MATH. SCAND. 58 (1986), 239–254

AN INTEGRAL REPRESENTATION FOR THE HELLINGER DISTANCE

ESKO VALKEILA AND LJUDMILLA VOSTRIKOVA

Summary.

We give an integral representation for the Hellinger distance between two probability measures on a stochastic basis. This representation uses the Hellinger process, which depends on the measures involved and of the filtration of the basis. We give bounds for the Hellinger distance in terms of the Hellinger process. These results can be applied to the strong convergence of stochastic processes.

1. Introduction.

Suppose that there is a stochastic basis $(\Omega, \mathscr{F}, \boldsymbol{F})$, i.e. a measurable space $(\Omega, \mathscr{F})$ with a filtration $\boldsymbol{F} = (\mathscr{F}_t)_{t \geqq 0}$, an increasing right-continuous family of sub-σ-fields of $\mathscr{F}$, and $\mathscr{F} = \bigvee_{t \geqq 0} \mathscr{F}_t$.

Given two probability measures P and $\tilde{P}$ on $(\Omega, \mathscr{F})$ we denote by P_t and $\tilde{P}_t$ their restrictions on $\mathscr{F}_t$, $P_t = P|\mathscr{F}_t$ and $\tilde{P}_t = \tilde{P}|\mathscr{F}_t$. Define a probability measure Q by $Q = (P + \tilde{P})/2$ and denote by Q_t the restriction of Q on $\mathscr{F}_t$. We suppose that $\mathscr{F}_0$ contains all the sets of $\mathscr{F}$ with Q-measure zero.

Let $\boldsymbol{D}$ be the space of the right-continuous functions with left-hand limits. Denote by $\zeta = (\zeta_t, \mathscr{F}_t)_{t \geqq 0}$ and by $\tilde{\zeta} = (\tilde{\zeta}_t, \mathscr{F}_t)_{t \geqq 0}$ the Randon–Nikodym derivatives of the measures P_t and $\tilde{P}_t$ with respect to Q_t. We can take versions of ζ and $\tilde{\zeta}$, which have paths in $\boldsymbol{D}$, and for an arbitrary stopping time T

$$\zeta_T = \frac{dP_T}{dQ_T} \quad \text{and} \quad \tilde{\zeta}_T = \frac{d\tilde{P}_T}{dQ_T}, \tag{1}$$

for this we refer to [4, § 3, Lemma 2, p. 649].

DEFINITION 1. Let T be a stopping time. The Hellinger distance $\rho(P_T, \tilde{P}_T)$ between the measures P_T and $\tilde{P}_T$ is defined by the formula

$$\rho^2(P_T, \tilde{P}_T) = E_Q(\sqrt{\zeta_T} - \sqrt{\tilde{\zeta}_T})^2$$

Received April 18, 1984; in revised form February 6, 1985.

where E_Q is expectation with respect to Q.

In particular when $T = \infty$ we have

$$\rho^2(P, \tilde{P}) = E_Q(\sqrt{\zeta_\infty} - \sqrt{\tilde{\zeta}_\infty})^2$$

where $\zeta_\infty = \lim_{t\to\infty} \zeta_t$, $\tilde{\zeta}_\infty = \lim_{t\to\infty} \tilde{\zeta}_t$.

To formulate the main results we introduce the notion of Hellinger process associated with P, $\tilde{P}$, and $\boldsymbol{F}$.

Let Y be a process with paths in $\boldsymbol{D}$. Then we denote by Y^c and Y^d the continuous and discontinuous part of the process Y, i.e.

$$Y_t = Y_t^c + Y_t^d, \quad Y_t^d = \sum_{0<s\leqq t} \Delta Y_s \quad \text{and} \quad \Delta Y_s = Y_s - Y_{s-}.$$

We denote by μ_Y the jump measure of Y and by $\nu_{Y,Q}$ the compensator (dual predictable projection) of μ_Y with respect to $(\boldsymbol{F}, Q)$.

Let M be a square integrable local martingale. Denote by $\langle M^c\rangle$ the quadratic characteristic of M^c, i.e. the increasing predictable process such that $(M^c)^2 - \langle M^c\rangle$ is local martingale (for details see [1] or [12]).

For $t > 0$ and $x \in \mathsf{R}$ let

$$(2) \qquad \lambda_t(x) = \left(1 + x(\zeta_{t-})^{\oplus}\right) \vee 0, \quad \tilde{\lambda}_t(x) = \left(1 - x(\tilde{\zeta}_{t-})^{\oplus}\right) \vee 0,$$

$$(3) \qquad \beta_t = \left((\zeta_{t-})^{\oplus} + (\tilde{\zeta}_{t-})^{\oplus}\right)^2$$

where for $-\infty \leqq a \leqq \infty$

$$a^{\oplus} = \begin{cases} 0 & \text{if } a = 0, \\ a^{-1} & \text{if } a = 0,\ |a| \neq \infty, \\ 0 & \text{if } |a| = \infty. \end{cases}$$

DEFINITION 2. (Compare with [9, § 1, Definition 2, p. 388] and also [2]). Predictable process $H = (H_t, \mathscr{F}_t)_{t\geqq 0}$, where

$$(4) \quad H_t = \tfrac{1}{4}\int_{]0,t]} \beta_s d\langle\zeta^c\rangle_s + \int_{]0,t]}\int_{]0,\infty[} (\sqrt{\tilde{\lambda}_s(x)} - \sqrt{\lambda_s(x)})^2\, \nu_{\zeta,Q}(ds, dx)$$

is the Hellinger process.

Hellinger process characterizes the strong convergence between two sequences of probability measures. For this we refer to [9].

We give the connection between Hellinger distance and Hellinger process in the first theorem. When the Hellinger process is deterministic then we can give an explicit expression for the Hellinger distance in terms of the Hellinger process. In Theorems 2 and 3 we give upper and lower bounds for the Hellinger distance in terms of the Hellinger process.

