Bochner and Bernstein theorems via the nuclear integral representation theorem

Erik G.F. Thomas

Mathematics Institute, Postbus 800, 9700 AV Groningen, The Netherlands

Received 9 February 2004

Submitted by R.M. Aron
To John Horváth on his 80th birthday

Abstract

We show how, using the nuclear integral representation theorem, the Bernstein–Choquet theorem and Bochner–Schwartz theorem may be derived. In the case of the Bernstein–Choquet theorem we give an example, determining the representing measure explicitly.

© 2004 Elsevier Inc. All rights reserved.

1. Introduction

The elegance of the theory of integral representations, Choquet theory, is well known. Suffice it to recall that starting at the Krein–Millman theorem according to which a compact convex set in a locally convex Hausdorff space is the closed convex hull of its extreme points, Choquet proved that for a compact convex metrizable set $K$ things can be formulated much more precisely: every point of $K$ is the resultant of a probability measure concentrated on the set of extreme points. Moreover the measures are uniquely determined if and only if the set $K$ is a simplex. Every compact convex set $K$ can be regarded as the intersection of a convex cone $F$ and a non homogeneous hyperplane. Then $K$ is a simplex if and only if $F$ is a lattice for its proper order. Choquet then proceeded with a vast generalization, in which convex cones, with or without compact base, but mostly weakly complete, were the primary object of attention.
Meanwhile, in the theory of group representations the relevance of Grothendieck’s nuclear spaces was demonstrated in connection with spectral theory and the decomposition of a unitary representation into a direct integral of irreducible representations [6,9,10,16]. In a cone consisting of Hilbert spaces or their reproducing kernels, irreducible representations correspond to extremal kernels [16]. In the context of group representation theory uniqueness meant uniqueness up to isomorphism. However, to obtain generalized Fourier analysis, with the property, as in Fourier analysis on \( \mathbb{R}^n \) or on the torus \( \mathbb{T}^n \), that operators commuting with the group action are necessarily diagonal, the uniqueness in the sense of Choquet theory, turned out to be essential [17].

These two parallel theories gave rise to the nuclear integral representation theorem, in which conuclear spaces play an essential role, but which is formulated as in the Choquet theory. The usefulness of the theorem comes from the ubiquity of nuclear and conuclear spaces. We assume knowledge of the basic theory of topological vector spaces such as in [8], but we recall a definition and examples of nuclear and conuclear spaces below.

In the paper [3], Choquet shows how the integral representation theory can be applied to prove Bernstein’s theorem on absolutely monotonic functions [2], and its higher dimensional analogue, and Bochner’s theorem. The purpose of the present paper is similarly to show how the nuclear integral representation theorem can be used to obtain the Bernstein–Choquet theorem on completely monotonic functions, as well as the Bochner–Schwartz theorem. In Choquet’s paper [3] the finite difference definition of absolutely monotonic functions is used, leading to a weakly complete cone; in the present paper the equivalent \( C^\infty \)-definition of completely monotonic functions is used. The completely monotonic functions considered here correspond to absolutely monotonic functions by a simple change of variable.

2. Nuclear and conuclear spaces

Given a locally convex space \( E \) one associates two kinds of Banach space with \( E \), and correspondingly two classes of powerful spaces, the nuclear and the conuclear spaces.

If \( p \) is a continuous seminorm on \( E \) the quotient space \( E_p = E / \{ x \in E : p(x) = 0 \} \) is naturally a normed space, the norm of the image of \( x \in E \) under the quotient map being \( p(x) \). We denote \( \hat E_p \) the Banach space obtained on completing this normed space. If \( q \) is another continuous seminorm, and such that \( p \leq q \), the natural map \( E_q \rightarrow E_p \) is continuous, with norm at most 1, giving rise to the canonical map \( \hat E_q \rightarrow \hat E_p \).

Given a closed convex balanced (stable under multiplication by scalars of modulus \( \leq 1 \)) bounded subset \( A \subset E \), the subspace \( E_A = \bigcup_{\lambda \geq 0} \lambda A \subset E \) is naturally a normed space, the norm being the gauge of \( A \), \( x \mapsto \inf_{\lambda \geq 0, x \in \lambda A} \lambda \). \( A \) being the unit ball. We denote \( \hat E_A \) the completion. If \( B \) is another such subset of \( E \) and so that \( A \subset B \), the inclusion map \( E_A \subset E_B \) is continuous, with norm at most 1, and gives rise to the canonical map \( \hat E_A \rightarrow \hat E_B \).

