*The Annals of Probability*
1992, Vol. 20, No. 3, 1420–1435

# QUASI-INVARIANCE OF PRODUCT MEASURES UNDER LIE GROUP PERTURBATIONS: FISHER INFORMATION AND $L^2$-DIFFERENTIABILITY

BY MAURO S. DE F. MARQUES AND LUIZ SAN MARTIN

*Universidade Estadual de Campinas*

A sequence of measures on a topological space is perturbed by a sequence of elements of a Lie group acting on that space. Criteria are given for the singularity and equivalence of the corresponding product measures. These criteria extend the results of Shepp and Steele. In particular, Fisher information comes into the scene and its role is further clarified.

**Introduction.** This work investigates the mutual Lebesgue decomposition of two product measures, one with a fixed marginal and the other with each marginal perturbed by an element of a Lie group. For general product measures the problem was studied by Kakutani (1948), who proved the following dichotomy: two product measures are either equivalent or singular, and his criterion is as follows. Let $\mu_1, \mu_2, \dots, \tilde{\mu}_1, \tilde{\mu}_2, \dots$ be measures and

$$\mu = \prod_{n=1}^{\infty} \mu_n, \qquad \tilde{\mu} = \prod_{n=1}^{\infty} \tilde{\mu}_n$$

be their product measures. Let $H$ denote the Hellinger product between measures, that is, $H(\mu, \tilde{\mu}) = \int (d\mu/d\nu)^{1/2}(d\tilde{\mu}/d\nu)^{1/2}\, d\nu$, where $\nu$ is a $\sigma$-finite measure with $\mu$ and $\tilde{\mu}$ absolutely continuous with respect to it [see Le Cam (1970)]. Kakutani showed that: (i) $\mu$ and $\tilde{\mu}$ are singular ($\mu \perp \tilde{\mu}$) if and only if $\prod_{n=1}^{\infty} H(\mu_n, \tilde{\mu}_n) = 0$. (ii) If $\mu_n$ and $\tilde{\mu}_n$ are equivalent ($\mu_n \sim \tilde{\mu}_n$) for all $n$, that is, they are mutually absolutely continuous, then $\mu \sim \tilde{\mu}$ if and only if $\prod_{n=1}^{\infty} H(\mu_n, \tilde{\mu}_n) > 0$.

More specific results on equivalence and singularity of product measures were considered by Feldman (1961), Shepp (1965), Rényi (1967), LeCam (1970), Chatterji and Mandrekar (1977), Steele (1986) and Marques (1987). In particular, the problem of translates of product measures in $\mathbb{R}^{\infty}$ was settled by Shepp, who brought the role of Fisher information to the scene. Extensions of Shepp's result to the groups of rigid motions and affine transformations in Euclidean spaces were obtained by Steele and Marques, respectively.

Here we show that Shepp's result holds under a more general scenario. We extend the results to Lie group perturbations acting on certain topological spaces, and the role of Fisher information in the problem is further clarified.

Apart from facts in probability theory, our techniques of proof involve mainly differential calculus of maps into Hilbert spaces and some representa-

Received March 1990; revised February 1991.
*AMS* 1980 *subject classifications*. Primary 60G30; secondary 60B15.
*Key words and phrases*. Quasi-invariance, absolute continuity, singularity, product measures, Lie groups, Fisher information.

tion theory of Lie groups. The differential calculus and its relation to Fisher information is discussed in Section 2. For group representations, we refer to the books of Warner (1972) or Bourbaki (1963). Let us, recall here however, some basic notions of this theory.

A representation of a group $G$ on a vector space $E$ is a homomorphism $g \to U(g)$ of $G$ into the group of invertible linear maps of $E$. In case $E$ is a topological vector space and $G$ a Lie group (or even a topological group), the representation $U$ is said to be continuous in case the map $(g, v) \in G \times E \to U(g)v \in E$ is continuous. In the body of the paper, only unitary representations of Lie groups will appear. These are representations in which $E$ is a Hilbert space and each $U(g)$, $g \in G$, is a unitary operator. For representations of Lie groups on Hilbert (or Banach) spaces, the above continuity condition is equivalent to the weaker one that every $v \in E$ be a continuous vector, that is, the map $g \in G \to U(g)v \in E$ is continuous [cf. Bourbaki (1963), Chapter VIII, Section 1].

If $U_i$ are representations on $E_i$ for $i = 1, 2$ of the same Lie group, a continuous operator $A$: $E_1 \to E_2$ is said to be an intertwining operator for $U_1$ and $U_2$ if $A_0 U_1(g) = U_2(g)_0 A$ for all $g \in G$. The representations $U_1$ and $U_2$ are equivalent in case there exists an intertwining operator which is also a bicontinuous bijection. Of course, $U_1$ and $U_2$ can be regarded as the same representation if they are equivalent.

A vector $v$ in the representation space $E$ (assumed to be a Banach space) of a Lie group $G$ is a $C^\infty$ ($C^k$, continuous, analytic, etc.) vector in case the map $\psi_v$: $g \in G \to \psi_v(g) = U(g)v \in E$ is $C^\infty$ ($C^k$, etc.). The set $E_\infty$ of $C^\infty$-vectors in $E$ is a dense subspace [cf. Warner (1972), Section 4.4.1]. If $v \in E_\infty$ and $X$ is an element of the Lie algebra $\mathbf{g}$ of $G$, it makes sense to define

$$U_\infty(X)v = \lim_{t \to 0} \frac{U(\exp tX)v - v}{t}.$$

The vector $U_\infty(X)$ is also $C^\infty$ and the assignment $v \in E_\infty \to U_\infty(X)v$ defines an operator of $E$ which is in general unbounded if $E$ is infinite-dimensional. Clearly, the derivative of $\psi_v(g) = U(g)v$ at the identity in the direction of $X$ is $U_\infty(X)v$. Also, $U(g)v \in E_\infty$ for all $g \in G$, if $v \in E_\infty$, $(d\psi_v)_g(X(g)) = (d/dt)(U(ge^{tX})v)_t = U(g)U_\infty(X)v$ if $X$ is regarded as a left invariant vector field, and $(d\psi_v)_g(X(g)) = (d/dt)(U(e^{tX}g)v)_{t=0} = U_\infty(X)U(g)v$ in case $X$ is regarded as a right invariant vector field.

Since the image of $U_\infty(X)$, $X \in \mathbf{g}$, is contained in $E_\infty$, it is possible to take compositions and consider higher order operators, that is, linear combinations of operators of the form $U_\infty(X_1) \circ \cdots \circ U_\infty(X_k)$, $X_i \in \mathbf{g}$. These are also densely defined operators in $E$. Later on in Section 5, a second order operator of this kind will appear.