THEOREM 1. *Let T be a stopping time and let $\rho(P_T, \tilde{P}_T)$ be the Hellinger distance between two probability measures P_T and $\tilde{P}_T$. Then we have the following integral representation*

$$\rho^2(P_T, \tilde{P}_T) = \rho^2(P_0, \tilde{P}_0) + E_Q \int_0^T \sqrt{\zeta_{s-} \tilde{\zeta}_{s-}}\, dH_s. \tag{5}$$

In particular when $T = \infty$ we get

$$\rho^2(P, \tilde{P}) = \rho^2(P_0, \tilde{P}_0) + E_Q \int_0^\infty \sqrt{\zeta_{s-} \tilde{\zeta}_{s-}}\, dH_s.$$

Denote by E (respectively $\tilde{E}$) the expectation with respect to P (respectively $\tilde{P}$).

COROLLARY 1. *The integral representation* (5) *can be written in terms of P or $\tilde{P}$:*

$$\rho^2(P_T, \tilde{P}_T) = 2E(1 - \sqrt{z_0}) + E \int_0^T \sqrt{z_{s-}}\, dH_s \tag{6}$$

or

$$\rho^2(P_T, \tilde{P}_T) = 2\tilde{E}(1 - \sqrt{\tilde{z}_0}) + \tilde{E} \int_0^T \sqrt{\tilde{z}_{s-}}\, dH_s \tag{7}$$

where $z_s = \tilde{\zeta}_s/\zeta_s$ and $\tilde{z}_s = \zeta_s/\tilde{\zeta}_s$ (assume that $0/0 = 0$).

Let A and B be functions with bounded variation and paths in $\boldsymbol{D}$. Denote by $\mathscr{E}^B(A)$ the unique solution of the equation

$$dZ_t = Z_{t-}\, dA_t + dB_t$$

[1, Theorem 6.8, p. 192].

COROLLARY 2. *Suppose that $\mathscr{F}_0 = \{\varnothing, \Omega\}$ and that the Hellinger process H is deterministic. Then for fixed $t \geq 0$ we have*

$$\rho^2(P_t, \tilde{P}_t) = \mathscr{E}_t^H(-\tfrac{1}{2}H). \tag{8}$$

REMARK 1. Define the processes $B = (B_t, \mathscr{F}_t)_{t \geqq 0}$ and $\tilde{B} = (\tilde{B}_t, \mathscr{F}_t)_{t \geqq 0}$,

$$B_t = \tfrac{1}{4}\int_{]0,t]} \beta_s d\langle \zeta^c \rangle_s + \int_{]0,t]}\int_{]0,\infty[} (1 - \sqrt{Y(x,s)})^2 \nu_{z,P}(ds,dx)$$

and

$$\tilde{B}_t = \tfrac{1}{4}\int_{]0,t]} \beta_s d\langle \zeta^c \rangle_s + \int_{]0,t]}\int_{]0,\infty[} (1 - \sqrt{\tilde{Y}(x,s)})^2 \nu_{\tilde{z},\tilde{P}}(ds,dx),$$

where $Y(x,s) = \big(1 + x(z_{s-})^{\oplus}\big) \vee 0$, $\tilde{Y}(x,s) = \big(1 + x(\tilde{z}_{s-})^{\oplus}\big) \vee 0$.

In general the Hellinger process H is not P-equivalent to B as stated in [9] (see [9, §1, Remark 2, p. 389]). But when $\tilde{P} \overset{\text{loc}}{\ll} P$ (respectively $P \overset{\text{loc}}{\ll} \tilde{P}$) the Hellinger process H and the process B (respectively $\tilde{B}$) are P- (respectively $\tilde{P}$-equivalent) and we may replace H in (6) (respectively (7)) by B (respectively $\tilde{B}$).

Using Theorem 1 and Corollary 1 we obtain upper and lower bounds for $\rho^2(P_T, \tilde{P}_T)$ with respect to the measures Q and P.

THEOREM 2. *For every stopping time T and every $\varepsilon > 0$ we have*

$$\rho^2(P_T, \tilde{P}_T) \leqq \rho^2(P_0, \tilde{P}_0) + \varepsilon + 2Q(H_T \geqq \varepsilon), \tag{9}$$

$$Q(H_T \geqq \varepsilon) \leqq \varepsilon^2 + (1 + 2/\varepsilon^2)\rho(P_T, \tilde{P}_T). \tag{10}$$

THEOREM 3. *For every stopping time T and every $\varepsilon > 0$ we get*

$$\rho^2(P_T, \tilde{P}_T) \leqq 4\varepsilon + 2P(|1 - \sqrt{z_0}| \geqq \varepsilon) + 2P^{1/2}(H_T \geqq \varepsilon), \tag{11}$$

$$P(|1 - \sqrt{z_0}| \geqq \varepsilon) + P(H_T \geqq \varepsilon) \leqq \varepsilon^2 + (1 + 2/\varepsilon^2)\rho(P_T, \tilde{P}_T). \tag{12}$$

REMARK 2. Because of the symmetry of the Hellinger distance and the Hellinger process we can write formulas (11) and (12) with $\tilde{P}$ and $\tilde{z}_0$ instead of P and z_0.

REMARK 3. The inequalities in Theorems 2 and 3 are useful in the case when $\rho^2(P_T, \tilde{P}_T)$ is near to zero.

REMARK 4. In the discrete time when $\mathscr{F}_0 = \{\varnothing, \Omega\}$ and $\tilde{P} \overset{\text{loc}}{\ll} P$ the upper bound for the Hellinger distance may be improved by

$$\rho^2(P_T, \tilde{P}_T) \leqq 2\varepsilon + 2P(H_T \geqq \varepsilon);$$

we refer to [13] for details.

REMARK 5. F. Liese [7] has obtained a similar bound to (12) between the variation distance of P and $\tilde{P}$ corresponding point processes, and Ju. Kabanov [3] obtained this bound in the general case.

From Theorem 3 we obtain necessary and sufficient conditions for convergence of probability measures in variation.

