The Bochner-Minlos theorem

Jordan Bell

May 13, 2014

1 Introduction

We take $\mathbb{N}$ to be the set of positive integers. If A is a set and $n \in \mathbb{N}$, we typically deal with the product A^n as the set of functions $\{1, \ldots, n\} \to A$.

In this note I am following and greatly expanding the proof of the Bochner-Minlos theorem given by Barry Simon, *Functional Integration and Quantum Physics*, p. 11, Theorem 2.2.

2 The Kolmogorov extension theorem

If X is a topological space, and for $m \geq n$ the maps $\pi_{m,n} : X^m \to X^n$ are defined by

$$(\pi_{m,n}(x))(j) = x(j), \qquad j \in \{1, \ldots, n\},$$

then the spaces X^n and maps $\pi_{m,n}$ constitute a projective system, and its limit in the category of topological spaces is $X^{\mathbb{N}}$ with the maps $\pi_n : X^{\mathbb{N}} \to X^n$, where $X^{\mathbb{N}}$ has the initial topology for the family $\{\pi_n : n \in \mathbb{N}\}$ (namely, the product topology). We say that a function $f : X^{\mathbb{N}} \to \mathbb{R}$ *depends on only finitely many coordinates* if there is some n and some function $g : X^n \to \mathbb{R}$ such that $f = g \circ \pi_n$. We denote by $C_{\text{fin}}(X^{\mathbb{N}})$ the set of all continuous functions $X^{\mathbb{N}} \to \mathbb{R}$ that depend on only finitely many coordinates.

If (X, τ_X) is a noncompact locally compact Hausdorff space, write $\dot{X} = X \cup \{\infty\}$, and let τ be the collection of all subsets U of $\dot{X}$ such that either (i) $U \in \tau_X$ or (ii) $\infty \in U$ and $X \setminus U$ is compact in (X, τ_X). One proves[1] that $(\dot{X}, \tau)$ is a compact Hausdorff space and that the inclusion map $\iota : X \to \dot{X}$ is a homeomorphism $X \to \iota(X)$, where $\iota(X)$ has the subspace topology inherited from $\dot{X}$. Also, if $f \in C(X)$ then there is some $F \in C(\dot{X})$ whose restriction to X equals f if and only if there is some $g \in C_0(X)$ and some constant c such that $f = g + c$, in which case

$$F(x) = \begin{cases} f(x) & x \in X \\ c & x = \infty. \end{cases}$$

---

[1] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, second ed., p. 132, Proposition 4.36.

We call $\dot{X}$ the *one-point compactification of X*. For example, one checks that the one-point compactification of $\mathbb{R}^n$ is homeomorphic to S^n.

Theorem 1 (Kolmogorov)**.** Suppose that for each $n \in \mathbb{N}$, μ_n is a Borel probability measure on $\mathbb{R}^n$, and that for any n and any Borel set A in $\mathbb{R}^n$ we have

$$\mu_{n+1}(A \times \mathbb{R}) = \mu_n(A);$$

equivalently, $\pi_{m,n_*}\mu_m = \mu_n$ for $m \geq n$. There is then a Borel probability measure μ on $\mathbb{R}^{\mathbb{N}}$ such that for any n and any Borel set A in $\mathbb{R}^n$,

$$\mu\left(\left\{x \in \mathbb{R}^{\mathbb{N}} : \pi_n(x) \in A\right\}\right) = \mu_n(A);$$

equivalently, $\pi_{n_*}\mu = \mu_n$.

Proof. Let $X = \dot{\mathbb{R}}$, the one-point compactification of $\mathbb{R}$, and let $X^{\mathbb{N}}$ have the product topology, with which it is a compact Hausdorff space. For each n, if A is a Borel set in X^n, we define $\nu_n(A) = \mu_n(A \cap \mathbb{R}^n)$. This is a Borel measure on X^n. If $g \in C(X^n)$, $m \geq n$, and $h = g \circ \pi_{m,n}$, then

$$\int_{X^m} h d\nu_m = \int_{X^m} g \circ \pi_{m,n} d\nu_m = \int_{X^n} g d(\pi_{m,n_*}\nu_m) = \int_{X^n} g d\nu_n.$$

We define $L : C_{\text{fin}}(X^{\mathbb{N}}) \to \mathbb{R}$ in the following way. For $f \in C_{\text{fin}}(X^{\mathbb{N}})$, there is some n and some $g \in C(X^n)$ such that $f = g \circ \pi_n$. We define

$$L(f) = \int_{X^n} g d\nu_n.$$

If $h \in C(X^m)$ and $f = h \circ \pi_m$ with $m \geq n$, then $h = g \circ \pi_{m,n}$, giving

$$\int_{X^m} h d\nu_m = \int_{X^m} g d\nu_n,$$

so the definition of $L(f)$ makes sense.