These constructions can be applied to the dual space \( E' \) of continuous linear forms \( x \mapsto \langle x, x' \rangle \) on \( E \). If \( A = \{ x' \in E' : |\langle x, x' \rangle| \leq p(x) , x \in E \} \) then the dual of \( \hat E_p \) is canonically isomorphic to the space \( E_A' \).
Finally, recall that a linear map $u : E_1 \to E_2$ from a Banach space $E_1$ to a Banach space $E_2$ is said to be nuclear if there exists a sequence $(x_k')_{k \in \mathbb{N}}$ in $E_1'$ and a sequence $(y_k)_{k \in \mathbb{N}}$ in $E_2$ so that
\[
\sum_{k \in \mathbb{N}} \|x_k'\| \|y_k\| < +\infty \quad \text{and such that}
\]
\[
u(x) = \sum_{k \in \mathbb{N}} \langle x, x_k' \rangle y_k, \quad x \in E_1.
\]
A locally convex Hausdorff space is said to be nuclear if for every continuous seminorm $p$ there exists a continuous seminorm $q \geq p$ such that the canonical map $\hat{E}_q \to \hat{E}_p$ is a nuclear map.

A locally convex Hausdorff space $E$ is said to be conuclear if for every closed convex balanced bounded subset $A \subset E$ there exists another, $B \subset E$, such that $A \subset B$ and such that the canonical map $\hat{E}_A \to \hat{E}_B$ is a nuclear map.

These dual notions are practically dual also, at least if $E$ is a barreled space: then $E$ is nuclear if and only if $E'$ is conuclear. It can be shown [7,12] that the dual of a nuclear Fréchet space is nuclear. Consequently every nuclear Fréchet space, being reflexive, is also conuclear. It follows by the stability theorems on nuclear spaces, that many important spaces are both nuclear and conuclear.1

The spaces from the theory of distributions $\mathcal{D}(V), \mathcal{E}(V), \mathcal{S}(\mathbb{R}^n), V$ a $C^\infty$ manifold, and their duals $\mathcal{D}'(V), \mathcal{E}'(V), \mathcal{S}'(\mathbb{R}^n)$ are all nuclear and conuclear, as is the space of holomorphic functions $\mathcal{H}(V)$ on a complex manifold $V$. The sequence spaces $C^\infty_n, s(n), s'(n)$ also are nuclear and conuclear. For details we refer to [7, Chapter II, §2, no. 3], [14, Chapter IV, §2].

3. The nuclear integral representation theorem

We recall the elements of integral representation theory.

Let $F$ be a locally convex Hausdorff space, and let $\Gamma \subset F$ be a closed convex cone in $F$, with vertex $\{0\}$, assumed proper: $\Gamma \cap -\Gamma = \{0\}$. One orders the space $F$ by putting
\[
f \leq g \iff g - f \in \Gamma.
\]
In particular
\[
\Gamma = \{f \in F: \ f \geq 0\}.
\]
The order thus defined on $\Gamma$ is called the proper order of $\Gamma$. If any two elements $f$ and $g$ in $\Gamma$ have a smallest common majorant, the cone $\Gamma$ is said to be a lattice.

If $f \in \Gamma$ the face generated by $f$ is the set $\Gamma(f) = \{g \in \Gamma: \ \exists \lambda \in \mathbb{R}_+, \ 0 \leq g \leq \lambda f\}$. This is a convex subcone whose proper order equals the order induced on $\Gamma(f)$ by $\Gamma$.

Clearly, $\Gamma$ is a lattice iff $\Gamma(f)$ is a lattice for all $f \in \Gamma$.

For $f \in \Gamma$ we denote the order interval:
\[
I(f) = \{g \in \Gamma: \ 0 \leq g \leq f\} = \Gamma \cap (f - \Gamma).
\]

1 An infinite dimensional Banach space can be neither nuclear nor conuclear, the identity not being compact. An infinite dimensional Banach space, equipped with its weak topology, is nuclear but not conuclear.
Recall that an element $e \in \Gamma$ is extremal if $g + h = e$ with $g, h \in \Gamma$ implies that $g = \lambda e$ for some number $\lambda \geq 0$ and similarly for $h$. Equivalently: the face $\Gamma(e)$ is a halfline $\mathbb{R}_+ e$.

Any element proportional to $e$, i.e., of the form $\lambda e$, $\lambda \geq 0$, is then extremal also. We denote $\text{ext}(\Gamma)$ the set of extremal elements of $\Gamma$ and $\text{ext}(\Gamma)_* = \text{ext}(\Gamma) \setminus \{0\}$.