Let the measurable space, filtration and the measures depend on the parameter $n \geqq 1$, that is, $(\Omega, \mathscr{F}, F) = (\Omega^n, \mathscr{F}^n, F^n)$, $P = P^n$, $\tilde{P} = \tilde{P}^n$. Denote by $\| P^n - \tilde{P}^n \|$ the variation distance between the measures P^n and $\tilde{P}^n$, i.e.

$$\| P^n - \tilde{P}^n \| = \sup_{A \in \mathscr{F}^n} |P^n(A) - \tilde{P}^n(A)|.$$

According to Kraft's inequality [9]

$$\rho^2(P^n, \tilde{P}^n) \leqq 2 \| P^n - \tilde{P}^n \| \leqq 2\rho(P^n, \tilde{P}^n). \tag{13}$$

From Theorem 3 and (13) we have the following result of R. Liptser and A. Shiryayev [9] and Ju. Kabanov [3].

COROLLARY 3. *Condition*

$$\lim_{n \to \infty} P^n\big((\sqrt{z_0^n} - 1)^2 + H_\infty^n \geqq \varepsilon\big) = 0, \quad \forall \varepsilon > 0,$$

is necessary and sufficient for $\| P^n - \tilde{P}^n \| \to 0$ *as* $n \to \infty$.

REMARK 6. Corollary 3 was proved in discrete time in [14, Theorem 2].

After finishing the first draft of this work we learned that similar work has been made by Kabanov, Liptser, and Shiryayev [5] and also by Memin and Shiryayev [10]. The interested reader should also see the paper by Jacod [2].

2. Integral representation for the Hellinger distance.

In this paragraph we prove Theorem 1, Corollary 1, and Corollary 2.

PROOF OF THEOREM 1. We consider the process $U = (U_t, \mathscr{F}_t)_{t \geqq 0}$ where

$$U_t = (\sqrt{\zeta_t} - \sqrt{\tilde{\zeta}_t})^2. \tag{14}$$

We define stopping times τ, $\tilde{\tau}$, σ and τ_k, $\tilde{\tau}_k$, σ_k for $k \geqq 1$ by the formulas

$$\begin{aligned} \tau &= \inf\{t \geqq 0 : \zeta_t = 0\}, \quad \tilde{\tau} = \inf\{t \geqq 0 : \tilde{\zeta}_t = 0\}, \\ \tau_k &= \inf\{t \geqq 0 : \zeta_t \leqq 2/(k+1)\}, \quad \tilde{\tau}_k = \inf\{t \geqq 0 : \tilde{\zeta}_t \leqq 2/(k+1)\}, \\ \sigma &= \tau \wedge \tilde{\tau}, \quad \sigma_k = \tau_k \wedge \tilde{\tau}_k \end{aligned} \tag{15}$$

where $\inf\{\varnothing\} = \infty$.

First of all we show that for every stopping time T

$$\rho^2(P_T, \tilde{P}_T) = E_Q U_{T\wedge\sigma} = \lim_{t\to\infty} \lim_{k\to\infty} E_Q U_{T\wedge\sigma_k\wedge t}. \tag{16}$$

Because for every $t \geqq 0$

$$\zeta_t + \tilde{\zeta}_t = 2, \tag{17}$$

we get from (14) that

$$U_t = 2 - 2\sqrt{\zeta_t \tilde{\zeta}_t}. \tag{18}$$

Since the point $\{0\}$ is an absorbing state for regular martingales ζ and $\tilde{\zeta}$, the point $\{2\}$ is an absorbing state for the process U and by (15) and (18) we have

$$\sigma = \inf\{t \geqq 0 : U_t = 2\}.$$

Hence, from (14) and Definition 1 we have

$$\begin{aligned} \rho^2(P_T, \tilde{P}_T) &= E_Q U_T I(\sigma \geqq T) + E_Q U_T I(\sigma < T) \\ &= E_Q U_{T\wedge\sigma} I(\sigma \geqq T) + E_Q U_\sigma I(\sigma < T) = E_Q U_{T\wedge\sigma} \end{aligned} \tag{19}$$

where $I(\cdot)$ is the indicator function.

Because $\sigma_k \leqq \sigma$, $\lim_{k\to\infty} \sigma_k = \sigma$ and since regular martingales ζ and $\tilde{\zeta}$ have left-hand limits, we obtain from (15)

$$\lim_{t\to\infty} \lim_{k\to\infty} \zeta_{T\wedge\sigma_k\wedge t} = \zeta_{T\wedge\sigma}, \quad \lim_{t\to\infty} \lim_{k\to\infty} \tilde{\zeta}_{T\wedge\sigma_k\wedge t} = \tilde{\zeta}_{T\wedge\sigma}. \tag{20}$$

According to (18), $U_t \leqq 2$ for every $t \geqq 0$. Therefore, from (14) and (20) by the Lebesque theorem we get

$$\lim_{t\to\infty} \lim_{k\to\infty} E_Q U_{T\wedge\sigma_k\wedge t} = E_Q U_{T\wedge\sigma}. \tag{21}$$

But (19) and (21) give (16).

Since ζ and $\tilde{\zeta}$ are regular martingales with respect to Q, by the Itô formula we have for $S_k = T \wedge \sigma_k \wedge t$

$$
\begin{aligned}
U_{S_k} = {} & U_0 - \int_0^{S_k} \sqrt{\frac{\zeta_{s-}}{\tilde{\zeta}_{s-}}}\, d\tilde{\zeta}_s - \int_0^{S_k} \sqrt{\frac{\tilde{\zeta}_{s-}}{\zeta_{s-}}}\, d\zeta_s + \\
& + \frac{1}{4}\left\{ \int_0^{S_k} \sqrt{\frac{\zeta_{s-}}{(\tilde{\zeta}_{s-})^3}}\, d\langle \tilde{\zeta}^c \rangle_s + \right. \\
& \left. + \int_0^{S_k} \sqrt{\frac{\tilde{\zeta}_{s-}}{(\zeta_{s-})^3}}\, d\langle \zeta^c \rangle_s + \int_0^{S_k} \sqrt{\frac{\zeta_{s-}}{(\tilde{\zeta}_{s-})^3}}\, d\langle \zeta^c \rangle_s - 2\int_0^{S_k} \frac{d\langle \tilde{\zeta}^c, \zeta^c \rangle}{\sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}} \right\} + \\
& + \sum_{0 < s \leqq S_k} \Delta U_s + \sum_{0 < s \leqq S_k} \left\{ \sqrt{\frac{\zeta_{s-}}{\tilde{\zeta}_{s-}}}\, \Delta\tilde{\zeta}_s + \sqrt{\frac{\tilde{\zeta}_{s-}}{\zeta_{s-}}}\, \Delta\zeta_s \right\}
\end{aligned}
\tag{22}
$$

where $\langle \tilde{\zeta}^c, \zeta^c \rangle$ is the mutual quadratic characteristic of $\tilde{\zeta}^c$ and ζ^c.