It is straightforward to check that $C_{\text{fin}}(X^{\mathbb{N}})$ is an algebra over $\mathbb{R}$. The algebra $C_{\text{fin}}(X^{\mathbb{N}})$ separates points in $X^{\mathbb{N}}$, and the constant map $x \mapsto 1$ belongs to $C_{\text{fin}}(X^{\mathbb{N}})$; the latter fact tells us that there is no $x \in X^{\mathbb{N}}$ such that $f(x) = 0$ for all $f \in C_{\text{fin}}(X^{\mathbb{N}})$. Therefore, the Stone-Weierstrass theorem[2] tells us that $C_{\text{fin}}(X^{\mathbb{N}})$ is dense in the Banach algebra $C(X^{\mathbb{N}})$.

If $f \in C_{\text{fin}}(X^{\mathbb{N}})$ and $f = g \circ \pi_n$, then $\|f\|_\infty = \|g\|_\infty$, which is finite because X^n is compact. Because each ν_n is a probability measure,

$$|L(f)| = \left|\int_{X^n} g d\nu_n\right| \leq \|g\|_\infty = \|f\|_\infty,$$

showing that $L : C_{\text{fin}}(X^{\mathbb{N}}) \to \mathbb{R}$ is a bounded linear map, with $\|L\| = 1$. Because $C_{\text{fin}}(X^{\mathbb{N}})$ is dense in $C(X^{\mathbb{N}})$, there is a bounded linear map $\Lambda : C(X^{\mathbb{N}}) \to \mathbb{R}$

---

[2] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, second ed., p. 141, Corollary 4.50.

whose restriction to $C_{\text{fin}}(X^{\mathbb{N}})$ is equal to L, and that satisfies $\|\Lambda\| = \|L\| = 1$. Moreover, if $f \in C_{\text{fin}}(X^{\mathbb{N}})$ satisfies $f \geq 0$, then it is apparent that $L(f) \geq 0$; we say that L is a *positive linear functional*. The fact that L is a positive linear functional implies that Λ is too. Because $\Lambda : C(X^{\mathbb{N}}) \to \mathbb{R}$ is a positive linear functional with $\|\Lambda\| = 1$, by the Riesz-Markov theorem[3] there is a Borel probability measure ν on $X^{\mathbb{N}}$ such that

$$\Lambda f = \int_{X^{\mathbb{N}}} f\, d\nu, \qquad f \in C(X^{\mathbb{N}}).$$

If A is a Borel set in $\mathbb{R}^{\mathbb{N}}$ with the product topology, define $\mu(A) = \nu(A)$. μ is a Borel probability measure on $\mathbb{R}^{\mathbb{N}}$.

Now that we have in our hands a Borel probability measure μ on $\mathbb{R}^{\mathbb{N}}$ it remains to verify that it does what we want it to do. □

3 Sequence spaces

For $x, y \in \mathbb{R}^{\mathbb{N}}$ and $m \in \mathbb{Z}$, we define

$$\langle x, y \rangle_m = \sum_{j=1}^{\infty} j^{2m} x(j) y(j),$$

and $\|x\|_m = \langle x, x \rangle_m^{1/2}$. We define $\mathfrak{S}_m$ to be the set of all those $x \in \mathbb{R}^{\mathbb{N}}$ for which $\|x\|_m < \infty$, and we take as granted that for each m, $\mathfrak{S}_m$ is a Hilbert space. For $m \geq n$, let $\iota_{m,n} : \mathfrak{S}_m \to \mathfrak{S}_n$ be the inclusion map. As

