The Annals of Probability
1986, Vol. 14, No. 1, 326–335

FISHER INFORMATION AND DETECTION OF A EUCLIDEAN PERTURBATION OF AN INDEPENDENT STATIONARY PROCESS

BY J. MICHAEL STEELE

Princeton University

An independent stationary process $\{X_t\}_{t=1}^{\infty}$ in $\mathbb{R}^n$ is perturbed by a sequence of Euclidean motions to obtain a new process $\{Y_t\}_{t=1}^{\infty}$. Criteria are given for the singularity or equivalence of these processes. When the distribution of the X process has finite Fisher information, the criteria are necessary and sufficient. Moreover, it is proved that it is exactly under the condition of finite Fisher information that the criteria are necessary and sufficient.

1. Introduction. The purpose of this article is to provide results which tell when an independent stationary process in $\mathbb{R}^d$, which has been perturbed by Euclidean motions, can be distinguished from the original process. The first results of this nature are due to Feldman (1961), Shepp (1965), and Renyi (1967). In particular, Shepp settled the question completely in the case of *translations* in $\mathbb{R}^1$.

Here we will obtain extensions of Shepp's results to $\mathbb{R}^d$, but, more pointedly, we extend the group of perturbations from translations to the whole group of proper rigid Euclidean motions (i.e., rotations, translations, and their compositions).

One benefit of this extension comes from having to spell out the proper analogue of finite Fisher information. A second benefit comes from the fact that the noncommuting perturbations studied here do not have a convenient harmonic analysis. This forces one to develop tools which are different from those used in the commutative case. A final benefit comes from seeing how several simple facts from the local theory of Lie groups can be put to work on a statistical problem.

Before stating the main results some notation needs to be developed. We will let G denote any closed continuous subgroup of the group of rigid motions of $\mathbb{R}^d$. It is known that such a G must actually be a differentiable manifold, and hence that there is a tangent space T_eG at the identity e. The elements $A \in T_eG$ can be viewed as matrices, and for all t one can define a new matrix $\exp(tA)$ by the converging sum $\sum_{n=0}^{\infty}(t^n/n!)A^n$, which we will denote by $p(t)$. The set $\{p(t): t \in \mathbb{R}\}$ can be verified to be a group, and it is called the one-dimensional subgroup generated by A.

To concretize these notions and to develop some facts which will be used later, we now consider the important special case where G is the full group of rigid

Received December 1981; revised March 1985.
AMS 1980 *subject classifications.* Primary 60G30; secondary 60B15
Key words and phrases. Fisher information, Kakutani's product theorem, product measures, Euclidean motions, singular processes, Hellinger integrals.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.* www.jstor.org

motions on $\mathbb{R}^2$. The usual real parametrization of G is given by the 3×3 matrices.

$$R(\theta, a, b) = \begin{pmatrix} \cos\theta & -\sin\theta & a \\ \sin\theta & \cos\theta & b \\ 0 & 0 & 1 \end{pmatrix}, \qquad -\pi < \theta \le \pi, \qquad -\infty < a, b < \infty.$$

Here, by parametrizing $\mathbb{R}^2$ by the two-dimensional set in $\mathbb{R}^3$ given by

$$\left\{ \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} : -\infty < x, y < \infty \right\},$$

one sees that matrix multiplication by $R(\theta, a, b)$ corresponds to a rotation by θ followed by a translation by a and b along the x and y axes. A basis for T_eG is given by

$$A_\theta = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_a = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_b = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

and direct power series computation establishes

$$\exp(\theta A_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \exp(aA_a) = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\exp(bA_b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}.$$

If $H = \{p(t): t \in \mathbb{R}\}$ is a one-parameter subgroup, we say a measure μ is invariant under H if $\mu(p(t)B) = \mu(B)$ for all measurable B and all t. Naturally, Lebesgue measure is invariant under any subgroup of the rigid motions. Also, note that if μ has a radially symmetric density (say in $\mathbb{R}^2$), then μ is invariant under the subgroup of rotations.

