leads to a contradiction. $\square$

REFERENCES

AEBI, R. (1996). *Schrödinger Diffusion Processes*. Birkhäuser, Basel.

BERNSTEIN, S. (1932). Sur les liaisons entre les grandeurs aléatoires. *Verh. Internat. Math. Kongr. Zürich* **1** 288–309.

BEURLING, A. (1960). An automorphism of product measures. *Ann. Math.* **72** 189–200.

BORWEIN, J. M. and LEWIS, A. S. (1992). Decomposition of multivariate functions. *Canad. J. Math.* **44** 463–482.

BORWEIN, J. M., LEWIS, A. S. and NUSSBAUM, R. (1994). Entropy minimization, DAD problems, and doubly stochastic kernels. *J. Funct. Anal.* **123** 264–307.

BROCKHAUS, O. (1995). Konditionierungen des Brownschen Blattes: Große Abweichungen und Schrödinger-Brücken. *Bonner Math. Schriften* **285**.

CSISZÁR, I. (1975). *I*-divergence geometry of probability distribution and minimization problems. *Ann. Probab.* **3** 146–158.

DAWSON, D., GOROSTIZA, L. and WAKOLBINGER, A. (1990). Schrödinger processes and large deviations. *J. Math. Phys.* **31** 2385–2388.

DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.* **29** 389–461.

DOOB, J. L. (1984). *Classical Potential Theory and Its Probabilistic Counterpart*. Springer, New York.

FÖLLMER, H. (1988). Random fields and diffusion processes. *Ecole d'Eté de Saint Flour XV–XVII. Lecture Notes in Math.* **1362** 101–203. Springer, New York.

FÖLLMER, H. (1991). Martin boundaries on Wiener Space. In *Diffusions Processes and Related Problems in Analysis. I* (M. Pinsky, ed.). 3–16. Birkhäuser, Basel.

FORTET, R. (1940). Résolution d'un systeme d'équations de M. Schrödinger. *J. Math. Pures Appl.* (9) 83–105.

GANTERT, N. (1994). Self-similarity of Brownian motion and a large deviation principle for random fields on a binary tree. *Probab. Theory Related Fields* **98** 7–20.

JAMISON, B. (1974). Reciprocal processes. *Z. Wahrsch. Verw. Gebiete* **30** 65–86.

JAMISON, B. (1975). The Markov process of Schrödinger. *Z. Wahrsch. Verw. Gebiete* **32** 323–331.

NAGASAWA, M. (1993). *Schrödinger Equations and Diffusion Theory*. Birkhäuser, Basel.

RÜSCHENDORF, L. and THOMSEN, W. (1993). Note on the Schrödinger equation and *I*-projections. *Statist. Probab. Lett.* **17** 369–375.

SCHRÖDINGER, E. (1931). Über die Umkehrung der Naturgesetze. *Sitzungsber. Preuss. Akad. Wiss. Berlin Phys. Math. Kl.* **8 / 9** 144–153.

SCHRÖDINGER, E. (1932). Sur la théorie relativiste de l'électron et l'interpretation de la mécanique quantique. *Ann. Inst. H. Poincaré* **2** 269–310.

THIEULLEN, M. (1993). Second order stochastic differential equations and non gaussian reciprocal diffusions. *Probab. Theory Related Fields* **97** 231–257.

THIEULLEN, M. and ZAMBRINI, J. C. (1995). Symmetries in stochastic calculus of variations. Preprint, Laboratoire de Probabilités, Univ. Paris VI.

HUMBOLDT-UNIVERSITÄT   TECHNISCHE UNIVERSITÄT
INSTITUT FÜR MATHEMATIK   FACHBEREICH MATHEMATIK
UNTER DEN LINDEN 6   STRASSE DES 17. JUNI 136
D-10099 BERLIN   D-10623 BERLIN
GERMANY   GERMANY
E-MAIL: foellmer@mathematik.hu-berlin.de