In below we write the process U as $U = U^c + U^d$, where U^c involves the continuous martingales ζ^c, $\tilde{\zeta}^c$ and their mutual quadratic characteristic, while U^d involves the jumps of U and discontinuous martingale part of U. From Lemma 2.3. in [9] we have for every $s \geqq 0$

$$
\langle \zeta^c \rangle_s = \langle \tilde{\zeta}^c \rangle_s = -\langle \tilde{\zeta}^c, \zeta^c \rangle_s. \tag{23}
$$

From (22), (23) and (3) we get

$$
U^c_{S_k} = U_0 - \int_0^{S_k} \sqrt{\frac{\zeta_{s-}}{\tilde{\zeta}_{s-}}}\, d\tilde{\zeta}^c_s - \int_0^{S_k} \sqrt{\frac{\tilde{\zeta}_{s-}}{\zeta_{s-}}}\, d\zeta^c_s + \tfrac{1}{4}\int_0^{S_k} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\, \beta_s d\langle \zeta^c \rangle_s. \tag{24}
$$

Because of (17) we have for every $s \geqq 0$

$$
\zeta^d_s + \tilde{\zeta}^d_s = 2 - \zeta^c_s - \tilde{\zeta}^c_s,
$$

and hence,

$$
\begin{aligned}
\zeta^d_s + \tilde{\zeta}^d_s &= \sum_{0 < t \leqq s} (\Delta\zeta_t + \Delta\tilde{\zeta}_t) = 0, \\
\Delta\zeta_s &= -\Delta\tilde{\zeta}_s.
\end{aligned}
\tag{25}
$$

According to (2), (18), and (25) we get for $s \leqq S_k$

$$
\begin{aligned}
\Delta U_s &= 2(\sqrt{\zeta_{s-}\tilde{\zeta}_{s-}} - \sqrt{\zeta_s\tilde{\zeta}_s}) \\
&= 2\sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\,(1 - \sqrt{\lambda_s(\Delta\zeta_s)\tilde{\lambda}_s(\Delta\zeta_s)}) \\
&= \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\,(\sqrt{\tilde{\lambda}_s(\Delta\zeta_s)} - \sqrt{\lambda_s(\Delta\zeta_s)})^2 - \sqrt{\frac{\zeta_{s-}}{\tilde{\zeta}_{s-}}}\, \Delta\tilde{\zeta}_s - \sqrt{\frac{\tilde{\zeta}_{s-}}{\zeta_{s-}}}\, \Delta\zeta_s.
\end{aligned}
\tag{26}
$$

From (22) and (26) we obtain

$$
\begin{aligned}
U^d_{S_k} = &- \int_0^{S_k} \left(\sqrt{\zeta_{s-}/\tilde{\zeta}_{s-}}\, d\tilde{\zeta}^d_s + \sqrt{\tilde{\zeta}_{s-}/\zeta_{s-}}\, d\zeta^d_s\right) + \\
&+ \int_{]0,S_k]} \int_{]0,\infty[} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}} \left(\sqrt{\tilde{\lambda}_s(x)} - \sqrt{\lambda_s(x)}\right)^2 \mu_\zeta(ds,dx).
\end{aligned} \tag{27}
$$

Since the martingale parts on the right-hand side of (24) and (27) have zero expectation and from the properties of compensators [12, §4, p. 211] we get

$$
E_Q U_{S_k} = E_Q U + E_Q \int_0^{S_k} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\, dH_s \tag{28}
$$

where $H = (H_s, \mathscr{F}_s)_{s \geqq 0}$ is the Hellinger process.

Because $S_k = T \wedge \sigma_k \wedge t$, from (16), (28) after $\lim_{t\to\infty} \lim_{k\to\infty}$ we have

$$
\rho^2(P_T, \tilde{P}_T) = E_Q U + E_Q \int_0^{T\wedge\sigma} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\, dH_s. \tag{29}
$$

But the point $\{0\}$ is an absorbing state for regular martingales ζ and $\tilde{\zeta}$, so that we may replace σ by ∞ in (29). Moreover, according to (16), $E_Q U_0 = \rho^2(P_0, \tilde{P}_0)$. Hence, from (29) it follows (5). Theorem 1 is proved.

Proof of Corollary 1. To prove (6) we show that for $k > 1$ and $S_k = T \wedge \sigma_k \wedge t$ we have

$$
E_Q \int_0^{S_k} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\, dH_s = E \int_0^{S_k} \sqrt{z_{s-}}\, dH_s, \tag{30}
$$

$$
E_Q U_0 = 2E(1 - \sqrt{z_0}) \tag{31}
$$

where E is expectation with respect to P.

Because ζ is a regular martingale with respect to Q we have for each stopping time T

$$
\begin{aligned}
E_Q \sqrt{\zeta_T \tilde{\zeta}_T} &= E_Q \sqrt{\zeta_T \tilde{\zeta}_T}\, I(\zeta_T > 0) \\
&= E_Q \zeta_T \sqrt{z_T}\, I(\zeta_T > 0) = E_0 \zeta_\infty \sqrt{z_T}\, I(\zeta_T > 0)
\end{aligned}
$$

Hence, according to Theorem 15 in [11, Chapter VII, §2, p. 144] we get

$$
E_Q \int_0^{S_k} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\, dH_s = E_Q \zeta_\infty \int_0^{S_k} \sqrt{z_{s-}}\, dH_s \tag{32}
$$

Because $P \ll Q$, for each non-negative $\mathscr{F}$-measurable random variable η we have

$$E_Q \zeta_\infty \eta = E\eta. \tag{33}$$

Therefore, from (32) and (33) we obtain (30).