**2. Immersions and Fisher information.** We start with some elementary considerations on maps into Banach spaces. Let $U$ be an open subset of $\mathbb{R}^d$ and $\varphi$: $U \to E$ a continuous injection into the Banach space $(E, \| \ \|)$. Define on $U$ the distance $d_E(x, y) = \|\varphi(x) - \varphi(y)\|$. Assume that $\varphi$ is differentiable at some $x \in U$ and that its differential $d\varphi_x\colon \mathbb{R}^d \to E$ is injective. Then for $v \in \mathbb{R}^d$, $\varphi(x+v) = \varphi(x) + d\varphi_x(v) + o(v)$ with $\lim_{v \to 0}(o(v)/|v|) = 0$, where $|\cdot|$ stands for the Euclidean norm in $\mathbb{R}^d$. Putting $M = \sup\{\|d\varphi_x(v)\|\colon |v| = 1\}$ and $m = \inf\{\|d\varphi_x(v)\|\colon |v| = 1\}$, we have $m > 0$ and for small $v \in \mathbb{R}^d$ and $\varepsilon > 0$, $(m - \varepsilon)|v| \le d_E(x, x+v) \le (M + \varepsilon)|v|$. Thus if $\varphi$ is differentiable, $d_E$ is locally equivalent to the Euclidean distance $d(x, y) = |x - y|$ restricted to $U$. So $\sum_{n \ge 1} d(x, y_n)^2 < \infty$ if and only if $\sum_{n \ge 1} d_E(x, y_n)^2 < \infty$ for $(y_n)_{n \ge 1}$ a sequence in $U$ with $y_n \to x$.

Shifting to manifolds, suppose that $M$ is a finite-dimensional smooth manifold endowed with some smooth Riemannian metric $\langle\ ,\ \rangle$. Let $\varphi = M \to E$ be a continuous injection. Through the arc length of curves joining points in $M$, a Riemannian metric defines a distance $d(x,y)$ in $M$. Localizing around $x \in M$ and taking some coordinate system, the existence of the exponential map guarantees the equivalence of this Riemannian distance with the Euclidean distance of the open subset of $\mathbb{R}^d$, where the coordinate system is defined. This leads us to the previous situation so that $d(x,y)$ becomes locally equivalent to $d_E(x,y) = \|\varphi(x) - \varphi(y)\|$. Again we have that for sequences $y_n \to x$, $\sum_{n \ge 1} d(x, y_n)^2 < \infty$ if and only if $\sum_{n \ge 1} d_E(x, y_n)^2 < \infty$, in case $\varphi$ is differentiable with $d\varphi_x$ one-to-one. This statement can be slightly improved in case $\varphi$ is a one-to-one immersion of $M$ into $E$. In fact, by a simple compactness argument one sees that the distances $d(\cdot, \cdot)$ and $d_E(\cdot, \cdot)$ are equivalent on compact subsets of $M$. Therefore, if $\varphi$ is an immersion, $\sum_{n \ge 1} d(x, y_n)^2 < \infty$ if and only if $\sum_{n \ge 1} d_E(x, y_n)^2 < \infty$, provided $y_n$ does not leave a compact.

Now, let $\{\mu_x\}_{x \in M}$ be a dominated model parameterized on $M$. Each $\mu_x$, $x \in M$, is a probability measure on the measurable space $(\Omega, \mathscr{F})$ and $\mu_x$ is absolutely continuous with respect to a basic measure $\mu$. Write $p(x, \omega) = (d\mu_x/d\mu)(\omega)$ for the set of Radon–Nikodim derivatives. For $x \in M$, $p(x, \cdot)$ is $\mu$-integrable so $x \to p(x, \cdot)$ defines a map $M \to L^1(\mu)$. Also, since $p(x, \cdot)$ is $\mu$-a.s. positive, $q(x, \omega) = p(x, \omega)^{1/2}$ makes sense and $x \to \varphi(x) = q(x, \cdot)$ defines a map from $M$ into the Hilbert space $L^2(\mu)$. We say that the model is $L^2$-continuous, $L^2$-differentiable and so on in case the specified property is satisfied by $\varphi$.

Suppose the model is $L^2$-differentiable at $x \in M$ and define the Fisher inner product $\langle\ ,\ \rangle_x$ on the tangent space $T_xM$ at $x$ by putting $\langle u, v \rangle_x = \langle d\varphi_x(u), d\varphi_x(v) \rangle$ with $u, v \in T_xM$ and $\langle\ ,\ \rangle$ on the right-hand side standing for the inner product in $L^2(\mu)$. Clearly, $\langle\ ,\ \rangle_x$ is positive semidefinite and it is positive definite if and only if $d\varphi_x$ is injective. We call $\langle\ ,\ \rangle_x$ the Fisher inner product because when it is expressed in coordinate systems, its matrix is nothing else than the Fisher information matrix in Ibragimov and Has'minskii (1981). Note that in the setting adopted here, the very definition of Fisher information requires differentiability of $\varphi$. However the usual way of dealing with Fisher information is by assuming the differentiability of $q(x, \omega)$ as a function of $x$ and putting

$$\langle u, v \rangle_x = \int D_u q(x, \omega) D_v q(x, \omega) \mu(d\omega),$$

provided the integral exists [here $D_v q(x, \omega)$ means directional derivative]. In this case, under some regularity conditions, usually called Cramér–Wald and Hajek's conditions, the existence of $\langle u, v\rangle_x$, $u, v \in T_x M$, implies the existence of $d\varphi_x$ [see Le Cam (1970)]. So in order to fix terminology, we understand that the existence of Fisher information at $x$ means the same as the existence of $d\varphi_x$.

Finally, set $d_H(x, y) = \|\varphi(x) - \varphi(y)\|_2$, where $\|\cdot\|_2$ denotes the $L_2(\mu)$-norm. Then $d_H(\cdot, \cdot)$ is the so-called Hellinger distance between $\mu_x$ and $\mu_y$, which can also be expressed as $d_H^2(x, y) = 2(1 - H(\mu_x, \mu_y))$. So taking into account Kakutani's theorem and the previous remarks, we have:

PROPOSITION 2.1. *In a dominated model $\{\mu_x\}_{x \in M}$, where $M$ is a smooth manifold with a distance given by any smooth Riemannian metric, suppose that $\varphi$: $M \to L^2(\mu)$ (notation as above) is differentiable at $x \in M$ and $d\varphi_x$ is one-to-one. Then for sequences $y_n \to x$,*

$$\prod_n \mu_{y_n} \sim \prod_n \mu_x$$

*if and only if $\sum_{n \geq 1} d(x, y_n)^2 < \infty$. In case $\varphi$ is an injective immersion, the result holds under the assumption that $y_n$ is confined in some compact subset.*

REMARK 1. In case $\varphi$ is a one-to-one smooth immersion, the Fisher information defines a smooth Riemannian metric. In this case, Fisher information itself can be used to measure the intrinsic distance in $M$.

REMARK 2. Smoothness of $M$ and $\langle\ ,\ \rangle$ is not essential; $C^2$ would do.