If $\text{ext}(\Gamma)$ has enough elements we consider a parametrization of the extreme rays as follows: it consists of a parameter space $T$ and a continuous map $T \to \text{ext}(\Gamma)_*$, $t \mapsto e_t$, such that every $e \in \text{ext}(\Gamma)_*$ is proportional to $e_t$ for precisely one $t \in T$.

We recall the nuclear integral representation theorem [18]:

**Theorem 1.** Let $F$ be a conuclear locally convex space.²

1. Let $\Gamma \subset F$ be a closed convex cone, such that the order intervals $I(f)$, $f \in \Gamma$, are bounded subsets of the topological vector space $F$. Then $\Gamma = \overline{\text{co}} \text{ext}(\Gamma)$ equals the closed convex hull of its extreme generators.

2. If $T \to \text{ext}(\Gamma)_*$, $t \mapsto e_t \in \text{ext}(\Gamma)_*$, is an admissible parametrization of the extreme rays then

   (A) For every $f \in \Gamma$ there is a Radon measure $m$ on $T$ such that
   
   $$f = \int_T e_t \, m(dt).$$

   (B) The measure $m$ is uniquely determined by $f$ iff the face $\Gamma(f)$ is a lattice. In particular, the representing measure is unique for every $f \in \Gamma$ iff $\Gamma$ is a lattice.

The term ‘admissible parametrization’ involves some measure theoretic technicalities [18]. But a parametrization in which $T$ is a Suslin space (e.g., a locally compact Hausdorff space with a countable base of open sets) is always admissible.

The vector integral in (3) just means, $F'$ denoting the dual of $F$,

$$\int_T |\langle e_t, \ell \rangle| \, m(dt) < +\infty, \quad \langle f, \ell \rangle = \int_T \langle e_t, \ell \rangle \, m(dt), \quad \ell \in F'.$$

By Radon measure we mean a positive, locally finite Borel measure, inner regular with respect to compact sets (cf. [4,14]).

If $\Gamma$ is a weakly complete proper convex cone in a locally convex space $F$, the order intervals $\Gamma \cap (f - \Gamma)$ are bounded (Choquet [4, Proposition 30.10]).

The condition that the order intervals be bounded in the topology of $F$ is not enough to ensure the existence of extreme rays if the space $F$ is not conuclear: for example the set of non-negative elements of $L_2[0, 1]$ is a weakly complete cone, but it has no extreme rays.

On the other hand a closed convex proper cone, with unbounded order intervals, in a conuclear space, does not necessarily have any extreme rays. An example is the cone $C_+^\infty(\mathbb{R})$ of non negative functions in $C_0^\infty(\mathbb{R})$. Here every non zero $f \in C_+^\infty(\mathbb{R})$ can be decomposed as a sum $f = g + h$, with $g, h \in C_+^\infty(\mathbb{R})$ not proportional to $f$ (partition of

² Assumed Hausdorff and quasi-complete, i.e., the closed bounded sets are complete.
unity), i.e., no \( f \neq 0 \) is extremal. The order interval \( I(f) \) is unbounded in the \( C^\infty \)-topology however, unless \( f = 0 \).

One of the first examples of integral representation was the following: let \( \Omega \) be an open subset of \( \mathbb{R}^n \). Let \( F = \text{Harm}(\Omega) \) be the space of harmonic functions \( \{ u : \Omega \to \mathbb{R}, \Delta u = 0 \} \) with the topology of uniform convergence on compact subsets of \( \Omega \), and let \( \Gamma = \{ u \in \text{Harm}(\Omega): u \geq 0 \} \). Then \( F \) is a nuclear Fréchet space, being a closed subspace of \( C^\infty(\Omega) \) by the ellipticity of \( \Delta \), and \( \Gamma \) is a closed convex cone obviously having bounded order intervals. Thus, it follows that \( \Gamma \) is the closed convex hull of its extreme generators, parameterized in good cases by the boundary points of \( \Omega \) via a Poison kernel. In this example the Laplace operator may be replaced by any hypo-elliptic operator, but the explicit representations by extremals are mostly unknown.

4. The Bernstein–Choquet theorem

The purpose of this section is to show how Bernstein’s theorem, and more generally the theorem of Bernstein–Choquet, may be deduced from the nuclear integral representation theorem, knowing that the space of \( C^\infty \) functions on an open subset of \( \mathbb{R}^m \) is a nuclear Fréchet space, and therefore a conuclear space.

For the traditional treatment in the case \( m = 1 \) see [2,19] and for the treatment using Choquet theory [3,4,11].