Finally, we recall that there is a neighborhood N of the identity and an $\varepsilon > 0$ such that each $g \in N$ can be written uniquely as $g = \exp(tA)$ for some $A \in T_eG$ with $\|A\| = 1$ and with $|t| < \varepsilon$. Here the norm $\|\cdot\|$ is computed by expressing $A = (a_{ij})$ with respect to a fixed basis and taking $\|A\| = (\Sigma a_{ij}^2)^{1/2}$. For any $g \in N$ we define $\|g\| = t$ where $g = e^{tA}$ is the canonical representation of g. If $g \notin N$ we just take $\|g\| = 1$. Referring to the previous examples, we see that $\|A_\theta\| = \sqrt{2}$, so for the rotation $g = \exp(\theta A_\theta)$ we have $\|g\| = |\theta|\sqrt{2}$. Similarly, for $g = \exp(aA_a)$ we get $\|g\| = |a|$. [For the facts used in this paragraph and subsequently, the handiest reference seems to be Auslander and MacKensie (1977), Chapter 7, pages 117–134.] With these conventions it is now possible to state the first results.

THEOREM 1. *Suppose that $X_1, X_2, \ldots$ is a sequence of independent random variables with distribution μ on $\mathbb{R}^d$, which is not invariant under any continuous one dimensional subgroup of rigid motions. If g_i is any sequence of rigid motions converging to the identity, but such that $\Sigma_{i=1}^{\infty}\|g_i\|^2 = \infty$, then the processes $\{X_1, X_2, \ldots\}$ and $\{g_1X_1, g_2X_2, \ldots\}$ are mutually singular.*

Before stating the next result, it is worth observing that the case of radial symmetry in $\mathbb{R}^2$ shows the necessity of ruling out invariant distributions μ. The possibility of discrete symmetry underlies the necessity of restricting attention to g_i converging to e.

To state the second theorem we need the notion of finite Fisher information. To motivate our definition we recall that if $f(\cdot)$ is a smooth density on $\mathbb{R}$, the translation family $f_\theta(x) = f(x - \theta)$ has Fisher information

$$I = E_\theta\left(\frac{\partial}{\partial\theta}\log f_\theta(x)\right)^2 = \int_{-\infty}^{\infty}\{f'(x)^2/f(x)\}\,dx$$
$$= 4\int_{-\infty}^{\infty}\left(\frac{d}{dx}\sqrt{f(x)}\right)^2 dx.$$

This last equality suggests that we generalize the notion that the derivative of $h(x) = \sqrt{f(x)}$ is in L^2. For any continuous one-parameter subgroup $H = \{p(t): t \in \mathbb{R}\}$ we define an operator on $C_0^\infty(\mathbb{R}^d)$ by setting

$$(L\phi)(x) = \frac{d}{dt}\phi(p(t)x)\bigg|_{t=0}.$$

This operator is called the infinitesimal operator associated with the subgroup H. For example, if H is the subgroup of rotations in $\mathbb{R}^2$, then one can easily check that $L = x(\partial/\partial y) - y(\partial/\partial x)$.

The infinitesimal operators can be extended from C_0^∞ in the usual way to the class of distributions (generalized functions) on $\mathbb{R}^d$. In particular, Lh is well-defined for any function h, although Lh may not necessarily be a function. We can now give the main definition.

We say that a density f on $\mathbb{R}^d$ has *finite Fisher information*, provided for $h = \sqrt{f}$ we have $Lh \in L^2(\mathbb{R}^d)$ for all infinitesimal operators associated with continuous one-parameter subgroups of rigid motions.

Here it is useful to recall that a distribution ν is said to be in L^2, provided $\sup \nu(\phi) < \infty$ for all $\phi \in C_0^\infty$ with $\|\phi\|_2 = 1$. We also note that such a ν must be in the dual of L^2 which is just L^2 again, so the statement $Lh \in L^2(\mathbb{R}^d)$ entails the conclusion that Lh is a function and is in $L^2(\mathbb{R}^d)$ in the usual sense.

The next result shows that l^2 Euclidean perturbations of distributions with finite Fisher information are equivalent. This result thus provides large class of examples of a particularly strong type of quasi-invariance [c.f. Feldman (1961)].

THEOREM 2. *If $X_1, X_2, \ldots$ are independent with density f on $\mathbb{R}^d$ with finite Fisher information, then the processes $\{X_1, X_2, \ldots\}$ and $\{g_1X_1, g_2X_2, \ldots\}$ are mutually absolutely continuous whenever $\sum_{k=1}^{\infty}\|g_k\|^2 < \infty$.*

The final result shows that finite Fisher information is much more than a convenience in the detection problem. In fact, the class of distributions with finite Fisher information is precisely the class for which l^2 perturbations never permit certain detection.