**3. Lie group perturbation of measures.** From now on, we consider only dominated models obtained by the action of Lie groups on quasi-invariant probability measures. Let $G$ be a connected finite-dimensional Lie group and $X$ a topological space which is assumed to be locally compact and to satisfy the second axiom of enumerability. Let $(g, x) \to gx$ or $g(x)$ denote a continuous action $G \times X \to X$ of $G$ on $X$ and lift it to the space $M(X)$ of Borel measures on $X$ by putting $(g\mu)(A) = \mu(g^{-1}A)$ for any Borel subset $A \subset X$ and $\mu \in M(X)$. The mapping $(g, \mu) \to g\mu$ defines in fact an action of $G$ on $M(X)$, which is moreover continuous when $M(X)$ is provided with the vague topology.

We recall that a $G$ quasi-invariant measure is a measure $\mu \in M(X)$ for which $g\mu$ is equivalent to $\mu$ for every $g \in G$. Thus a dominated model $\{g\mu\}_{g \in G}$ is defined if $\mu$ is a quasi-invariant probability. In this model a dominant measure may be taken to be $\mu$ itself. This is the combined transformational model in Barndorff-Nielsen (1987). In the sequel we shall put

$$(3.1)\quad p(g, x) = \frac{dg\,\mu}{d\mu}(x) \quad\text{and}\quad q(g, x) = p(g, x)^{1/2}; \qquad g \in G, x \in X,$$

from which the map $\psi$: $g \in G \to q(g, \cdot) \in L^2(\mu)$ is defined. The parameter manifold of this model is $G$. When defined this way, the model need not be one-to-one. So in order to avoid duplication of parameters it is convenient to replace $G$ by its quotient with a subgroup. Set

$$H = \{g \in G: g\mu = \mu\}.$$

Since $(g, \mu) \to g\mu$ is a continuous action, $h$ is a closed subgroup of $G$. It is thus a Lie subgroup of $G$ so that the coset space $G/H$ has a structure of an analytic manifold on which $G$ acts analytically. Since for $h \in H$, $(gh)\mu = g\mu$, we have that $q(gh, x) = q(g, x)$ $\mu$ a.s. hence the map $\psi$: $G \to L^2(\mu)$ defines a one-to-one map $\varphi$: $gH \in G/H \to \psi(g) \in L^2(\mu)$. We take $G/H$ as our parameter manifold. For the mapping $\varphi$, assumptions like that appearing in Proposition 2.1 are acceptable.

In $G/H$, take an arbitrary smooth Riemannian metric $\langle\ ,\ \rangle$ with associated Riemannian distance $d(\cdot, \cdot)$. By definition, the set of sequences $(g_n)_{n \geq 1} \subset G$ that are square summable in $G/H$ is

$$l_H^2 = \left\{(g_n)_{n \geq 1} \subset G: \sum_{n \geq 1} d(g_n H, H)^2 < \infty\right\}. \tag{3.2}$$

We put

$$E(\mu) = \left\{(g_n)_{n \geq 1} \subset G: \prod_n g_n\mu \sim \prod_n \mu\right\}$$

and

$$E_o(\mu) = \{(g_n)_{n \geq 1} \in E(\mu); d(g_n H, H) \to 0\}.$$

It is easily checked that $E(\mu)$ is a subgroup of the group $G^{\mathbb{N}}$ of sequences in $G$.

Now the main result can be stated.

THEOREM 3.1. *With $G$ and $X$ as above, let $\mu$ be a quasi-invariant probability in $M(X)$ and $H$ the closed subgroup which fixes $\mu$. Then:*

(a) $E_o(\mu) \subset l_H^2$, *if $\varphi$ is continuous,*
(b) $E_o(\mu) = l_H^2$, *if $\varphi$ is differentiable at the origin $H \in G/H$,*
(c) *$\varphi$ is continuous and differentiable on $G/H$, if $E_o(\mu) = l_H^2$.*

REMARK. In (c) it is not necessary to assume in advance that $\mu$ is quasi-invariant. In fact, $g\mu \sim \mu$ for every $g \in G$, if $E_o(\mu) \supset l_H^2$, that is, $\mu$ is quasi-invariant.

The proof of this theorem follows in this and later sections. Part (b) is essentially Section 2 above. It involves only the notion of differentiability, which will be clarified soon. Parts (a) and (c) are more delicate and form the contents of later sections. By now we introduce a representation of $G$ associated with $q(g, x)$, from which we derive our main techniques.

An alternative way of defining $g\mu$ is by requiring

$$\int_X f(x) g\mu(dx) = \int_X f(gx)\mu(dx)$$

for integrable $f$. From this equality one gets quickly that $g\mu_1 \sim g\mu_2$ if $\mu_1 \sim \mu_2$ and that

$$\frac{dg\,\mu_1}{dg\,\mu_2}(x) = \frac{d\mu_1}{d\mu_2}(g^{-1}x), \qquad \mu_2\text{-a.s.}$$

Hence, if we put as before $p(g,x) = (dg\,\mu/d\mu)(x)$, then $p$ satisfies (for every $g, h \in G$ and $\mu$-a.a. $x$)

$$p(gh,x) = p(h, g^{-1}x)p(g,x), \tag{3.3}$$

that is, $p$ is a cocycle on $G$ over $X$. If $H$ is the subgroup that fixes $\mu$, then $p$ is $H$-invariant in the sense that $p(h,x) = 1$ if $h \in H$ (this and other equalities appearing below are to be taken $\mu$-a.s.). Moreover $p(h,x) = 1$ if $h \in H$. Note that $H$-invariance and (3.3) imply that $p(gh,x) = p(g,x)$, if $h \in H$ as was already remarked before.

Clearly by putting $q(g,x) = p(g,x)^{1/2}$, $q$ also becomes a cocycle over $X$. This fact permits the introduction of the following representation of $G$: For $g \in G$ and $f \in L^2(\mu)$, define the function

$$(U(g)f)(x) = q(g,x)f(g^{-1}x). \tag{3.4}$$

From

$$\|U(g)f\|_2 = \int_X q(g,x)^2 |f(g^{-1}x)|^2 \mu(dx) = \int_X |f(g^{-1}x)|^2 g\mu(dx) = \|f\|_2,$$

we see that $f \to U(g)f$ defines a unitary operator in $L^2(\mu)$. Moreover, the cocycle condition (3.3) for $q$ leads to $U(gh) = U(g)U(h)$, $g, h \in G$, so that the mapping $g \to U(g)$ becomes a unitary representation of $G$ on $L^2(\mu)$.

Note that $\psi(g) = U(g)1$, where 1 denotes the constant function $1(x) = 1$, so the family $q(g,x)$ is nothing else than the orbit of 1 under the action of $G$ on $L^2(\mu)$ defined by the representation $U$.