Let \( C \) be an open convex cone in \( \mathbb{R}^m \), with vertex \( \{0\} \) (stable under addition and multiplication by strictly positive scalars), with proper closure. Let \( C^\circ = \{ \lambda \in \mathbb{R}^m: \lambda x \geq 0, \forall x \in C \} \) be the polar, a closed convex cone in the space \( \mathbb{R}^m \), which is identified to its dual by means of the inner product

\[
\langle \lambda x, y \rangle = \sum_{i=1}^m \lambda_i x_i,
\]

For instance, if \( C = (0, +\infty)^m \) then \( C^\circ = [0, +\infty)^m \).

The function \( f : C \to \mathbb{R} \) is said to be completely monotonic if \( f \) is of class \( C^\infty \), \( f \geq 0 \), if for all \( h \in C \) the directional derivative \( D_h f(x) = \frac{d}{dt} f(x + th)|_{t=0} \) is \( \leq 0 \), and if more generally, for any finite sequence \( h_1, \ldots, h_n \) of vectors in \( C \), one has

\[
(-1)^k D_{h_1,\ldots,h_n} f(x) \geq 0 \text{ for all } x \in C, \quad D_{h_1,\ldots,h_n} f \text{ being an abbreviation for } D_{h_n} D_{h_{n-1}} \ldots D_{h_1} f.
\]

Thus, a function \( f : (0, +\infty)^m \to \mathbb{R} \) is completely monotonic if \( (-1)^k D^k f(x) \geq 0 \) for all \( x \in (0, +\infty)^m \), where \( D^k \), with \( k = (k_1, \ldots, k_m) \), as usual indicates \( k_i \) derivatives with respect to \( x_i \), and \( |k| = k_1 + \cdots + k_m \).

Let \( \Gamma \) be the set of completely monotonic functions on \( C \).

It is a closed convex cone in the space \( C^\infty(C) \), equipped with the \( C^\infty \) topology, defined by the seminorms

\[
p_{n,H}(f) = \sup_{|k| \leq n} \sup_{x \in H} |D^k f(x)|,
\]

where \( H \) is a compact subset of\( C \), and \( n \in \mathbb{Z}_+ \), \( D^0 f = f \). Because \( C - C = \mathbb{R}^m \), an equivalent system of seminorms is the following:

\[
\sup_{x \in H} |D_{h_1,\ldots,h_n} f(x)|
\]

with \( h_1, \ldots, h_n \in C, n \in \mathbb{Z}_+ \), where for \( n = 0 \), \( D_{h_1,\ldots,h_n} f = f \).
Theorem 2 (Bernstein–Choquet). For every completely monotonic \( f : C \to \mathbb{R} \) there exists a unique Radon measure \( m \geq 0 \) on \( C^\circ \) such that
\[
f(x) = \int_{C^\circ} e^{-\lambda x} m(d\lambda), \quad x \in C.
\] Conversely, if \( m \) is a Radon measure on \( C^\circ \) such that these integrals are finite for all \( x \in C \), then \( f \) is a completely monotonic function.

Remark 3. A function \( f : C \to \mathbb{R} \) is completely monotonic if and only if the function \( x \mapsto F(x) = f(-x) \) defined on the cone \(-C\), is absolutely monotonic, i.e., has all its derivatives \( D_{k_1,\ldots,k_n} F(x) \), \( k_i \in -C \), positive. In [2,3] the representation analogous to (7) was obtained for absolutely monotonic functions.

Proof. The space \( C^\infty(C) \) is conuclear and the order intervals
\[
I(f) = \left\{ g \in C^\infty(C) : 0 \leq (-1)^n D_{h_1,\ldots,h_n} g(x) \leq (-1)^n D_{h_1,\ldots,h_n} f(x), \quad x \in C, \quad h_j \in C, \quad n \in \mathbb{Z}_+ \right\}
\] are bounded in the \( C^\infty \) topology, so \( \Gamma \) is the closed convex hull of its extreme generators. Let \( h \in C \), and let \( f_h(x) = f(x+h) \). Then obviously \( f_h \in \Gamma \). Since \( f(x+h) = f(x) + \int_0^1 D_h f(x + th) \, dt \), we have \( 0 \leq f_h(x) \leq f(x) \). Applying this to the functions \((-1)^n D_{h_1,\ldots,h_n} f\), noting that differentiations and translations commute, it follows that \( 0 \leq f_h \leq f \) for the order defined by \( \Gamma \). Thus, if \( f \) is extremal, \( f_h = cf \), the constant \( c \) depending on \( h \). Differentiating, it is seen that \( D_h f = kf \), where \( k \) is a constant, depending linearly on \( h \), and \( \leq 0 \) because \( f \) belongs to \( \Gamma \). Therefore there exists \( \lambda \in C^\circ \) such that \( D_h f = -\lambda h f \) for all \( h \in C \). It follows that \( D_h(e^{\lambda x} f) = 0 \) for all \( h \in C \), which implies that \( e^{\lambda x} f(x) = c \) is constant, i.e., \( f(x) = ce^{-\lambda x} \).