From (18)

$$E_Q U_0 = 2 - 2E_Q\sqrt{\zeta_0 \tilde{\zeta}_0}.$$

According to (33)

$$\begin{aligned} E_Q\sqrt{\zeta_0 \tilde{\zeta}_0} &= E_Q\sqrt{\zeta_0 \tilde{\zeta}_0}\, I(\zeta_0 > 0) = E_Q \zeta_0 \sqrt{z_0}\, I(\zeta_0 > 0) \\ &= E_Q \zeta_\infty \sqrt{z_0}\, I(\zeta_0 > 0) = E\sqrt{z_0}\, I(\zeta_0 > 0) = E\sqrt{z_0} \end{aligned}$$

since $P(\zeta_0 > 0) = 1$.

From (16), (28), (30), and (31) after $\lim_{t\to\infty} \lim_{k\to\infty}$ we get

$$\rho^2(P_T, \tilde{P}_T) = 2E(1 - \sqrt{z_0}) + E\int_0^{T\wedge\sigma} \sqrt{z_{s-}}\, dH_s. \tag{34}$$

Because the point $\{0\}$ is an absorbing state for the process $z = (z_s, \mathscr{F}_s)_{s \geqq 0}$ and $z_{\tilde{\tau}} = 0$, we may substitute $\sigma = \tau \wedge \tilde{\tau}$ by τ in (34).

By definition

$$\tau = \inf\{t \geqq 0 : \zeta_t = 0\} = \inf\{t \geqq 0 : z_t = \infty\}.$$

According to Lemma 2.2 in [4]

$$P(\sup_{0 \leqq s < \infty} z_s = \infty) = 0.$$

Hence, $P(\tau < \infty) = 0$ and we may replace in (34) σ by ∞. The representation (6) is proved.

Because of the symmetry of the Hellinger distance and the Hellinger process from (6) it follows (7). Corollary 1 is proved.

Proof of Corollary 2. We fix $t > 0$. Suppose also that $\mathscr{F}_0 = \{\varnothing, \Omega\}$ and H is deterministic.

From (18) we have

$$\sqrt{\zeta_t \tilde{\zeta}_t} = 1 - \tfrac{1}{2} U_t. \tag{35}$$

Denote by $Z_t = \rho^2(P_t, \tilde{P}_t)$. Then we have from (5) and (35)

$$Z_t = \int_0^t E_Q\sqrt{\zeta_{s-} \tilde{\zeta}_{s-}}\, dH_s = H_t - \tfrac{1}{2}\int_0^t Z_{s-}\, dH_s, \tag{36}$$

because H is deterministic. But from (36) we get (8). Corollary 2 is proved.

3. Upper and lower bounds for the Hellinger distance.

In this section we prove Theorem 2 and Theorem 3.

PROOF OF THEOREM 2. Let T be a finite stopping time. To prove (9) we show that for every finite stopping time S

$$\rho^2(P_T, \tilde{P}_T) \leqq \rho^2(P_{T\wedge S}, \tilde{P}_{T\wedge S}) + 2Q(S \leqq T). \tag{37}$$

Really, since $U_T \leqq 2$ for every stopping time T, from (14) and Definition 1 we have

$$\begin{aligned}\rho^2(P_T, \tilde{P}_T) &= E_Q U_T = E_Q U_T I(T < S) + E_Q U_T I(S \leqq T)\\ &\leqq E_Q U_{T\wedge S} + E_Q U_T I(S \leqq T) < \rho^2(P_{T\wedge S}, \tilde{P}_{T\wedge S}) + 2Q(S \leqq T).\end{aligned}$$

Let stopping time τ_ε be defined by the formula

$$\tau_\varepsilon = \inf\{0 \leqq t \leqq T : H_t \geqq \varepsilon\} \tag{38}$$

where $\inf\{\varnothing\} = \infty$.

Because the Hellinger process $H = (H_t, \mathscr{F}_t)_{t\geqq 0}$ is predictable, then the stopping time τ_ε is predictable, too. According to [1, Proposition 1.9, p. 9] there is a sequence of stopping times $(\tau_\varepsilon^k)_{k\geqq 1}$ such that for every $k \geqq 1$, $\tau_\varepsilon^k \leqq \tau_\varepsilon^{k+1}$, $\tau_\varepsilon^k < \tau_\varepsilon$ and

$$\lim_{k\to\infty} \tau_\varepsilon^k = \tau_\varepsilon \quad (Q-\text{a.s.}).$$

Using (37) with $S = \tau_\varepsilon^k$ we have

$$\rho^2(P_T, \tilde{P}_T) \leqq \rho^2(P_{T\wedge \tau_\varepsilon^k}, \tilde{P}_{T\wedge \tau_\varepsilon^k}) + 2Q(\tau_\varepsilon^k \leqq T). \tag{39}$$

Since by (17), $\zeta_{s-}\tilde{\zeta}_{s-} \leqq 1$ for every $s > 0$ and $H_{T\wedge\tau_\varepsilon^k} \leqq \varepsilon$, from Theorem 1 we get

$$\begin{aligned}\rho^2(P_{T\wedge \tau_\varepsilon^k}, \tilde{P}_{T\wedge \tau_\varepsilon^k}) &= \rho^2(P_0, \tilde{P}_0) + E_Q \int_0^{T\wedge\tau_\varepsilon^k} \sqrt{\zeta_{s-}\tilde{\zeta}_{s-}}\, dH_s\\ &\leqq \rho^2(P_0, \tilde{P}_0) + \varepsilon.\end{aligned} \tag{40}$$

Because the Hellinger process H is non-decreasing, from the equality

$$\bigcap_k \{\tau_\varepsilon^k \leqq T\} = \{\tau_\varepsilon \leqq T\} \tag{41}$$

we have

$$\begin{aligned}&\lim_{k\to\infty} Q(\tau_\varepsilon^k \leqq T) = Q(\tau_\varepsilon \leqq T) = Q\left(\sup_{0\leqq s\leqq T} H_s \geqq \varepsilon\right)\\ &= Q(H_T \geqq \varepsilon)\end{aligned} \tag{42}$$

Hence, from (39), (40), and (42) after $\lim_{k\to\infty}$ we get (9) for every finite stopping time T.