Combining all this we can complete the proof of (b) in Theorem 3.1. Suppose that $\varphi$ is differentiable at the origin of $G/H$. By Section 2, it is enough to show that its differential is injective. In terms of the representation $U$, differentiability of $\varphi$ at the origin means differentiability of the vector $1 \in L^2(\mu)$, that is, differentiability at the identity of $G$ of the mapping $\psi(g) = U(g)1$. As is known, this is enough to assure that $\psi$ is everywhere differentiable. In fact, $d\psi_g = U(g)_0\, d\psi_1$. Let $A \in \mathbf{g}$ be such that $d\psi_1(1) = 0$, that is, such that $(d/dt)U(e^{tA})(1)|_{t=0} = 0$. Then

$$(d/dt)U(e^{tA})(1) = U(e^{tA})\big((d/ds)U(e^{sA})(1)|_{s=0}\big) = 0.$$

Hence $U(e^{tA})(1) = 1$ for all $t \in \mathbb{R}$ and therefore $A \in \mathbf{h}$, the Lie algebra of $H$. This suffices to show that the differential of $\varphi$ at the origin of $G/H$ is injective, which in turn completes the proof of (b).

We now present some cases in which the continuity of $\varphi$ or $\psi$ can be taken for granted, thus clarifying the condition in (a). First note that, since $\psi(g) = U(g)1$, $\psi$ and hence $\varphi$ is continuous if $U$ is a continuous representation. Actually, continuity of $\psi$ is equivalent to continuity of $U$, as is shown by:

PROPOSITION 3.2. *$\psi$: $G \to L^2(\mu)$ is continuous if and only if the representation $U$ is continuous.*

PROOF. Suppose $\psi$ is continuous and let $f \in L^2(\mu)$. To show the continuity of $U$ it suffices to show that $g \to U(g)f$ is continuous when $f$ is assumed to be continuous with compact support [cf. Bourbaki (1963), Chapter VIII, Section 2]. In turn it suffices to show the continuity at the identity of $G$. In this case

$$\begin{aligned}\|U(g)f - f\|_2 &= \|q(g,x)f(g^{-1}x) - f(x)\|_2\\ &\le \|(q(g,x) - 1)f(x)\|_2 + \|q(g,x)\big(f(g^{-1}x) - f(x)\big)\|_2,\end{aligned}$$

which converges to zero as $g \to 1$. □

PROPOSITION 3.3. *If either:*

(a) *$q(g,x)$ is jointly continuous in $(g,x)$, or*
(b) *$q(g,x)$ is bounded,*

*then $\psi$ is continuous.*

PROOF. (a) is Proposition 8 and (b) is Proposition 9 in Bourbaki [(1963), Chapter VIII, Section 2]. □

Here is another case of guaranteed continuity of $U$.

PROPOSITION 3.4. *If $G$ acts transitively on $X$, that is, $X = G/L$ for some closed subgroup $L$, then the representation $U$ is continuous.*

PROOF. It is known [cf. Bourbaki (1963), Chapter VII, Section 2] that two quasi-invariant measures on homogeneous spaces are equivalent. It is also known [see, for instance, Bruhat's Lemma A.1.1 in Warner (1972)] that in $G/L$, a quasi-invariant measure $\nu$ with $C^\infty$ cocycle $(dg\,\nu/d\nu)(x) = s(g,x)$, $g \in G$, $x \in G/L$ exists. Put $\rho(x) = d\mu/d\nu$. Then $\rho > 0$, $\nu$-a.s. and

$$\begin{aligned}p(g,x) &= \frac{dg\,\mu}{d\mu}(x) = \frac{dg(\rho\nu)}{d\rho\,\nu}(x) = \frac{\rho(g^{-1}x)}{\rho(x)}\frac{dg\,\nu}{d\nu}(x)\\ &= \frac{\rho(g^{-1}x)}{\rho(x)}s(g,x).\end{aligned}\tag{3.5}$$

Let $A$: $L^2(\mu) \to L^2(\nu)$ be the operator defined by $A(f) = \rho^{1/2}f$. It is easy to check that $A$ is an isometry. Moreover,

$$\begin{aligned}(A \circ U(g) f)(x) &= \rho^{1/2}(x) q(g, x) f(g^{-1}x)\\ &= s(g, x)^{1/2}(\rho^{1/2} f)(g^{-1}x)\\ &= \big(\tilde{U}(g) \circ A(f)\big)(x),\end{aligned}$$

where $(\tilde{U}f)(x) = s(g,x)^{1/2} f(g^{-1}x)$ is the representation associated with the cocycle $s(g,x)^{1/2}$. Therefore $A$ intertwines the representation $U$ and $\tilde{U}$. Hence $U$ is continuous if and only if $\tilde{U}$ is. The continuity of $\tilde{U}$ follows from Proposition 3.3(a) above. □

Finally, let us make some comments about the Fisher information of the model. As observed at the end of the last section, the metric that measures intrinsic distances on the parameters manifold can be taken to be Fisher information itself in case $\varphi$ is smooth. For the model $\{g\mu\}_{g \in G}$ considered before, the Fisher information metric on $G/H$, necessarily nondegenerate, may become very convenient. One reason is the formula $d\psi_g = U(g) \circ d\psi_1$, which shows that Fisher information is $G$-invariant, that is, $\langle dg_\xi(v), dg_\xi(w)\rangle_{g\xi} = \langle v, w\rangle_\xi$, where $\langle\ ,\ \rangle_\xi$ is Fisher information at $\xi \in G/H$, $v, w \in T_\xi(G/H)$ and $dg$ is the differential of the mapping $g$: $G/H \to G/H$ induced by $g \in G$.

It is worthwhile to compare this approach with the notion of Fisher information in the form exploited by Steele (1986). There Lebesgue measure on $\mathbb{R}^d$ is taken as the dominant measure for the models. These are parameterized by the group of rigid motions on $\mathbb{R}^d$. Their natural representation is by the group $G$ of rigid motions on the $L^2$-space of Lebesgue measure on $\mathbb{R}^d$, say $L^2(\mathbb{R}^d)$, determined by $(U(g)f)(x) = f(gx)$, $x \in \mathbb{R}^d$, $g \in G$, $f \in L^2(\mathbb{R}^d)$. In Steele (1986), a density $f$ is said to have finite Fisher information if $Lf \in L^2(\mathbb{R}^d)$ for certain operators defined from this representation. In our earlier notation, these operators are extensions of $U_\infty(X)$, $X \in \mathbf{g}$, from $E_\infty$ to a larger space.

Now it is known [cf. Ibragimov and Has'minskii (1981), page 65] that the existence of Fisher information does not depend on the dominant measure. We take $\mu$ itself as the dominant measure. This being so, note that if $v \in T_\xi(G/H)$, then $v = X(\xi)$ for some $X \in \mathbf{g}$ so that $d\varphi_\xi(x)$ is given by evaluation of $U_\infty(X)$ (or rather an extension of it) on $\varphi(\xi)$. Therefore, existence of Fisher information is equivalent to $U_\infty(X)(\varphi(\xi)) \in L^2(\mu)$ as in Steele (1986).