Conversely, it is obvious that the functions \( e^{\lambda x}(x) = e^{-\lambda x} \), with \( \lambda \in C^\circ \), belong to \( \Gamma \). We shall prove that they are extremal.

In the case where \( C = (0, +\infty) \) this follows most easily by an elegant method due to Choquet: the transformations \( f \mapsto (r, f) \), defined for \( r > 0 \) by \( (r, f)(x) = f(rx) \) are automorphisms of \( \Gamma \) so leave \( \text{ext}(\Gamma) \) invariant. Since \( \Gamma \) is the closed convex hull of its extreme generators, there exists \( e_{\lambda_0} \in \text{ext}(\Gamma) \) with \( \lambda_0 > 0 \), and so all exponentials \( e^{\lambda x} \) with \( \lambda > 0 \) belong to \( \text{ext}(\Gamma) \). On the other hand \( e_1 = 1 \) is extremal also.

In the multidimensional case this method does not seem to be as convenient, because the boundary of \( C \) may be more complex. Instead we use the following notion and lemma.

Let \( K \) be the set of symmetric kernels \( K(x, y) \in \mathbb{R}, \ x \in C, \ y \in C \), which are positive semi-definite: \( \sum_{1 \leq i, j \leq n} K(x_i, x_j)\alpha_i\alpha_j \geq 0, \ n \in \mathbb{N}, \ (x_1, \ldots, x_n) \in C^n, \ (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n \).

Lemma 4. The set \( K \) is a convex cone in \( \mathbb{R}^{C \times C} \) whose extremal generators are the rank one kernels \( f \otimes f : (x, y) \mapsto f(x)f(y) \).

Proof. It is well known that these kernels are precisely the reproducing kernels of Hilbert subspaces of \( \mathbb{R}^C \) (cf. [1]). The extremal Hilbert subspaces are obviously those of dimension one, \([f]\). But \([f]\) has the reproducing kernel \( f \otimes f \). \( \square \)
For \( f \in \Gamma \), let \( K_f(x, y) = f(x + y) \). Then if \( f \) is extremal, we have seen that it is proportional to an exponential, which implies that \( K_f(x, y) = c f(x) f(y) \) for some \( c \geq 0 \). In particular \( K_f \) belongs to \( \mathcal{K} \). Thus, \( \Gamma \) being the closed convex hull of its extremal generators, \( K_f \) belongs to \( \mathcal{K} \) for all \( f \in \Gamma \). The map \( f \mapsto K_f \) being one-to-one, we may view \( \Gamma \) as a subcone of \( \mathcal{K} \).

5. A special case

Consider, in the case \( m = 2 \) the function, defined, for \( \alpha \geq 0 \), on the positive quadrant \( Q = [0, +\infty)^2 \)

\[
f_\alpha(x, y) = \frac{1}{1 + x + y + \alpha xy}.
\]

This function is obviously completely monotonic for \( \alpha = 0 \) and \( \alpha = 1 \):

\[
\frac{1}{1 + x + y} = \int_0^{+\infty} e^{-sx} e^{-y} ds,
\]

\[
\frac{1}{1 + x + y + xy} = \int_0^{+\infty} \int_0^{+\infty} e^{-sx} e^{-ty} e^{-t} ds dt.
\]

These are representations as in (7), with in the first case the measure concentrated on the diagonal.

**Proposition 5.** The function \( f_\alpha \) is completely monotonic if and only if \( 0 \leq \alpha \leq 1 \).

**Proof.** We shall use the ‘product rule’ according to which the product of two completely monotonic functions is completely monotonic. Let \( \Gamma(m) \) the cone of completely
monotonic functions on \((0, +\infty)^m\). To prove that \(f_\alpha\) is completely monotonic if \(0 \leq \alpha \leq 1\) we have to prove

\[
(-1)^{k+\ell} \frac{\partial^{k+\ell}}{\partial x^k \partial y^\ell} f_\alpha(x, y) \geq 0.
\] (12)

Obviously

\[
\left(-\frac{\partial}{\partial y}\right)^\ell f_\alpha(x, y) = \ell! \frac{(1 + \alpha x)^\ell}{((1 + \alpha x) y + 1 + x)^{\ell+1}}.
\] (13)