THEOREM 3. *If* $\{X_1, X_2, \dots\}$ *is an independent stationary process on* $\mathbb{R}^d$ *and* $\{X_1, X_2, \dots\}$ *and* $\{g_1X_1, g_2X_2, \dots\}$ *are mutually absolutely continuous for all sequences* $\{g_i\}$ *such that* $\sum_{i=1}^{\infty}\|g_i\|^2 < \infty$, *then the* X_is *have a strictly positive density* $f(x)$ *with finite Fisher information.*

The proofs of these theorems are given in the next three sections. The fifth section discusses some related literature and mentions some open problems.

2. Proof of Theorem 1. The main tool used in the proofs which follow is the theorem of Kakutani (1948) on the singularity and equivalence of product measures. If μ and ν are any two probability measures which are absolutely continuous with respect to a measure m, then the Hellinger integral $H(\mu, \nu)$ is defined by

$$H(\mu, \nu) = \int (fg)^{1/2}\, dm,$$

where $f = d\mu/dm$ and $g = d\nu/dm$ are the Radon–Nikodým derivatives. One can check without difficulty that $H(\mu, \nu)$ does not depend upon m and that $0 \le H(\mu, \nu) \le 1$.

Now if μ_k, $k = 1, 2, \dots$ and ν_k, $k = 1, 2, \dots$ are any two sequences of probability measures such that μ_k and ν_k are mutually absolutely continuous for each k, then the theorem of Kakutani states that the infinite product measures

$$\mu = \prod_{k=1}^{\infty} \mu_k \quad \text{and} \quad \nu = \prod_{k=1}^{\infty} \nu_k$$

are either mutually singular or mutually absolutely continuous accordingly as

$$\prod_{k=1}^{\infty} H(\mu_k, \nu_k) = 0$$

or

$$\prod_{k=1}^{\infty} H(\mu_k, \nu_k) > 0.$$

To prove Theorem 1 we first note that since g_i are converging to the identity e, there is no loss in assuming that all the g_i are in the neighborhood N where each $g \in N$ can be written as $g = \exp(tA)$ for a unique $A \in T_eG$ with $\|A\| = 1$ and a unique t, $0 \le t \le \varepsilon < \infty$.

We now take a sequence Z_i which are i.i.d. $N(\mathbf{0}, I)$ and consider the sequences $\{X_i'\} = \{X_i + Z_i\}$ and $\{Y_i'\} = \{g_i(X_i + Z_i)\}$. It is intuitive that if $\{X_i'\}$ and $\{Y_i'\}$ are singular, then so are $\{X_i\}$ and $\{g_iX_i\}$. To establish this rigorously we first note that $\{g_i(X_i + Z_i)\} =_d \{g_iX_i + Z_i\}$ by the affine character of g_i and the spherical symmetry of Z_i. By Kakutani's theorem we see therefore that the singularity of $\{X_i'\}$ and $\{Y_i'\}$ implies the singularity of $\{X_i + Z_i\}$ and $\{g_iX_i + Z_i\}$. Now, we use a Fubini argument. By the singularity of $\{X_k + Z_k\}$ and $\{g_kX_k + Z_k\}$ there is a measurable subset $B \subset \mathbb{R}^\infty$ such that the events $E_1 = \{\{X_k + Z_k\} \in B\}$ and $E_0 = \{\{g_kX_k + Z_k\} \in B\}$ have measures 1 and 0, respectively under the product measure $P \times P'$, where P is the measure on Ω given by the $\{X_k\}$ process and P' is the measure on Ω' given by the $\{Z_k\}$ process. By Fubini's theorem there is a subset $\Omega_0' \subset \Omega'$ of P' measure one such that for all $\omega' \in \Omega_0'$ the events $E_1(\omega') = \{\{X_k + Z_k(\omega')\} \in B\}$ and $E_0(\omega') = \{\{g_k X_k + Z_k(\omega')\} \in B\}$ have P measures 1 and 0, respectively. We choose some fixed $\omega_0' \subset \Omega_0'$ and define a new measurable subset $\tilde{B} \subset \mathbb{R}^\infty$ by $\tilde{B} = B - \{Z_k(\omega_0')\}$. We then have $P(\{X_k\} \in \tilde{B}) = 1$ and $P(\{g_k X_k\} \in \tilde{B}) = 0$ which proves that $\{X_k\}$ and $\{g_k X_k\}$ are singular.