$$
\begin{aligned}
\|\iota_{m,n} x\|_n^2 &= \sum_{j=1}^{\infty} j^{2n} (\iota_{m,n} x)(j)^2 \\
&= \sum_{j=1}^{\infty} j^{2m} j^{2(n-m)} x(j)^2 \\
&\leq \sum_{j=1}^{\infty} j^{2m} x(j)^2,
\end{aligned}
$$

the map $\iota_{m,n}$ is a bounded operator. In fact, if $m > n$ we now demonstrate that $\iota_{m,n}$ is a Hilbert-Schmidt operator, and so compact, which is the conclusion of *Rellich's theorem*. For $i \in \mathbb{N}$, define $e_i \in \mathbb{R}^{\mathbb{N}}$ by

$$e_i(j) = j^{-m} \delta_{i,j}.$$

---

[3] Walter Rudin, *Real and Complex Analysis*, third ed., p. 40, Theorem 2.14.

These e_i are an orthonormal basis for $\mathfrak{S}_m$, and

$$
\begin{aligned}
\sum_{i=1}^{\infty} \|\iota_{m,n} e_i\|_n^2 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} j^{2n} (\iota_{m,n} e_i)(j)^2 \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} j^{2n} \left(j^{-m} \delta_{i,j} \right)^2 \\
&= \sum_{i=1}^{\infty} i^{2n} i^{-2m} \\
&= \sum_{i=1}^{\infty} i^{2(n-m)}.
\end{aligned}
$$

Because $m > n$, this last expression is $< \infty$. This shows that $\iota_{m,n} : \mathfrak{S}_m \to \mathfrak{S}_n$ is a Hilbert-Schmidt operator.

For $x \in \mathfrak{S}_m$ and $\lambda \in \mathfrak{S}_{-m}$, define

$$
\langle x, \lambda \rangle = \sum_{j=1}^{\infty} x(j) \lambda(j), \tag{1}
$$

and using the Cauchy-Schwarz inequality,

$$
\begin{aligned}
|\langle x, \lambda \rangle| &\leq \sum_{j=1}^{\infty} |x(j)| |\lambda(j)| \\
&= \sum_{j=1}^{\infty} j^m |x(j)| j^{-m} |\lambda(j)| \\
&\leq \left(\sum_{j=1}^{\infty} j^{2m} |x(j)|^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} j^{-2m} |\lambda(j)|^2 \right)^{1/2} \\
&= \|x\|_m \|\lambda\|_{-m}.
\end{aligned}
$$

$\mathfrak{S}_{-m}$ is thus the dual space of the Banach space $\mathfrak{S}_m$. That is, as a vector space $\mathfrak{S}_m^* = \mathfrak{S}_{-m}$, but we shall be interested in $\mathfrak{S}_m^*$ with the weak-* topology with which it is a locally convex space, rather than with the norm topology with which it is a Banach space.

Since $\iota_{m,n} : \mathfrak{S}_m \to \mathfrak{S}_n$ is a bounded linear map for $m \geq n$, the spaces $\mathfrak{S}_m$ and the maps $\iota_{m,n}$ are a projective system of Banach spaces, and this projective system has a limit $\mathfrak{S}$ in the category of locally convex spaces. This limit $\mathfrak{S}$ is a Fréchet space. The duals $\mathfrak{S}_m^*$ with the weak-* topology are locally convex spaces and constitute a direct system with the maps $(\iota_{m,n})^* : \mathfrak{S}_n^* \to \mathfrak{S}_m^*$, where $(\iota_{m,n})^*(\lambda) = \lambda \circ \iota_{m,n}$ for $\lambda \in \mathfrak{S}_n^*$. This direct system has a colimit in the category of locally convex spaces which is equal to $\mathfrak{S}^*$ with the weak-* topology.[4] As

---

[4] Paul Garrett, http://www.math.umn.edu/~garrett/m/fun/notes_2012-13/04_blevi_sobolev.pdf, p. 15.

sets,

$$\mathfrak{S} = \bigcap_{m \in \mathbb{Z}} \mathfrak{S}_m, \qquad \mathfrak{S}^* = \bigcup_{m \in \mathbb{Z}} \mathfrak{S}_m^* = \bigcup_{m \in \mathbb{Z}} \mathfrak{S}_m.$$

We also denote by $\langle \cdot, \cdot \rangle$ the dual pairing of $\mathfrak{S}$ and $\mathfrak{S}^*$: for $x \in \mathfrak{S}$ and $\lambda \in \mathfrak{S}^*$,

$$\lambda(x) = \langle x, \lambda \rangle = \sum_{j=1}^{\infty} x(j)\lambda(j).$$

For any $\lambda \in \mathfrak{S}^*$, there is some m for which $\lambda \in \mathfrak{S}_m^* = \mathfrak{S}_{-m}$. But if $x \in \mathfrak{S}$ then $x \in \mathfrak{S}_m$, and so this dual pairing coincides with (1).