Rather than calculating the derivatives of this with respect to \(x\) we prove that this function belongs to \(\Gamma(1)\) by using the product rule. The function in (13) is the product of the function

\[
\frac{1}{1 + x + y + \alpha xy} = \frac{1}{a + bx}
\] (14)

with \(a > 0, b > 0\), which obviously is completely monotonic with respect to \(x\), and functions

\[
\frac{1 + \alpha x}{1 + x + y + \alpha xy}
\] (15)

We pose \(t = 1 + x + y + \alpha xy\). Then this expression is

\[
\frac{1}{1 + \alpha x} \frac{1 - a + at}{t}
\] (16)

which is a constant times a convex combination of 1 and 1/t, hence completely monotonic in \(t = a + bx\), with \(a > 0\) and \(b > 0\), and therefore in \(x\). This shows that (13) is completely monotonic in \(x\), so that \(f_\alpha\) is completely monotonic.

To finish the proof, we have to show that for \(\alpha > 1\) the function \(f_\alpha\) is not completely monotonic. We omit the details of this, but it is not hard to see, suggested by a calculation in Mathematica, that some low order derivative (e.g., \(k = 2, \ell = 2\)) fails to be \(\geq 0\).

Next we show that it is possible to calculate the representing measure \(\nu_\alpha\) explicitly. We assume \(0 < \alpha \leq 1\). First regard \(f_\alpha\) as a completely monotonic function of \(y\) and determine the measure \(\mu_x\) so that

\[
\frac{1}{1 + x + y + \alpha xy} = \int_0^{+\infty} e^{-s y} \mu_x(ds).
\]

This is easily obtained by writing the function in the form

\[
\frac{1}{A + By} = \frac{1}{(1 + x) + (1 + \alpha x)y} = \int_0^{+\infty} e^{-s(1+x) - s(1+\alpha x)y} ds.
\]

Changing to the variable \(u = s(1 + \alpha x)\) this becomes

\[
\frac{1}{1 + \alpha x} \int_0^{+\infty} e^{-uy} e^{-\frac{x}{1+\alpha x}} du.
\]
so that
\[ \mu_x(ds) = \frac{1}{1 + \alpha x} e^{-\frac{1 + \alpha x}{1 + \alpha x}} ds. \]

This is a completely monotonic function of \( x \) with values in the positive measures, and we have to represent it as such. For convenience we first work with the variable \( t = 1 + \alpha x \).

Then \( x = (t - 1)/\alpha \) and
\[ \frac{1 + x}{1 + \alpha x} = \frac{1 - \alpha}{\alpha} \frac{1}{t - \alpha}. \]

Thus we have
\[ \frac{1}{1 + \alpha x} e^{-\frac{1 + \alpha x}{1 + \alpha x}} = \frac{1}{t} e^{-\frac{x}{\alpha}} e^{\beta t} \]

with \( \beta = (1 - \alpha)/\alpha \). To represent this as a function of \( t \) we develop the second exponential and write this as
\[ e^{-s/\alpha} \sum_{n=0}^{+\infty} \frac{1}{n!} (\beta s)^n \frac{1}{t^{n+1}}. \]

Thus we need only represent \( 1/t^{n+1} \). We have
\[ \frac{1}{t^{n+1}} = \frac{1}{n!} \int_0^{+\infty} \lambda^n e^{-\lambda t} d\lambda. \]

Therefore our expression becomes
\[ e^{-s/\alpha} \int_0^{+\infty} \sum_{n=0}^{+\infty} \frac{1}{(n!)^2} (\beta s)^n e^{-\lambda t} d\lambda = e^{-s/\alpha} \int_0^{+\infty} \varphi(\beta s \lambda) e^{-\lambda t} d\lambda. \]

where
\[ \varphi(u) = \sum_{n=0}^{+\infty} \frac{u^n}{(n!)^2}. \]

Substituting \( t = 1 + \alpha x \) we get
\[ e^{-s/\alpha} \int_0^{+\infty} \varphi(\beta s \lambda) e^{-\lambda} e^{-\alpha \lambda x} d\lambda. \]

Changing to \( \tau = \alpha \lambda \) this becomes
\[ \frac{1}{\alpha} e^{-s/\alpha} \int_0^{+\infty} \varphi(\beta s \tau / \alpha) e^{-\tau/\alpha} e^{-\tau x} d\tau. \]

Thus we have
\[ \mu_x(ds) = \int_0^{+\infty} e^{-\tau x} v_\alpha(ds, d\tau). \]
where

\[ \nu_\alpha(ds,d\tau) = \frac{1}{\alpha} e^{-\alpha s/\alpha} e^{-\tau/(\alpha)} \left( \frac{1}{\alpha} \right)^{\frac{\tau}{\alpha}} ds
d\tau. \]