Our objective is now to start computing Hellinger integrals in order to use the assumption $\sum_{k=1}^\infty \|g_k\|^2 = \infty$ to show $\{X_i'\}$ and $\{Y_i'\}$ are singular.

We let $f(\cdot)$ denote the common density of the X_i' and consider the Hellinger integrals

$$H_k = \int \sqrt{f(x) f(g_k^{-1} x)}\, dx = \int h(x) h(e^{-tA} x)\, dx,$$

where $h(x) = \sqrt{f(x)}$ and $g_k = e^{tA}$. Setting $\psi_A(t) = \int h(x) h(e^{-tA} x)\, dx$ we note that $\psi_A(0) = 1$ and that it is not difficult to show that $\psi_A(\cdot)$ is infinitely differentiable.

By the change of variables $y = e^{-tA} x$ one also sees that $\psi_A(\cdot)$ is an even function and that consequently the two-term Taylor series with remainder is just

$$\psi_A(t) = 1 + \int_0^t \psi_A''(u)(t-u)\, du. \tag{2.1}$$

By computing the first derivative and then changing variables before computing the second, one has

$$\psi_A''(t) = -\int \{\nabla h(e^{tA} y) \cdot A e^{tA} y\}\{\nabla h(y) \cdot Ay\}\, dy, \tag{2.2}$$

which simplifies for $t = 0$ to

$$\psi_A''(0) = -\int \{\nabla h(y) A y\}^2\, dy = -\int \{Lh\}^2\, dy.$$

Now, if $\psi_A''(0) = 0$, we will show that h is invariant under the one-parameter group $p(t) = e^{tA}$. To see this, first note that $\psi_A''(0) = 0$ implies $1 - \psi_A(t) = O(t^4)$, since $\psi_a(t)$ is even. But setting $0 = t_0 < t_1 < \cdots < t_n = t$ we have

$$\begin{aligned}
0 \le 1 - \int h(x) h(p(t)x)\, dx &= \sum_{k=1}^n \int h(x)\{h(p(t_{k-1})x) - h(p(t_k)x)\}\, dx \\
&\le \sum_{k=1}^n \left(\int \{h(p(t_{k-1})x) - h(p(t_k)x)\}^2\, dx \right)^{1/2} \\
&\le \sum_{k=1}^n \left(\int \{h(p(t_{k-1} - t_k)x) - h(x)\}^2\, dx \right)^{1/2} \\
&\le \sum_{k=1}^n (2 - 2\psi_A(t_{k-1} - t_k))^{1/2}.
\end{aligned}$$

Now the fact that $1 - \psi_A(t) = O(t^4)$ shows the last term above can be made as small as we like. Thus $\int h(x)h(p(t)x)\,dx = 1$ for all t; and since $H(\nu, \mu) = 1$ only if ν and μ are equal, we see $f(x) = f(p(t)x)$ for almost every x. Our assumption that f is not invariant under a one-parameter subgroup therefore implies that $\psi_A''(0) < 0$ for all $A \in T_eG$, $\|A\| = 1$.

By the continuity of $\psi_A''(t)$ as a function of (t, A) and by the compactness of the set $K = \{(0, A): \|A\| = 1\}$, we obtain an open neighborhood $\mathcal{O}$ containing K and a $\delta > 0$ such that $\psi_A''(t) \leq -\delta$ for all $(t, A) \in \mathcal{O}$.

Applying this bound in the Taylor expansion (2.1) we see

$$\psi_A(t) \leq 1 - \delta t^2/2 \tag{2.3}$$

for all A, $\|A\| = 1$ and all t in an ε neighborhood of 0. Since the g_k are by assumption converging to e, there is no loss in assuming $g_k = e^{t_k A_k}$, $\|A_k\| = 1$, $|t_k| \equiv \|g_k\| \leq \varepsilon$. If μ_k is the measure corresponding to the $\{X_k + Z_k\}$ and ν_k corresponds to $\{g_k(X_k + Z_k)\}$, then

$$\prod_{k=1}^{\infty} H(\mu_k, \nu_k) \leq \prod_{k=1}^{\infty} \{1 - \delta\|g_k\|^2/2\}. \tag{2.4}$$