4   Positive-definite functions

If X is a vector space and $\Phi : X \to \mathbb{C}$ is a function, we say that Φ is *positive-definite* if for any positive integer r and for any $x_1, \ldots, x_r \in X$ and $c_1, \ldots, c_r \in \mathbb{C}$, we have

$$\sum_{j,k=1}^{r} c_j \overline{c_k} \Phi(x_j - x_k) \geq 0.$$

If Φ is positive-definite, one proves that $\Phi(0) \geq 0$, $\Phi(-x) = \overline{\Phi(x)}$, and $|\Phi(x)| \leq \Phi(0)$.

If μ is a probability measure on $(\mathfrak{S}^*, \mathrm{Cyl}\,(\mathfrak{S}^*))$, we define the *Fourier transform* of μ to be the function $\hat{\mu} : \mathfrak{S} \to \mathbb{C}$ defined by

$$\hat{\mu}(x) = (\mathscr{F}\mu)(x) = \int_{\mathfrak{S}^*} \exp(-iL_x)d\mu, \qquad x \in \mathfrak{S};$$

because $L_x : \mathfrak{S}^* \to \mathbb{R}$ and μ is a probability measure, this integral is finite. Using the dominated convergence theorem, one checks that $\hat{\mu} : \mathfrak{S} \to \mathbb{C}$ is continuous. It is apparent that $\hat{\mu}(0) = 1$. If $x_1, \ldots, x_r \in \mathfrak{S}$ and $c_1, \ldots, c_r \in \mathbb{C}$, then

$$
\begin{aligned}
\sum_{j,k=1}^{r} c_j \overline{c_k} \hat{\mu}(x_j - x_k) &= \sum_{j,k=1}^{r} c_j \overline{c_k} \int_{\mathfrak{S}^*} \exp(-i\lambda(x_j - x_k))d\mu(\lambda) \\
&= \int_{\mathfrak{S}^*} \sum_{j,k=1}^{r} c_j \exp(-i\lambda x_j)\overline{c_k \exp(-i\lambda x_k)}d\mu(\lambda) \\
&= \int_{\mathfrak{S}^*} \left(\sum_{j=1}^{r} c_j \exp(-i\lambda x_j) \right) \overline{\left(\sum_{k=1}^{r} c_k \exp(-i\lambda x_k) \right)} d\mu(\lambda) \\
&= \int_{\mathfrak{S}^*} \left| \sum_{j=1}^{r} c_j \exp(-i\lambda x_j) \right|^2 d\mu(\lambda) \\
&\geq 0,
\end{aligned}
$$

so $\hat{\mu} : \mathfrak{S} \to \mathbb{C}$ is positive-definite.

5

5 Cylinder sigma-algebras

$\mathfrak{S}^*$ is a topological space and so has a Borel σ-algebra. We shall now define a σ-algebra on $\mathfrak{S}^*$, called the *cylinder σ-algebra of $\mathfrak{S}^*$* and denoted $\mathrm{Cyl}\,(\mathfrak{S}^*)$, that is strictly contained in the Borel σ-algebra of $\mathfrak{S}^*$. For $x \in \mathfrak{S}$, define $L_x : \mathfrak{S}^* \to \mathbb{R}$ by $L_x(\lambda) = \lambda(x) = \langle x, \lambda \rangle$. We define $\mathrm{Cyl}\,(\mathfrak{S}^*)$ to be the coarsest σ-algebra such that for each $x \in \mathfrak{S}$, the map $L_x : \mathfrak{S}^* \to \mathbb{R}$ is measurable, where $\mathbb{R}$ has the Borel σ-algebra. Since each L_x is continuous, L_x is measurable with respect to the Borel σ-algebra on $\mathfrak{S}^*$, so $\mathrm{Cyl}\,(\mathfrak{S}^*)$ is contained in the Borel σ-algebra of $\mathfrak{S}^*$; it is not obvious that the cylinder σ-algebra is strictly contained in the Borel σ-algebra. Unless we say otherwise, when we speak of measurable functions on $\mathfrak{S}^*$ or measures on $\mathfrak{S}^*$ we mean with respect to the cylinder σ-algebra.