Note that the function \( \varrho \) is essentially a Bessel function:

\[ \varrho(u) = J_0(2\sqrt{-u}) = I_0(2\sqrt{-u}). \]

6. The Bochner–Schwartz theorem

We prove, using the nuclear integral representation theorem, the theorem of Bochner–
Schwartz [13, Chapter VII, Theorem XVIII] according to which a distribution
\( T \in D'(\mathbb{R}^n) \) is the Fourier transform \( f \) of a positive temperate measure \( m \)
if and only if \( T \) is of positive type, i.e., noting \( \tilde{\phi}(x) = \phi(-x) \),

\[ T(\phi \ast \tilde{\phi}) \geq 0, \quad \phi \in D(\mathbb{R}^n). \]

We denote this

\[ T \gg 0. \]

It is easy to see that a temperate distribution of positive type is the Fourier transform
of a positive measure, essentially because the inverse Fourier transform makes sense. But
the fact that distributions of positive type are bounded, hence temperate, is not obvious
[13, pp. 195, 196, 201].

Here we apply the theory of integral representations with the conuclear space \( F = D'(\mathbb{R}^n) \)
and the cone of positive definite distributions on \( \mathbb{R}^n \):

\[ \Gamma = \{ T \in D'(\mathbb{R}^n) : T(\phi \ast \tilde{\phi}) \geq 0, \quad \phi \in D(\mathbb{R}^n) \}. \]

(1) The extremals are the distributions proportional to the characters \( e_\lambda, \lambda \in \mathbb{R}^n, e_\lambda(x) = e^{-2\pi i \lambda \cdot x} \). The map \( \lambda \mapsto e_\lambda \) is an admissible parametrization of the extreme rays.

**Proof.** An element \( T \in \Gamma \) defines a Schwartz reproducing kernel \( K_T(\phi,\psi) = T(\phi \ast \tilde{\psi}) \).

(2) It is bounded order intervals. Let \( I(S) = \{ T \in D'(\mathbb{R}^n) : 0 \ll T \ll S \} = \Gamma \cap (S - \Gamma) \). It suffices to prove that \( I(S) \) is weakly bounded, i.e.,

\[ \sup_{T \in I(S)} |T(\phi)| < +\infty \]

for all \( \phi \in D(G) \). Obviously (19) is true if \( \phi = \psi \ast \tilde{\psi} \) because then \( 0 \leq T(\phi) \leq S(\phi) \) for
all \( T \in I(S) \). By polarization (19) is still true if \( \phi \) is a convolution product \( \psi \ast \chi \) of two
test functions. Now, by a theorem of Dixmier and Malliavin [5], any \( \phi \in \mathcal{D} \) is a finite sum of such convolution products (for \( G = \mathbb{R}^n \) even the sum of two such convolution products). Below we give an elementary proof of (19) avoiding this remarkable theorem.

(3) By the nuclear integral representation theorem we have for every \( T \in \mathcal{F} \), an integral representation

\[
T = \int e_\lambda \, m(d\lambda)
\]

with

\[
\int |\langle e_\lambda, \phi \rangle| \, m(d\lambda) < +\infty, \quad \phi \in \mathcal{D},
\]

i.e.,

\[
\int |\hat{\phi}(\lambda)| \, m(d\lambda) < +\infty, \quad \phi \in \mathcal{D}.
\]

It remains to prove that \( m \) is temperate. By Fatou’s lemma the map \( \phi \mapsto \int |\hat{\phi}(\lambda)| \, m(d\lambda) \) is lower semi-continuous on \( \mathcal{D}_K \) for every compact \( K \). By the uniform boundedness principle it is continuous, i.e., if \( p_N(\phi) = \sup_{|\lambda| \leq N} ||D^k \phi||_{\infty} \), there exist \( N \in \mathbb{N} \) and \( M \geq 0 \) such that

\[
\int |\hat{\phi}(\lambda)| \, m(d\lambda) \leq Mp_N(\phi), \quad \phi \in \mathcal{D}_K,
\]

\( \mathcal{D}_K \) denoting the Fréchet space of test functions having their support in the compact set \( K \).