Since $\sum_{k=1}^{\infty}\|g_k\|^2 = \infty$ by assumption, the right side of (2.4) diverges to zero. By Kakutani's theorem this shows $\{X_i'\}$ and $\{Y_i'\}$ are singular and by the earlier reduction this completes the proof. □

3. Proof of Theorem 2. Let ϕ_ε denote a normal density with mean zero and covariance matrix εI. Under the hypothesis of Theorem 2 it is easy to check that

$$\lim_{\varepsilon \to 0} \int \sqrt{\phi_\varepsilon^* f(x) \cdot f(x)}\, dx = \int f(x)\, dx = 1.$$

We can therefore choose $\varepsilon_k \downarrow 0$ so rapidly that

$$\prod_{k=1}^{\infty} \int \sqrt{\phi_{\varepsilon_k}^* f(x) \cdot f(x)}\, dx > 0. \tag{3.1}$$

If the $Z_k \sim N(\mathbf{0}, \varepsilon_k I)$ $1 \leq k < \infty$ are independent, then (3.1) and Kakutani's theorem will give

$$\{X_k\} \sim \{X_k + Z_k\} \tag{3.2}$$

and

$$\{g_k X_k\} \sim \{g_k(X_k + Z_k)\} \sim \{g_k X_k + Z_k\}. \tag{3.3}$$

Here " $\sim$ " is used to denote measure theoretic equivalence or mutual absolute continuity of the processes. One should note that the second equivalence of (3.3) comes from the fact that $g(X_k + Z_k) =_d gX_k + Z_k$ which is due to the spherical symmetry of $N(\mathbf{0}, I)$.

Now letting $h_\varepsilon(x) = \sqrt{f * \phi_\varepsilon(x)}$; $g = e^{tA}$ and $\psi(t, \varepsilon, A) = \int h_\varepsilon(x) h_\varepsilon(e^{-tA}x)\, dx$, we see, as in (2.1) and (2.2), that

$$\psi(t, \varepsilon, A) = 1 + \int_0^t \psi''(u, \varepsilon, A)(t - u)\, du \tag{3.4}$$

and

$$\psi''(t, \varepsilon, A) = -\int \{\nabla h_\varepsilon(e^{tA}y) \cdot Ae^{tA}y\}\{\nabla h_\varepsilon(y) \cdot Ay\}\, dy. \tag{3.5}$$

Applying the Schwarz inequality to (3.5) and using the invariance of Lebesgue measure we have

$$|\psi''(t, \varepsilon, A)| < \int \{(\nabla h_\varepsilon)(y) \cdot Ay\}^2\, dy. \tag{3.6}$$

To bound this last integral, first re-express it and then use Schwarz's inequality

$$\int \{(\nabla h_\varepsilon)(y) \cdot Ay\}^2\, dy = \frac{1}{4}\int \left(\frac{d}{dt} f * \phi_\varepsilon(e^{tA}x)\right)^2 \Big/ f * \phi_\varepsilon(e^{tA}x)\, dx$$

and

$$\begin{aligned}\left(\frac{d}{dt} f * \phi_\varepsilon(e^{tA}x)\right)^2 &= \left(\int \frac{d}{dt} f(e^{tA}x - y)\phi_\varepsilon(y)\, dy\right)^2 \\ &\le \int \phi_\varepsilon(y)\left(\frac{d}{dt} f(e^{tA}x - y)\right)^2 \Big/ f(e^{tA}x - y)\, dy \\ &\quad \cdot \int f(e^{tA}x - y)\phi_\varepsilon(y)\, dy.\end{aligned}$$

Hence

$$\begin{aligned}\int \{(\nabla h_\varepsilon)(y) \cdot Ay\}^2\, dy &\le \frac{1}{4}\int\int \phi_\varepsilon(y)\left(\frac{d}{dt} f(e^{tA}x - y)\right)^2 \Big/ f(e^{tA}x - y)\, dy\, dx \\ &= \frac{1}{4}\int \left(\frac{d}{dt} f(e^{tA}x)\right)^2 \Big/ f(e^{tA}x)\, dx \\ &= \int \left\{\left(\nabla \sqrt{f(y)}\right) \cdot Ay\right\}^2 dy < \infty.\end{aligned}$$

The finiteness of the last integral naturally is just the application of the hypothesis of finite Fisher information.