To prove (9) for an arbitrary stopping time T we consider a finite stopping time $T \wedge t$. From above we have

$$\begin{aligned}\rho^2(P_{T\wedge t}, \tilde{P}_{T\wedge t}) &\leqq \rho^2(P_0, \tilde{P}_0) + \varepsilon + 2Q(H_{T\wedge t} \geqq \varepsilon)\\ &\leqq \rho^2(P_0, \tilde{P}_0) + \varepsilon + 2Q(H_T \geqq \varepsilon)\end{aligned}$$

because the Hellinger process is non-decreasing. From this after $\lim_{t\to\infty}$ we get (9).

Here we will not prove (10), since this inequality directly follows from (12) and from the analog of (12) with P and z_0 substituted by $\tilde{P}$ and $\tilde{z}_0$.

PROOF OF THEOREM 3. We prove inequality (11). Let T be a finite stopping time. First of all we show that for every finite stopping time S

$$\rho^2(P_T, \tilde{P}_T) \leqq \rho^2(P_{T\wedge S}, \tilde{P}_{T\wedge S}) + 2P^{1/2}(S \leqq T). \tag{43}$$

Really, by (14) and Definition 1

$$\rho^2(P_T, \tilde{P}_T) = \rho^2(P_{T\wedge S}, \tilde{P}_{T\wedge S}) + E_Q(U_T - U_{T\wedge S}). \tag{44}$$

According to (18) we have

$$E_Q(U_T - U_{T\wedge S}) = 2E_Q(\sqrt{\zeta_{T\wedge S}\tilde{\zeta}_{T\wedge S}} - \sqrt{\zeta_T\tilde{\zeta}_T}). \tag{45}$$

Since ζ and $\tilde{\zeta}$ are regular martingale with respect to Q and

$$P(\inf_{0\leqq s<\infty}\zeta_s > 0) = 1,$$

we get from (33)

$$\begin{aligned}E_Q\sqrt{\zeta_T\tilde{\zeta}_T} &= E_Q\sqrt{\zeta_T\tilde{\zeta}_T}I(\zeta_T > 0) = E_Q\zeta_T\sqrt{z_T}I(\zeta_T > 0)\\ &= E_Q\zeta_\infty\sqrt{z_T}I(\zeta_T > 0) = E\sqrt{z_T}I(\zeta_T > 0) = E\sqrt{z_T}.\end{aligned} \tag{46}$$

In the same way we have

$$E_Q\sqrt{\zeta_{T\wedge S}\tilde{\zeta}_{T\wedge S}} = E\sqrt{z_{T\wedge S}}. \tag{47}$$

By Lemma 2.2 in [9], $Ez_{T\wedge S} \leqq 1$ and then from (45), (46), and (47) it follows that

$$
\begin{aligned}
E_Q(U_T - U_{T\wedge S}) &= 2E(\sqrt{z_{T\wedge S}} - \sqrt{z_T}) \\
&= 2E(\sqrt{z_{T\wedge S}} - \sqrt{z_T}) I(S \leqq T) \\
&\leqq 2E\sqrt{z_{T\wedge S}} I(S < T) \\
&\leqq 2(Ez_{T\wedge S})^{1/2} P^{1/2}(S < T) \\
&\leqq 2P^{1/2}(S \leqq T).
\end{aligned} \tag{48}
$$

But (44) and (48) give (43).

Let stopping time τ_ε be defined by (38).

Since τ_ε is predictable stopping time, there is non-decreasing sequence $(\tau_\varepsilon^k)_{k\geqq 1}$ of stopping times such that $\tau_\varepsilon^k < \tau_\varepsilon$ and $\lim_{k\to\infty} \tau_\varepsilon^k = \tau_\varepsilon$.

Using (43) with $S = \tau_\varepsilon^k$ we have

$$
\rho^2(P_T, \tilde{P}_T) \leqq \rho^2(P_{T\wedge\tau_\varepsilon^k}, \tilde{P}_{T\wedge\tau_\varepsilon^k}) + 2P^{1/2}(\tau_\varepsilon^k \leqq T). \tag{49}
$$

Since $H_{T\wedge\tau_\varepsilon^k} \leqq \varepsilon$, from Corollary 1 we get

$$
\begin{aligned}
\rho^2(P_{T\wedge\tau_\varepsilon^k}, \tilde{P}_{T\wedge\tau_\varepsilon^k}) &= 2E(1 - \sqrt{z_o}) + E\int_0^{T\wedge\tau_\varepsilon^k} \sqrt{z_{s-}}\, dH_s \\
&\leqq 2\varepsilon + 2P(|1 - \sqrt{z_0}| \geqq \varepsilon) + E\Big(\sup_{0\leqq s<\infty} \sqrt{z_s} H_{T\wedge\tau_\varepsilon}\Big) \\
&\leqq 2\varepsilon + 2P(|1 - \sqrt{z_0}| \geqq \varepsilon) + \varepsilon E\Big(\sup_{0\leqq s<\infty} \sqrt{z_s}\Big).
\end{aligned} \tag{50}
$$

By Lemma 2.2 in [9] we have for every $a > 0$ that

$$
P\Big(\sup_{0\leqq s<\infty} z_s \geqq a\Big) \leqq 1/a.
$$

Hence, $E\sup_{0\leqq s<\infty}\sqrt{z_s} \leqq 2$ and from (50) it follows that

$$
\rho^2(P_{T\wedge\tau_\varepsilon^k}, \tilde{P}_{T\wedge\tau_\varepsilon^k}) \leqq 4\varepsilon + 2P(|1 - \sqrt{z_0}| \geqq \varepsilon. \tag{51}
$$

From (41) we get

$$
\begin{aligned}
\lim_{k\to\infty} P(\tau_\varepsilon^k \leqq T) &= P(\tau_\varepsilon \leqq T) = P\Big(\sup_{0\leqq s\leqq T} H_s \geqq \varepsilon\Big) \\
&= P(H_T \geqq \varepsilon)
\end{aligned} \tag{52}
$$

since the Hellinger process is non-decreasing.

But from (43), (51), and (52) after $\lim_{k\to\infty}$ we obtain (12) for a finite stopping time T. To prove (12) for an arbitrary stopping time T it is enough to consider a finite stopping time $T\wedge t$ and make $\lim_{t\to\infty}$.