6 Minlos's theorem

In the following theorem we obtain a Borel probability measure μ on $\mathbb{R}^{\mathbb{N}}$ with the product topology. We denote by $\mathfrak{B}$ the Borel σ-algebra of $\mathbb{R}^{\mathbb{N}}$. The collection $\mathfrak{B}_0 = \{B \cap \mathfrak{S} : B \in \mathfrak{B}\}$ is a σ-algebra on $\mathfrak{S}$. We assert that $\mathfrak{B}_0 \subseteq \mathrm{Cyl}\,(\mathfrak{S})$, and that $\mathrm{Cyl}\,(\mathfrak{S})$ does not contain the Borel σ-algbera of $\mathfrak{S}$, and thus that the restriction of μ to $\mathfrak{S}$ is not a Borel measure.

Theorem 2 (Minlos)**.** If $\Phi : \mathfrak{S} \to \mathbb{C}$ is positive-definite, continuous, and $\Phi(0) = 1$, then there is some probability measure μ on $(\mathfrak{S}^*, \mathrm{Cyl}\,(\mathfrak{S}^*))$ such that $\Phi = \hat{\mu}$.

Proof. For $M \geq N$, define $\pi_{M,N} : \mathbb{R}^M \to \mathbb{R}^N$ by

$$(\pi_{M,N}x)(j) = x(j), \qquad j \in \{1, \ldots, N\}.$$

The Banach spaces $\mathbb{R}^N$ and the maps $\pi_{M,N}$ constitute a projective system in the category of locally convex spaces, with the limit $\mathbb{R}^{\mathbb{N}}$, which is thus a Fréchet space, with the maps $\pi_N : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^N$,

$$(\pi_N x)(j) = x(j), \qquad j \in \{1, \ldots, N\}.$$

The dual maps $\pi_{M,N}^* : (\mathbb{R}^N)^* \to (\mathbb{R}^M)^*$ are defined for $\lambda \in (\mathbb{R}^N)^*$ by

$$(\pi_{M,N}^*)(\lambda) = \lambda \circ \pi_{M,N}.$$

$(\mathbb{R}^N)^* = \mathbb{R}^N$ and the maps $\pi_{M,N}^*$ constitute a direct system, and their colimit in the category of locally convex spaces is denoted

$$\mathbb{R}^{\infty} = \bigoplus_{N \in \mathbb{N}} \mathbb{R}.$$

$\mathbb{R}^{\infty} = (\mathbb{R}^{\mathbb{N}})^*$, and the maps $\pi_N^* : (\mathbb{R}^N)^* \to \mathbb{R}^{\infty}$ satisfy

$$\pi_N^*(\lambda) = \lambda \circ \pi_N.$$

The function $\Phi_N = \Phi \circ \pi_N^* : (\mathbb{R}^N)^* \to \mathbb{C}$ satisfies, for $\lambda_1, \ldots, \lambda_r \in (\mathbb{R}^N)^*$ and $c_1, \ldots, c_r \in \mathbb{C}$,

$$\sum_{j,k=1}^{r} c_j \overline{c_k} (\Phi \circ \pi_N^*)(\lambda_j - \lambda_k) = \sum_{j,k=1}^{r} c_j \overline{c_k} \Phi(\lambda_j \circ \pi_N - \lambda_k \circ \pi_N) \geq 0,$$

because $\lambda_1 \circ \pi_N, \ldots, \lambda_r \circ \pi_N \in \mathfrak{S}$ and $\Phi : \mathfrak{S} \to \mathbb{C}$ is positive-definite. Furthermore, $(\Phi \circ \pi_N^*)(0) = \Phi(0) = 1$, and $\Phi_N = \Phi \circ \pi_N^*$ is a composition of continuous functions so is itself continuous. Therefore, for each $N \in \mathbb{N}$ we have by Bochner's theorem that there is one and only Borel probability measure μ_N on $\mathbb{R}^N$ that satisfies $\Phi_N = \hat{\mu}_N$. If $M \geq N$, for $\xi \in (\mathbb{R}^N)^* = \mathbb{R}^N$ we have