**4. The $l$-condition and the proof of (c).** In order to prove (c) we make use of the $l$-condition introduced in Le Cam (1970), which for families of the from $\{g\mu\}_{g \in G}$ turns out to be equivalent to differentiability. We take the setup of Theorem 3.1 and denote by $\mathbf{g}$ the Lie algebra of the group $G$. In $\mathbf{g}$ let $|\cdot|$ be any norm. In this context, the $l$-condition reads

$$(4.1)\qquad (l)\colon \limsup_{A \to 0} \frac{\|\psi(ge^{tA}) - \psi(g)\|_2}{|A|} < \infty; \qquad g \in G,\ A \in \mathbf{g}.$$

In principle this condition should be checked at every $g \in G$. However, by the unitarity of the representation $U$, it is fulfilled everywhere if it is satisfied at some $g \in G$. In fact, $\psi(g) = U(g)(1)$ and $\psi(ge^A) = U(g)U(e^A)(1)$ so $\|\psi(ge^A) - \psi(g)\|_2 = \|U(g)(\psi(e^A) - 1)\|_2 = \|\psi(e^A) - 1\|_2$ because $U$ is unitary. Hence the $l$-condition at some $g$ is equivalent to the $l$-condition at the identity of $G$.

The connection between differentiability and the $l$-condition is provided by a result of Le Cam (1970), which says that $\psi$ is almost always differentiable on a measurable set $S$ with respect to the Lesbegue measure of the parameter space if the $l$-condition is satisfied on $S$. From this result we see that if our model satisfies the $l$-condition at just one point of $G$, then it is differentiable everywhere. In fact, as already observed in the proof of (b), for differentiability of $\psi$ everywhere it suffices that $\psi$ be differentiable at some point. Conversely, it is clear that the $l$-condition is satisfied at $g \in G$ if $\psi$ is differentiable at $g$. Thus we get the equivalence of the $l$-condition and differentiability for models of the type $\{g\mu\}_{g \in G}$.

This being so, we can prove (c) by showing that $\psi$ satisfies the $l$-condition in case $\prod_n g_n\mu \sim \prod_n \mu$ for every sequence $(g_n)_{n \geq 1} \in l_H^2$. Before proceeding, let us note that by the local equivalence of Riemannian distances, $e^{An} \in l_H^2$ if $(A_n)_{n \geq 1}$ is a sequence in the Lie algebra $\mathbf{g}$ with $\sum_{n \geq 1} |A_n|^2 < \infty$. For the proof of (c) it is possible to handle only sequences of the form $e^{A_n}$.

We now need a lemma proved in Shepp (1965) as Lemma 4.

LEMMA 4.1. *Let $T: \mathbb{R}^d \to \mathbb{R}$ be a positive function satisfying*

$$\limsup_{a \to 0} T(a) = \infty.$$

*Then there exists a sequence $(a_n)_{n \geq 1} \subset \mathbb{R}^d$ such that $\sum_{n \geq 1} |a_n|^2 < \infty$ and $\sum_{n \geq 1} |a_n|^2 T(a_n) = \infty$.*

To prove (c) in Theorem 3.1, suppose that the $l$-condition for $\psi$ is not satisfied (at the identity of $G$) and take $T: \mathbf{g} \to \mathbb{R}$ in the lemma above to be

$$T(A) = \frac{\|\psi(e^A) - 1\|_2^2}{|A|^2}.$$

Since the $l$-condition is not satisfied, there exists a sequence $(A_n)_{n \geq 1} \subset \mathbf{g}$ such that $\sum_{n \geq 1} |A_n|^2 < \infty$ and $\sum_{n \geq 1} |A_n|^2 T(A_n) = \sum_{n \geq 1} \|\psi(e^{A_n}) - 1\|_2^2 = \sum_{n \geq 1} d_H^2(e^{A_n}\mu, \mu) = 2\sum_{n \geq 1}(1 - H(e^{A_n}\mu, \mu)) = \infty$, which is equivalent to $\prod_{n \geq 1} H(e^{A_n}\mu, \mu) = 0$. By Kakutani's theorem, this implies that $\prod_n e^{A_n}\mu$ and $\prod_n \mu$ are singular, which contradicts the hypothesis.

**5. Proof of (a).** We now turn to the proof of (a). By Kakutani's theorem, we need to show that for sequences $g_n \in G$ with $g_n H \to H$, $\sum_{n \geq 1} \|\psi(g_n) - 1\|_2^2 = \sum_{n \geq 1} \|U(g_n)(1) - 1\|_2^2 = \infty$ if $\sum_{n \geq 1} d(g_n H, H)^2 = \infty$. We know from (b) that this holds in the differentiable case, so we try to regularize $\psi$ (or $\varphi$) by a linear operator $T: L^2(\mu) \to L^2(\mu)$ such that: (i) $T$ is bounded; (ii) the mapping

$g \to T(U(g)(1))$ is smooth ($C^\infty$); (iii) $T(U(g)(1)) = T(1)$ iff $U(g)1 = 1$, that is, iff $g \in H$; or (iii′) $T$ is one-to-one.

Suppose such a $T$ exists. By (ii) and (iii), the mapping $\theta$: $gH \in G/H \to T(\psi(g))$ is one-to-one and smooth. Therefore by a known fact about differentiable mappings, there exists an open subset $W \subset G/H$ such that when restricted to $W$, $\theta$ becomes an immersion. Fix $gH \in W$ and take a sequence $g_n$ with $g_n H \to H$ and $\sum_{n\geq 1} d(g_n H, H)^2 = \infty$. Then $gg_n H \to gH$ and by the local equivalence of Riemannian distances we have $\sum_{n \geq 1} d(gg_n H, gH)^2 = \infty$. Hence, from the differentiable case of Section 2 we have that

$$\sum_{n\geq 1} \|TU(gg_n)(1) - TU(g)(1)\|_2^2 = \infty.$$

Now by the continuity of $T$ as required in (i),

$$\|TU(gg_n)(1) - TU(g)(1)\|_2 \leq \|T\|\,\|U(gg_n)(1) - U(g)(1)\|_2,$$

so that $\sum_{n\geq 1}\|U(gg_n)(1) - U(g)(1)\|_2^2 = \infty$, which by the unitarity of the representation is equivalent to $\sum_{n\geq 1}\|U(g_n)(1) - 1\|_2^2 = \infty$. This proves (a).

One way of getting an operator $T$ satisfying (i), (ii) and (iii) is by representing Brownian motion on $G$ as convolution operators in $L^2(\mu)$. Actually the whole construction to follow is not specific for the representation $U$ on $L^2(\mu)$ but works for any continuous unitary representation of $G$ on a Hilbert space. Here continuity of the representation is essential for the theory to be applicable. This is the reason why continuity of $\varphi$ is required.