We prove the inequality (12), i.e. the lower bound for the Hellinger distance. Because $Ez_0 \leqq 1$.

$$2E(1-\sqrt{z_0}) \geqq E(1-\sqrt{z_0})^2. \tag{53}$$

Hence, according to Corollary 1 we have

$$\rho^2(P_T, \tilde{P}_T) \geqq E(1-\sqrt{z_0})^2 + E\left(\inf_{0\leqq s\leqq T}\sqrt{z_s}\,H_T\right). \tag{54}$$

By the Chebyshev inequality we get for every $\varepsilon > 0$

$$E(1-\sqrt{z_0})^2 \geqq \varepsilon^2 P(|1-\sqrt{z_0}| \geqq \varepsilon) \tag{55}$$

$$E\left(\inf_{0\leqq s\leqq T}\sqrt{z_s}\,H_T\right) \geqq \varepsilon^2 P\left(\inf_{0\leqq s\leqq T}\sqrt{z_s}\,H_T \geqq \varepsilon^2\right). \tag{56}$$

For arbitrary positive random variables X and Y and constants $a > 0$ and $b > 0$ we have

$$P(XY \geqq a) \geqq P(X \geqq a/b) - P(Y \leqq b),$$

so we get

$$P\left(\inf_{0\leqq s\leqq T}\sqrt{z_s}\,H_T \geqq \varepsilon^2\right) \geqq P(H_T \geqq \varepsilon) - P\left(\inf_{0\leqq s\leqq T}\sqrt{z_s} \leqq \varepsilon\right). \tag{57}$$

From Lemma 2.2 in [9] it follows that

$$\begin{aligned} P\left(\inf_{0\leqq s\leqq T}\sqrt{z_s} \leqq \varepsilon\right) &\leqq \|P_T - \tilde{P}_T\| + \tilde{P}\left(\inf_{0\leqq s\leqq T}\sqrt{z_s} \leqq \varepsilon\right) \\ &\leqq \|P_T - \tilde{P}_T\| + \varepsilon^2. \end{aligned} \tag{58}$$

From Kraft's inequality, (54), (55), (56), (57), and (58) we get

$$\begin{aligned} 2\|P_T - \tilde{P}_T\| \geqq \rho^2(P_T, \tilde{P}_T) \geqq \varepsilon^2 P(|1-\sqrt{z_0}| \geqq \varepsilon) + \\ + \varepsilon^2 P(H_T \geqq \varepsilon) - \varepsilon^2 \|P_T - \tilde{P}_T\| - \varepsilon^4. \end{aligned}$$

Again from Kraft's inequality we get

$$\begin{aligned} P(|1-\sqrt{z_0}| \geqq \varepsilon) + P(H_T \geqq \varepsilon) &\leqq \varepsilon^2 + (1+2/\varepsilon^2)\|\tilde{P}_T - P_T\| \\ &\leqq \varepsilon^2 + (1+2/\varepsilon^2)\rho(P_T, \tilde{P}_T), \end{aligned}$$

i.e. the inequality (12). Theorem 3 is proved.

4. Examples.

We give some examples of the Hellinger process. In these examples the stopping time σ defined in (15) plays an essential role.

EXAMPLE 1. Let $\xi = (\xi_k)_{k \geqq 1}$ and $\tilde{\xi} = (\tilde{\xi}_k)_{k \geqq 1}$ be two sequences of independent random variables and $P, \tilde{P}, P_k, \tilde{P}_k$ be the distributions of $\xi, \tilde{\xi}, \xi_k, \tilde{\xi}_k$. We set

$$\mathscr{F}_k = \sigma\{\xi_1, \xi_2, \dots \xi_k, \tilde{\xi}_k\}, \quad \mathscr{F} = \mathscr{F}_\infty.$$

Then the Hellinger process $H = (H_k, \mathscr{F}_k)_{k \geqq 1}$ is defined by

$$H_k = \sum_{i=1}^{k \wedge \sigma} \rho^2(P_i, \tilde{P}_i). \tag{59}$$

EXAMPLE 2. Let $(\Omega, \mathscr{F})$ be the measurable space of piece-wise constant right-continuous functions $X = (X_t)_{t \geqq 0}$ with $X_0 = 0$, $X_t = X_{t-} + \Delta X_t$, $\Delta X_t \in \{0,1\}$. Let $\mathscr{F}$ be Borel σ-algebra and $\mathscr{F}_t = \sigma\{X_s | s \leqq t\}$ for every $t \geqq 0$.

Let X be counting process on $(\Omega, \mathscr{F})$ and P and $\tilde{P}$ be probability measures on $(\Omega, \mathscr{F})$. Denote by A, $\tilde{A}$ and G the compensators of X with respect to the measures P, $\tilde{P}$ and $Q = \frac{1}{2}(P + \tilde{P})$. Note that $G = \frac{1}{2}(A + \tilde{A})$.

According to [8] the Hellinger process $H = (H_t, \mathscr{F}_t)_{t \geqq 0}$ is given by

(60)
$$H_t = \int_0^{t \wedge \sigma} (\sqrt{\tilde{\rho}_s} - \sqrt{\rho_s})^2 dG_s + \sum_{0 < s \leqq t \wedge \sigma} (\sqrt{1 - \Delta \tilde{A}_s} - \sqrt{1 - \Delta A_s})^2$$

where

$$\rho_s = \frac{dA_s}{dG_s} \quad \text{and} \quad \tilde{\rho}_s = \frac{d\tilde{A}_s}{dG_s}.$$

We suppose that the compensators A and B are deterministic and, for simplicity, continuous. Then the Hellinger process (60) is continuous and deterministic, too, and from (8) we have that

$$\rho^2(P_t, \tilde{P}_t) = e^{-\frac{1}{2}H_t} \int_0^t e^{\frac{1}{2}H_s} dH_s = 2 - 2e^{-\frac{1}{2}H_t},$$

when $t \geqq 0$.