Since the last integral is just a quadratic function of A, we have a uniform bound

$$\sup_{\|A\|=1} \int \left\{\left(\nabla \sqrt{f(y)}\right) \cdot Ay\right\}^2 dy = B < \infty,$$

which in (3.4) and (3.6) yields

$$1 - Bt^2/2 \le \psi(t, \varepsilon, A) \le 1 \tag{3.7}$$

for all $0 \le \varepsilon < \infty$, $\|A\| = 1$, and $-\infty < t < -\infty$. From (3.7) it follows immediately that $\prod_{k=1}^{\infty}\psi(t_k, \varepsilon_k, A) > 0$ whenever $\sum_{i=1}^{\infty}\|g_i\|^2 < \infty$. By Kakutani's theorem we then have $\{X_i + Z_i\} \sim \{g_i(X_i + Z_i)\}$, so by the equivalences (3.2) and (3.3) the proof of Theorem 2 is complete. □

4. Proof of Theorem 3. If μ is any measure on $\mathbb{R}^d$, we define the translated measure μ_a, $a \in \mathbb{R}^d$, by $\mu(A) = \mu(A + a)$ for all Borel $A \subset \mathbb{R}^d$.

LEMMA 4.1 [c.f. Shepp (1965), Lemma 5]. *If $\mu \sim \mu_a$ for all $a \in \mathbb{R}^n$, then μ is absolutely continuous with respect to Lebesgue measure and corresponds to a strictly positive density.*

PROOF. Letting λ denote Lebesgue measure we have by Fubini's theorem that

$$\int_{\mathbb{R}^d} \mu(A + a)\, da = \lambda(A). \tag{4.1}$$

Now, if $\mu(A) = 0$, then $\mu_a \sim \mu$ for all a implies $\mu(A + a) = 0$ for all a, which by (4.1) shows $\lambda(A) = 0$. The Radon–Nikodým theorem then shows μ has a density $f(\cdot)$.

Now, to prove Theorem 3 we suppose that $\{X_1, X_2, \dots\}$ is equivalent to $\{g_1X_1, g_2X_2, \dots\}$ for all $(g_i)_{i=1}^{\infty}$ such that $\Sigma_{i=1}^{\infty}\|g_i\|^2 < \infty$. Since we can take g_1 to be an arbitrary translation and then take $g_i = e$ for $i \geq 2$, we see from Lemma 4.1 that the X_i must have a positive density f.

Next, let L be any infinitesimal operator associated with a one-parameter subgroup H rigid motions. We have to show that $Lh \in L^2(\mathbb{R}^d)$ where $h = \sqrt{f}$.

There are two cases to consider: H compact and H noncompact. It is well known that if H is compact, it must be conjugate to the subgroup of rotations in the x_1–x_2 plane. Further, if H is noncompact, it must be conjugate the subgroup of translations along the x_1-axis.

We consider first the harder case of H compact. Writing $H = \{e^{tA}\colon t \in \mathbb{R}\}$ we known there is a rigid motion M so that

$$\exp(tA)M = M\exp(tA_\theta), \tag{4.2}$$

where $\exp(tA_\theta)$ corresponds to rotation in the x_1–x_2 plane.

Now we let $h_0(x) = h(Mx)$ and calculate

$$\begin{aligned} I(t) &= \int h(x)h(\exp(tA)x)\,dx = \int h(Mx)h(\exp(tA)Mx)\,dx \\ &= \int h(Mx)h(M\exp(tA_\theta)x)\,dx = \int h_0(x)h_0(\exp(tA_\theta)x)\,dx. \end{aligned} \tag{4.3}$$

We write $(x_1, x_2, \dots, x_d) = (\rho\cos\theta, \rho\sin\theta, x_3, \dots, x_d)$, and define $h_{00}(\theta, \rho, x_3, \dots, x_d) = h_0(x_1, x_2, \dots, x_d)$ to obtain

$$\begin{aligned} I(t) &= \int h_0(x)h_0(\exp(tA_\theta)x)\,dx \\ &= \int h_{00}(\theta, \rho, x_3, \dots, x_d)h_{00}(\theta + t, \rho, x_3, \dots, x_d) \\ &\qquad \cdot \rho\, d\rho\, d\theta\, dx_3, \dots, dx_d. \end{aligned} \tag{4.4}$$