For the rest of this section, $U$ denotes a continuous unitary representation of $G$ on the Hilbert space $(E, \langle\ ,\ \rangle)$.

Recall that the representation $U$ lifts to the algebra (under convolution) of probability measures on the Borel sets of $G$ by putting, for a probability $\nu$ and $v \in E$,

$$U(\nu)v = \int_G U(g)v\nu(dg). \tag{5.1}$$

The integral here is the Bochner integral. Its existence is guaranteed by the fact that $U$ is unitary so $\|U(g)v\| = \|v\|$. From this we also have that $\|U(\nu)v\| \leq \|v\|$, so $U(\nu)$ is bounded with operator norm $\|U(\nu)\| \leq 1$. Thus if we take $T = U(\nu)$ for some $\nu$, then condition (i) is automatically satisfied.

Smoothness [condition (ii)] can also be obtained by operators of the form $U(\nu)$. In fact:

LEMMA 5.1. *Let $\pi$ be a left-invariant Haar measure on G and suppose that $\nu$ is a probability with smooth density $f(g) = (d\nu/d\pi)(g)$. Then for all $v \in E$, the mapping $g \in G \to U(\nu)\circ U(g)v$ is $C^\infty$.*

PROOF. Define $\tilde{v}$: $G \to E$ by $\tilde{v}(g) = U(g)v$. Since $\|\tilde{v}(g)\| \leq \|v\|$, $\tilde{v}$ is bounded so it is convolvable with any probability in $G$. Let $\check{\nu}$ be the image of $\nu$ under the mapping $g \in G \to g^{-1} \in G$. Then $\check{\nu}$ also has a smooth density and

$$\begin{aligned}\check{\nu} * \tilde{v}(g) &= \int_G \tilde{v}(h^{-1}g)\check{\nu}(dh)\\ &= \int_G U(h^{-1})U(g)v\check{\nu}(dh)\\ &= U(\nu)\circ U(g)v.\end{aligned}$$

Since convolutions with smooth functions are smooth, we conclude that $g \to U(\nu)\circ U(g)v$ is smooth [see Warner (1972) for convolutions of vector-valued functions]. □

In order to obtain a $\nu$ with smooth density and such that $U(\nu)$ is an injective operator on the representation space $E$, we consider Brownian motion on $G$.

Let $X_1, \dots, X_d$ be a basis for the Lie algebra $\mathbf{g}$ of $G$. Regarding the elements of $\mathbf{g}$ as right-invariant vector fields on $G$, consider the (Stratonovich) stochastic differential equation

$$dg_t = X_1(g)\circ dW_t^1 + \cdots + X_d(g)\circ dW_t^d, \tag{5.2}$$

where $W^1, \dots, W^d$ are independent Brownian motions.

A solution of (5.2) with initial condition $g_0 = 1$ is called a Brownian motion on $G$. It is known [cf. Kunita (1984), Theorem 5.1] that such a solution exists and is defined on the whole ray $[0,\infty)$. Take $t > 0$, denote by $P$ the $d$-dimensional Wiener measure and let $\nu_t = P\circ g_t^{-1}$ be the probability law of the $G$-valued random variable $g_t$ [$\nu_t$ becomes the transition probability of the Markov process associated to solutions of (5.2)]. The infinitesimal generator of (5.2) is the second order differential operator $L = \frac{1}{2}\sum_{i=1}^d X_i^2$. Since this is an elliptic operator, each $\nu_t$ has a smooth density with respect to Lebesgue measure in $G$. This follows, for example, from Malliavin calculus [see, for instance, Watanabe (1984)]. Therefore $\nu_t$ has also a smooth density with respect to any leftinvariant Haar measure $\pi$, so we are in the situation of the above lemma.

We now are going to prove that $U(\nu_t)$ for $t > 0$ is a one-to-one operator in the Hilbert space $E$.

Due to the right-invariance of the vector fields that occur as the coefficients in (5.2), it follows that a solution with initial condition $g$ is given by $g_t g$ with $g_t$ a solution starting at the identity. From this and the Markov property, it follows that if $\nu_t$ is as above, then $\nu_{t+s} = \nu_t * \nu_s = \nu_s * \nu_t$, $t, s \ge 0$.

Applying the representation $U$ to $\nu_t$ we obtain a semigroup $t \to U(\nu_t)$ of bounded operators in $E$, in fact a contraction semigroup since $\|U(\nu_t)\| \le 1$. The infinitesimal generator of $U(\nu_t)$ is of course $U(L)$. Specifically, let $v \in E$ be a $C^\infty$-vector for $U$. Then $g \to U(g)v$ is $C^\infty$ so Itô's formula applies. It gives

$$\begin{aligned}U(g_t)v &= v + \tfrac{1}{2}\sum_{j=1}^d \int_o^t U_\infty\big(X_j^2\big)(U(g_s)v)\,ds\\ &= v + \int_o^t U_\infty(L)(U(g_s)v)\,ds,\end{aligned} \tag{5.3}$$

where $g_t$ is the solution starting at the identity and $U_\infty$ stands for the representation of the right-invariant differential operators of $G$ on the $C^\infty$ vectors of $E$. From the definition of $\nu_t$, it follows that $U(\nu_t)v = E[U(g_t)v]$ for $v \in E$, where the expectation is taken with respect to the Wiener measure $P$. Therefore, (5.3) gives

$$U(\nu_t)v = v + \int_o^t U_\infty(L)(U(\nu_s)v)\, ds,$$

from which it is readily seen that the infinitesimal generator of $U(\nu_t)$ restricted to the $C^\infty$-vectors is exactly $U_\infty(L)$. Hence this is a closable operator and its closure is the infinitesimal generator of the semigroup $U(\nu_t)$.

Now we use a result by Nelson and Stinespring [see Warner (1972), Theorem 4.4.4.3] which assures that for elliptic operators of the form $L$, the closure of $U_\infty(L)$ is self-adjoint. This result and the following lemma on self-adjoint semigroups guarantees that $U(\nu_t)$ is one-to-one.

LEMMA 5.2. *Let $(T_t, t \geq 0)$ be a contraction semigroup on a Hilbert space $E$ with self-adjoint infinitesimal generator $A$. Then $\langle T_t v, v\rangle \geq 0$ for all $v \in E$, $t > 0$, with equality only if $v = 0$. In particular $T_t$ is injective.*

PROOF. By the spectral theorem, we may suppose that $E$ is the $L^2$-space of some measure space $(\Omega, m)$ and that there exists a measurable $h\colon \Omega \to \mathbb{R}$ such that

$$\mathrm{dom}(A) = \left\{ f\colon \int_\Omega \left(1 + h(x)^2\right)|f(x)|^2 m(dx) < \infty \right\}$$

and that

$$(Af)(x) = h(x) f(x), \qquad f \in \mathrm{dom}(A)$$

[see Davies (1980)]. Since $T_t$ is contracting, $A$ is negative semidefinite and hence $h(x) \leq 0$, $m$-a.s. This is enough to assure that

$$(S_t f)(x) = e^{th(x)} f(x)$$

defines a contraction semigroup $(S_t, t \geq 0)$ with $A$ as its infinitesimal generator. By the uniqueness of semigroups given their generators, $T_t = S_t$. For $f \in L^2(m)$, we have

$$\langle T_t f, f\rangle = \int_\Omega e^{th(x)} (f(x))^2 m(dx) \geq 0$$

with equality iff $f(x) = 0$ a.s. The lemma is thus proved. □

By this lemma we conclude the constructability of an operator satisfying (i), (ii) and (iii') and thus the proof of (a) in Theorem 3.1.