EXAMPLE 3. Let $(\Omega, \mathscr{F})$ be the measurable space of continuous functions $X = (X_t)_{t \geqq 0}$, $X_0 = 0$, $\mathscr{F}$ be Borel σ-algebra and $\mathscr{F}_t = \sigma\{X_s | s \leqq t\}$ for every $t \geqq 0$.

Let P and $\tilde{P}$ be unique probability measures on $(\Omega, \mathscr{F})$ which correspond to diffusion type processes satisfying the stochastic differential equations

$$dX_t = a(t, X)dt + dw_t, \quad X_0 = 0,$$
$$d\tilde{X}_t = \tilde{a}(t, \tilde{X})dt + dw_t, \quad \tilde{X}_0 = 0.$$

Here $w = (w_t)_{t>0}$ is Wiener process and $a(t,X)$, $\tilde{a}(t,X)$ are non-anticipating functionals with

$$\int_0^t a^2(s,X)ds < \infty, \quad \int_0^t \tilde{a}^2(s,X)ds < \infty \quad (Q-\text{a.s.}).$$

Then according to [8] the Hellinger process $H = (H_t, \mathscr{F}_t)_{t \geqq 0}$ is given by

$$H_t = \tfrac{1}{4}\int_0^{t\wedge\sigma} (a(s,X) - \tilde{a}(s,X))^2 ds \tag{61}$$

REMARK 7. It should be noted that in the case $P \sim \tilde{P}$ we have $\sigma = \infty$ $Q-$a.s. When $\sigma \neq \infty$ for practical purpose it is convenient instead of H to consider the process H' which is defined by (53), (60), and (61) accordingly with σ substituted by ∞.

Because $H_t \leqq H_t'$ for every t, the inequalities (10) and (11) of Theorems 2 and 3 remain true. Since for every stopping time T

$$P\Big(\inf_{0\leqq s\leqq T} \zeta_s = 0\Big) = 0, \quad \tilde{P}\Big(\inf_{0\leqq s\leqq T} \tilde{\zeta}_s = 0\Big) = 0,$$

we have

$$\begin{aligned} P(H_T \geqq \varepsilon) &\geqq P\Big(\{H_T \geqq \varepsilon\} \cap \Big\{\inf_{0<s<T} \zeta_s\tilde{\zeta}_s > 0\Big\}\Big) \\ &\geqq P(H_T' \geqq \varepsilon) - P\Big(\inf_{0\leqq s\leqq T} \zeta_s\tilde{\zeta}_s = 0\Big) \\ &\geqq P(H_T' > \varepsilon) - \|P_T - \tilde{P}_T\|. \end{aligned}$$

Hence, instead of (10) and (12) for the process H' we can write

$$Q(H_T' \geqq \varepsilon) < \varepsilon^2 + (2 + 2/\varepsilon^2)\rho(P_T, \tilde{P}_T),$$
$$P(|1 - \sqrt{z_0}| \geqq \varepsilon) + P(H_T' \geqq \varepsilon) \leqq \varepsilon^2 + (2 + 2/\varepsilon^2)\rho(P_T, \tilde{P}_T).$$

ACKNOWLEDGEMENTS. The authors want to thank Professor A. N. Shiryayev for improving an earlier version of this paper and also Professor R. Sh. Liptser for useful Remark 7.

The first author is grateful to the exchange program between the Finnish Academy and the Hungarian Academy of Sciences for a stay in Budapest in September, 1983.

REFERENCES

1. J. Jacod, *Calcul stochastique et problémes de martingales* (Lecture Notes in Math. 714), Springer-Verlag, Berlin - Heidelberg - New York, 1979.
2. J. Jacod, *Processus de Hellinger, absolute continuité, contiguité*, Tech. Rep., Dept. de Math., Université de Rennes, Rennes, 1984.
3. Ju. M. Kabanov, *The strong convergence of random processes*, Short Communications of the International Congress Math., vol. 8, p. 5, Warzawa, 1982.
4. Ju. M. Kabanov, R. Sh. Liptser, and A. N. Shiryayev, *Absolute continuity and singularity of locally absolutely continuous probability distributions*, I, Math. USSR-Sb. 35 (1979), 631–680.
5. Ju. M. Kabanov, R. Sh. Liptser, and A. N. Shiryayev, *On the variation distance for probability measures defined on a filtered space*, Probab. Theory Related Fields 71 (1986), 19–35.
6. C. H. Kraft, *Some conditions for consistency and uniform consistency of statistical procedures*, Univ. California Publ. Statist. 2 (1955), 135–141.
7. F. Liese, *An estimation of Hellinger integrals of point processes*, Friedrich Schiller-Universität Jena, Forschungsergebnisse. Eingang: 12.1.1983, Nr. N/83/1.
8. R. Sh. Liptser and A. N. Shiryayev, *On contiguity of probability measures corresponding to semimartingales*, Anal. Math. 11 (1985), 93–124.
9. R. Sh. Liptser and A. N. Shiryayev, *On the problem of "predictable" criteria of contiguity*, in *Probability theory and mathematical statistics* (Proc. Tbilisi, 1982), eds. K. Ito and J. V. Prokhorov (Lecture Notes in Math. 1021), pp. 386–418. Springer-Verlag, Berlin - Heidelberg - New York, 1983.
10. J. Memin and A. N. Shiryayev, *Distance de Helliger–Kakutani des lois correspondant à deux processus à accroisement independants: critères d'absolute continuité et de singularité*, Z. Wahrsch. Verw. Gebiete 70 (1985), 67–89.
11. P.-A. Meyer, *Probability and potentials*, Mir, Moscow, 1973 (in Russian).
12. A. N. Shiryayev, *Martingales: recent developments, results and applications*, Internat. Statist. Rev. 49 (1981), 199–233.
13. L. Ju. Vostrikova, *On criteria for C_n-consistency of estimators*, Stochastics 11 (1984), 265–290.
14. L. Ju. Vostrikova, *On necessary and sufficient conditions for convergence of probability measures in variation*, Stochastic Process. Appl. 18 (1984), 99–112.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HELSINKI
SF-00100 HELSINKI
FINLAND

COMPUTER CENTER OF EÖTVÖS
LORAND UNIVERSITY
BOGDÁNFY U. 10/B.
BUDAPEST
HUNGARY