Take \( \phi = \psi \ast \tilde{\psi} \), \( \hat{\phi}(\lambda) \geq 0 \) and assume \( \hat{\phi}(0) > 1 \). \( \hat{\phi}(\lambda) \geq 1 \) for all \( \lambda \in B(\delta) \), \( \hat{\phi}(\lambda/r) \geq 1 \) \( \lambda \in B(r\delta) \). The function \( \hat{\phi}(\lambda/r) \) is the Fourier transform of \( r^n \phi_r \), \( \phi_r(x) = \phi(rx) \), whose support is contained in the ball \( K/r \subset K \) if \( r \geq 1 \). We therefore get

\[
m(B(r\delta)) \leq Mp_N(r^n \phi_r) \leq M\lambda^{n+N} p_N(\phi).
\]

Replacing \( r \) by \( r/\delta \) we get \( m(B(r)) = O(r^{N+n}) \), i.e., \( m \) is a temperate measure [13, VII, 4; 7]. By (20) we have, for \( \phi \in \mathcal{D} \),

\[
\langle T, \phi \rangle = \int \langle e_\lambda, \phi \rangle \, m(d\lambda) = \int \hat{\phi}(\lambda) \, m(d\lambda) = \{m, \hat{\phi}\} = \{\hat{m}, \phi\}.
\]

Consequently \( T \) is a temperate distribution and \( T = \mathcal{F}m \). \( \square \)

Here is an elementary proof of the fact that the order intervals in \( \mathcal{F} \) are bounded in \( \mathcal{D}'(\mathbb{R}^n) \). Instead of a finite decomposition \( \phi = \sum_k \varphi_k * \psi_k \) we produce a series decomposition, which serves the same purpose. We abbreviate \( \mathbb{R}^n = G \).

**Lemma 6.** For every \( \varphi \in \mathcal{D}(G) \) there exist sequences \( (\varphi_n) \) and \( (\psi_n) \) bounded in \( \mathcal{D}(G) \) and a sequence \( (\lambda_n) \) in \( \ell^1 \) such that

\[
\varphi = \sum_{n=1}^{\infty} \lambda_n \varphi_n * \psi_n.
\]
Proof. Consider the map \( P : \Phi \mapsto \varphi \) from \( \mathcal{D}(G \times G) \) to \( \mathcal{D}(G) \) defined by:

\[
\varphi(x) = \int \Phi(x - y, y) \, dy.
\] (27)

This is obviously a continuous linear map. Moreover \( P \) is a surjection. In fact choose \( \psi \) so that \( \int \psi(x) \, dx = 1 \). To obtain the image \( \varphi \) let \( \Phi(x - y, y) = \varphi(x + y) \psi(y) \), i.e., choose \( \Phi(x, y) = \varphi(x + y) \psi(y) \).

Now every \( \Phi \in \mathcal{D}(G \times G) \) has a series expansion converging in \( \mathcal{D}(G \times G) \):

\[
\Phi(x, y) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \psi_n(y).
\] (28)

This follows from known facts about topological tensor products, but, in the case of \( \mathbb{R}^n \), the argument is entirely elementary. If \( \text{supp}(\Phi) \subset (-T, T)^{2n} \) consider \( \Phi \) as a \( 2T \)-periodic function, expand as a Fourier series and multiply by cutoff functions \( \alpha \) and \( \beta \) in \( \mathcal{D}(\mathbb{R}^n) \) such that \( \text{supp}(\Phi) \subset \text{supp}(\alpha) \times \text{supp}(\beta) \subset (-T, T)^{2n} \) and \( \alpha(x)\beta(y) = 1 \) on the support of \( \Phi \):

\[
\Phi(x, y) = \sum_{k, \ell} c_{k, \ell} \alpha(x)e_k(x)\beta(y)e \ell(y), \quad x, y \in G,
\] (29)

convergence in the \( C^\infty \) topology with supports in a fixed compact set, i.e., convergence in the sense of \( \mathcal{D}(G) \).

Applying \( P \) to (28), with \( P(\Phi) = \varphi \), we get (26), whence (19). \( \square \)

Remark 7. A distribution \( T \) is bounded if \( T \ast \varphi \) is a bounded function for all \( \varphi \in \mathcal{D} \). To prove, as is done in [13], that \( T \) is bounded if \( T \ast \varphi \ast \psi \) is bounded for all \( \varphi \) and \( \psi \) in \( \mathcal{D} \), is quite delicate. This property follows immediately from the theorem of Dixmier and Malliavin mentioned above, but it also follows easily from the above Lemma 6. For if \( T \ast \varphi \ast \psi \) belongs to \( L^\infty(G) \) for \( \varphi \) and \( \psi \) in \( \mathcal{D}_K \), the map \( (\varphi, \psi) \rightarrow T \ast \varphi \ast \psi \in L^\infty(G) \) which is separately continuous, is jointly continuous, \( \mathcal{D}_K \) being a Fréchet space, and consequently by (26) we have \( T \ast \varphi = \sum_{n \in \mathbb{N}} \lambda_n T \ast \varphi_n \ast \psi_n \in L^\infty(G) \).

References