**6. Unbounded sequences.** Theorem 3.1 covers only bounded sequences in $G$. It is not difficult, however, to give examples of quasi-invariant probabilities for which there are plenty of sequences $g_n \in G$ that are unbounded (in a sense to be made precise below) and such that $\prod_n g_n \mu \sim \prod_n \mu$ (see Example 7.3).

Here we present a result which assures in many situations that every unbounded sequence $g_n$ separates $\prod_n g_n \mu$ and $\prod_n \mu$. In such situations the condition in Theorem 3.1 that $g_n H \to H$ can be removed. The result we give here is inspired by the notion of recurrence, more specifically by Poincaré's recurrence theorem.

THEOREM 6.1. *Take the previous setting with $G$ acting on $X$ and $\mu$ a quasi-invariant probability on $X$. Suppose $(g_n)_{n \geq 1} \subset G$ is a sequence with $\prod_n g_n \mu \sim \prod_n \mu$. Then for every measurable $A$ with $\mu(A) > 0$, there exists an integer $i > 0$ and $x \in A$ such that $g_n x$ returns infinitely often to $g_i(A)$.*

PROOF. First of all, recall that we have also $\prod_n g_n^{-1} \mu \sim \prod_n \mu$ so by Kakutani's theorem, $\sum_{m \geq 1} d_H(g_n^{-1}\mu, \mu)^2 < \infty$, where $d_H(g_n^{-1}\mu, \mu) = \|\psi(g_n^{-1}) - 1\|_2$ is the Hellinger distance. Then $d_H(g_n^{-1}\mu, \mu) \to 0$ as $n \to \infty$. However, $\{g\mu\}$ is a model dominated by a probability. In these circumstances it is known that the Hellinger distance dominates the variational distance

$$\rho(g\mu, \mu) = \sup |g\mu(B) - \mu(B)|,$$

with the supremum taken over the measurable sets $B$ [cf. Le Cam (1970)]. We conclude that $\lim_{n \to \infty} |\mu(g_n B) - \mu(B)| = 0$ for every measurable $B$.

Now, let $A$ be as in the statement. Take $\varepsilon > 0$ with $\varepsilon < \mu(A)$ and an integer $N > 0$ such that $|\mu(g_n A) - \mu(A)| < \varepsilon$ if $n \geq N$. We must prove that for some integer $i > 0$, there is an infinite number of $n$'s such that $g_n^{-1} g_i A \cap A \neq \phi$, that is, $g_i A \cap g_n A \neq \phi$.

Suppose on the contrary that for every $i \geq N$, the set $\{j > i : g_i A \cap g_j A = \phi\}$ is finite. Then we can select a sequence $i_k \to \infty$ such that the intersection between any two of the sets $g_{i_k} A$ is empty. Since $\mu(g_{i_k} A) \geq \mu(A) - \varepsilon > 0$, this contradicts the hypothesis that $\mu$ is probability. $\square$

In order to demonstrate how to apply this theorem to our original problem, we include Corollary 6.2, which involves the following terminology: A sequence $y_n$ in a topological space $Y$ is said to converge to $\infty$ if this happens in the one-point compactification of $Y$, that is, if for any compact $K \subset Y$, $y_n \notin K$ if $n \geq N$ for some $N > 0$. The sequence is unbounded if some subsequence convergences to $\infty$. It is bounded otherwise.

COROLLARY 6.2. *Suppose the support of $\mu$ is all of $X$. Suppose also that for every sequence $g_n \in G$ with $g_n \to \infty$, $g_n x \in X$ is unbounded. Then $g_n \in E(\mu)$, only if $g_n$ is bounded. In this case Theorem 3.1 applies.*

PROOF. Since $X$ is assumed to be locally compact, there is a compact $K$ with $\mu(K) > 0$. Since $g_n x$ is unbounded, it can be assumed that $g_n x \to \infty$ by passing to a subsequence. But this contradicts the fact that $g_n x$ returns infinitely often to the compact $g_i(K)$. □

A particular instance of this corollary is when $X = G$, where $G$ acts by left translations.

Typically Theorem 6.1 applies when there is no recurrence of sequences $g_n \in G$ with $g_n \to \infty$. We do not formulate any general statement about this but include some examples of this phenomenon below.

## 7. Examples.

7.1. Take $G = R^d = X$ with the action given by translation: $(x, y) \in G \times X \to x + y \in X$. The case $d = 1$ is treated in the pioneering work by Shepp (1965). Theorem 3.1 above specializes to Theorem 1 of Shepp (1965) except that it is restricted to bounded sequences. However Corollary 6.2 applies so this restriction can in fact be removed. Note that the action of $G$ on $X$ is transitive so the continuity assumption in Theorem 3.1(a) holds by Proposition 3.4.

7.2. Let $G$ be the group of rigid motions on $\mathbb{R}^d$, acting canonically on $X = \mathbb{R}^d$. This is the situation covered by Steele (1986). Our Theorem 3.1(a) gives a slight extension of Theorem 1 in Steele (1986) (our formulation allows basic measures which are invariant by subgroups). Also (b) and (c) of Theorem 3.1 specializes verbatim to Theorems 2 and 3 of Steele (1986), respectively. We note that the restriction that the sequence converges to the identity is not necessary [Steele (1986), Theorem 1]. In fact, $G$ and its action on $X = \mathbb{R}^d$ are easily seen to satisfy the conditions of Corollary 6.2.

7.3. This is an example (more properly a family of examples) showing a way of constructing sequences $g_n \to \infty$ in $G$ that still lie in $E(\mu)$. Let $K$ be a compact group and $G$ a one-parameter irrational flow on $K$, that is, a one-parameter subgroup whose closure has dimension greater than 1. For instance $K$ could be the two-torus $T^2 = \mathbb{R}^2/Z^2$ and $G$ the subgroup $\{(t, \alpha t) \bmod Z^2$: $t \in \mathbb{R}$, $\alpha$ a fixed irrational$\}$. Let $K$ act by left-translation on itself and restrict this action to $G$. We take for $\mu$ a probability of the type $d\mu = cf\,d\lambda$, where $d\lambda$ is Haar measure on $K$, $f = 1 + \varphi$, with $\varphi$ a smooth positive function with support contained in some small neighborhood of the identity and $c$ a constant of normalization. With $\mu$ taken this way only the identity in $K$ fixes $\mu$. Theorem 3.1 applies to the family $\{k\mu\}_{k \in K}$ and since there are sequences $g_m \in G$ with $g_m \to \infty$ (in the topology of $G$) and $g_m \to 1$ (in the topology of $K$), it is clear that there are also sequences $g_n \subset G$ with $\prod_n g_n \mu \sim \prod_n \mu$ such that $g_n \to \infty$ in the intrinsic topology of $G$.