Setting

$$\hat{h}_{00}(n, \rho, x_3, \dots, x_d) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-in\theta} h_{00}(\theta, \rho, x_3, \dots, x_d)\, d\theta,$$

one obtains by Parseval's identity that

$$I(t) = \int \Big\{ \sum_{\mathbb{Z}} |\hat{h}_{00}(n, \rho, x_3, \dots, x_d)|^2 \cos(nt) \Big\} \rho\, d\rho\, dx_3, \dots, dx_d \tag{4.5}$$

and

$$\begin{aligned} 1 - I(t) &= \int \Big\{ \sum_{\mathbb{Z}} |\hat{h}_{00}(n, \rho, x_3, \dots, x_d)|^2 \cos(nt) \Big\} \rho\, d\rho\, dx_3, \dots, dx_d \\ &\geq \frac{t^2}{4} \int \Big\{ \sum_{|nt| \leq 1} n^2 |\hat{h}_{00}(n, \rho, x_3, \dots, x_d)|^2 \Big\} \rho\, d\rho\, dx_3, \dots, dx_d. \end{aligned} \tag{4.6}$$

Now, we note that $Lh \in L^2$ if and only if

$$\frac{d}{d\theta} h_{00}(\theta, \rho, x_3, \dots, x_d) \in L^2(\rho\, d\rho\, d\theta\, dx_3, \dots, dx_d).$$

Therefore, if $Lh \notin L^2$, the function

$$T(t) = \int \Big\{ \sum_{|nt| \leq 1} n^2 |\hat{h}_{00}(n, \rho, x_3, \dots, x_d)|^2 \Big\} \rho\, d\rho\, dx_3, \dots, dx_d$$

satisfies $T(t) \to \infty$ as $t \to 0$.

We can now use an elementary lemma on real sequences which is from Shepp (1965).

LEMMA 4.2. *If $T(x) \to \infty$ as $x \to 0$, there exists a real sequence a_k with $\Sigma a_k^2 < \infty$, but $\Sigma a_k^2 T(a_k) = \infty$.*

Applied to (4.6) this lemma shows that if $Lh \notin L^2$ there exist $(a_k) \in l^2$ such that $\Pi_{k=1}^{\infty}(1 - I(a_k)) = 0$. By Kakutani's theorem this says that the processes $\{X_1, X_2, \dots\}$ and $\{\exp(a_1 A)X_1, \exp(a_2 A)X_2, \dots\}$ are singular. This contradicts the hypothesis of Theorem 3, and therefore establishes the fact that $Lh \in L^2$ in the case that H is compact. For the noncompact case, one performs a similar reduction to the case of a one-dimensional translation. After that reduction the proof can be completed just as above except that the Fourier transform replaced the Fourier series. □

5. Final remarks. There is an intimate connection between equivalence under l^2 Euclidean perturbations and finite Fisher information. Shepp (1965) posed the question of determining the class of distributions F which are equivalent for all l^p translation perturbations with $p \neq 2$. This problem was settled definitively by Chatterji and Mandrekar (1977). It would be interesting to know if the results of Chatterji and Mandrekar (1977) can be extended to the case of l^p Euclidean perturbations with $p \neq 2$.

Acknowledgment. I would like to thank F. Huffer, M. Shahshahani, and L. A. Shepp for their comments on an earlier draft of this manuscript. This research was supported in part by National Science Foundation Grants MCS 80-24649 and DMS84-14069, and technical assistance from the Department of Statistics, University of Chicago.

REFERENCES

AUSLANDER, L. and MACKENSIE, R. E. (1977). *Introduction to Differentiable Manifolds*. Dover, New York.

CHATTERJI, S. D. and MANDREKAR, V. (1977). Quasi-invariant measures under translation. *Math. Z.* **154** 19–29.

FELDMAN, J. (1961). Examples of non-Gaussian quasi-invariant distributions in a Hibert space. *Trans. Amer. Math. Soc.* **99** 342–349.

KAKUTANI, S. (1948). On equivalence of product measures. *Ann. Math.* **49** 214–224.

RENYI, A. (1967). On some basic problems of statistics from the point of view of information theory, *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 531–543. Univ. California Press.

SHEPP, L. A. (1965). Distinguishing a sequence of random variables from a translate of itself. *Ann. Math. Statist.* **36** 1107–1112.

DEPARTMENT OF STATISTICS
FINE HALL
WASHINGTON ROAD
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08544