7.4. This example illustrates an application of Theorem 6.1 in a situation not covered by Corollary 6.2. Take $G = Gl^+(d, \mathbb{R})$, the group of invertible $d \times d$ matrices with positive determinant, $X = \mathbb{R}^d$ and the linear action of $G$

on $X$. Let $\mu$ be the normal distribution $N(0,1)$. Then $\{g\mu\}_{g \in G}$ becomes the family of zero-mean normal distributions and $H$ in our previous notation is the compact group $SO(d, \mathbb{R})$ of orthogonal matrices. Of course, Theorem 3.1 applies in full generality.

In order to handle unbounded sequences, embed $G$ in the space of $d \times d$ matrices with its $\mathbb{R}^{d^2}$-canonical topology. Then it is easy to see that if $g_n \to \infty$ in $G$, then either (i) there is a subsequence $g_{n_k} \to \infty$ in $\mathbb{R}^{d^2}$ or (ii) there is a subsequence $g_{n_k}$ converging in $\mathbb{R}^{d^2}$ to a singular matrix. As to the first case, an argument similar to the proof of Corollary 6.2 shows that $\prod_n g_n\mu \perp \prod_n \mu$. For the second case appeal directly to the recurrence approach of Theorem 6.1 to show that $\prod_n g_n \mu \perp \prod_n \mu$. For this, suppose that $g_n$ is a sequence in $Gl^+(d, \mathbb{R})$ converging to the singular matrix $P$. Put $V = \operatorname{im} P$. Clearly, $k = \dim V \leq d - 1$. If we check that $\bigcup_n g_n^{-1}(V)$ is not dense we are done. In fact, in this case there is a compact $K \subset \mathbb{R}^d$ with nonempty interior such that $g_n K \cap V = \phi$ for every $n$. Since $g_n \to P$, $g_n x \to Px \in V$ for all $x \in \mathbb{R}^d$, so it is impossible for $g_n x$, $x \in K$, to return infinitely often to $g_i K$ for any $i > 0$. Thus the condition of Theorem 6.1 is violated so that $(g_n)_{n \geq 1} \notin E(\mu)$.

To see that $\bigcup_n g_n^{-1}(V)$ is not dense, exploit the Grassmannian of $k$-planes in $\mathbb{R}^d$. Passing to a subsequence if necessary we can assume that $g_n^{-1}(V)$ converges in the Grassmannian. This is enough to ensure that the set $(\bigcup_n g_n^{-1}(V)) \cap S^{d-1} \subset \mathbb{R}^d$ is compact, thus showing that $\bigcup_n g_n^{-1}(V)$ is not dense in $\mathbb{R}^d$.

We stress that this method of dealing with unbounded sequences does not depend on the specific $\mu$ but only on the way $G$ acts on $X$.

7.5. Let $G = Gl^+(d, \mathbb{R})$ and $X = G/SO(d, \mathbb{R})$ the space of positive-definite $d \times d$ matrices. The action of $G$ on $X$ is given by $g(s) = gsg^t$, $s \in X$ (where $t$ indicates transposition). It is clear that if $g_n \to \infty$ in $G$, then $g_n s g_n^t$ is unbounded in $X$, so Corollary 6.2 applies and Theorem 3.1 works for bounded or unbounded sequences in $G$. Here the probability $\mu$ can be taken to be any quasi-invariant measure.

One specific $\mu$ is the Wishart distribution $W(1, d, n)$ with $n$ degrees of freedom and the identity as scale matrix. The family $\{g\mu\}_{g \in G}$ becomes the Wishart family of distributions $W(\Sigma, p, n) = g\mu$ with $\Sigma = gg^t$.

7.6. The above example is in fact typical for the following class: Let $G$ be a semisimple or reductive Lie group and $K$ a maximal compact subgroup. Take $X = G/K$ and $\mu$ any quasi-invariant probability on $X$. By considering, for example, a Cartan decomposition of $G$, one checks easily that the action of $G$ on $X$ satisfies the conditions of Corollary 6.2, so Theorem 3.1 applies for any sequence $g_n \in G$. Note that because of Proposition 3.4, the assumption in Theorem 3.1(a) need not be checked.

Another example in this class is the hyperboloid model [cf. Barndorff-Nielsen (1967)]. For this model $G$ is the semisimple Lie group $SO(1, d-1)$ and $K = SO(d-1, \mathbb{R})$ and the symmetric space $G/K$ is the unit hyperboloid $H^{d-1} = \{(x_0, \dots, x_{d-1}) \in \mathbb{R}^d = x_0^2 - (x_1^2 + \cdots + x_{d-1}^2) = 1\}$.

## REFERENCES


BARNDORFF-NIELSEN, O. E. (1987). *Differential and Integral Geometry in Statistical Inference. IMS Lecture Notes—Monograph Series* **10**. IMS, Hayward, Calif.

BOURBAKI, N. (1963). *Eléments de Mathématique: Intégration*. Hermann, Paris.

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

DAVIES, E. B. (1980). *One Parameter Semigroups*. Academic Press, New York.

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

IBRAGIMOV, I. A. and HAS'MINSKII, R. Z. (1981). *Statistical Estimation—Asymptotic Theory*. Springer, New York.

KAKUTANI, S. (1948). On equivalence of infinite product measures. *Ann. of Math. (2)* **49** 214–224.

KUNITA, H. (1984). Stochastic differential equations and stochastic flow of diffeomorphism. *Ecole d'Éte de Probabilités de Saint-Flour XII. Lecture Notes in Math.* **1097** 144–303. Springer, New York.

LE CAM, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates. *Ann. Math. Statist* **41** 802–828.

MARQUES, M. S. F. (1987). A study on the Lebesgue decomposition of measures induced by stable processes. Ph.D. dissertation, Univ. North Carolina, Chapel Hill.

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

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

STEELE, J. (1986). Fisher information and detection of a Euclidean perturbation of an independent stationary process. *Ann. Probab.* **14** 326–335.

WARNER, G. (1972). *Harmonic Analysis on Semi-simple Lie Groups*. Springer, New York.

WATANABE, S. (1984). *Stochastic Differential Equations and Malliavin Calculus*. Springer, New York.



INSTITUTO DE MATEMÁTICA,
ESTATÍSTICA E CIÊNCIA DA COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS
CAIXA POSTAL 6065
13081, CAMPINAS, S.P.
BRASIL