$$
\begin{aligned}
\mathscr{F}(\pi_{M,N_*}\mu_M)(\xi) &= \int_{\mathbb{R}^N} e^{-i\xi \cdot x} d(\pi_{M,N_*}\mu_M)(x) \\
&= \int_{\mathbb{R}^M} e^{-i\xi \cdot \pi_{M,N}(x)} d\mu_M(x) \\
&= \int_{\mathbb{R}^M} e^{-i\pi_{M,N}^*(\xi) \cdot x} d\mu_M(x) \\
&= \hat{\mu}_M(\pi_{M,N}^*(\xi)) \\
&= \Phi_M(\pi_{M,N}^*(\xi)) \\
&= (\Phi_M \circ \pi_{M,N}^*)(\xi) \\
&= \Phi_N(\xi) \\
&= \hat{\mu}_N(\xi).
\end{aligned}
$$

Since $\mathscr{F}(\pi_{M,N_*}\mu_M) = \mathscr{F}(\mu_N)$, it follows that $\pi_{M,N_*}\mu_M = \mu_N$. Therefore, the Borel probability measures μ_N satisfy the conditions of the Kolmogorov extension theorem, and so there is some Borel probability measure μ on $\mathbb{R}^{\mathbb{N}}$ such that $\pi_{N_*}\mu = \mu_N$. Now that we have our hands on the measure μ, one must prove that $\hat{\mu} = \Phi$.

Supposing that we have proved $\hat{\mu} = \Phi$, we now prove that $\mu(\mathfrak{S}^*) = 1$. Let $\epsilon > 0$. $\Phi : \mathfrak{S} \to \mathbb{C}$ is continuous at 0 and $\Phi(0) = 1$, and as $\mathfrak{S}$ has the locally convex topology induced by the family of seminorms $\|\cdot\|_m$, there is some $m \in \mathbb{Z}$ and some $\delta > 0$ such that $\|y\|_m \leq \delta$ implies that $|\Phi(y) - 1| \leq \epsilon$. Suppose that $y \in \mathfrak{S}$. On the one hand, if $\|y\|_m^2 \leq \delta^2$, then

$$1 - \operatorname{Re} \Phi(y) \leq |\operatorname{Re} \Phi(y) - 1| \leq |\Phi(y) - 1| \leq \epsilon.$$

On the other hand, if $\|y\|_m^2 > \delta^2$, using $|\Phi(y)| \leq \Phi(0) = 1$ and so $|\operatorname{Re} \Phi(y)| \leq |\Phi(y)| \leq 1$, we get

$$\operatorname{Re} \Phi(y) \geq -1 > 1 - 2\delta^{-2} \|y\|_m^2.$$

Therefore, for any $y \in \mathfrak{S}$,

$$\operatorname{Re} \Phi(y) \geq 1 - \epsilon - 2\delta^{-2} \|y\|_m^2.$$

Then, for $y \in \mathbb{R}^N$ we have

$$\operatorname{Re} \Phi(\pi_N^*(y)) \geq 1 - \epsilon - 2\delta^{-2} \|\pi_N^*(y)\|_m^2. \tag{2}$$

Let $\alpha > 0$, let $q_k = k^{-2m-2}$, and let $\sigma_{\alpha,N}$ be the measure on $\mathbb{R}^N$ with

$$d\sigma_{\alpha,N}(y) = \prod_{k=1}^{N} (2\pi\alpha q_k)^{-1/2} \exp\left(-\frac{y_k^2}{2\alpha q_k}\right) dy_k.$$

Using $\int_{\mathbb{R}} \exp(-x^2)dx = \sqrt{\pi}$, it is straightforward to check that $\sigma_{\alpha,N}$ is a probability measure on $\mathbb{R}^N$. Furthermore, we calculate, using respectively $\int_{\mathbb{R}} x\exp(-x^2)dx = 0$, $\int_{\mathbb{R}} \exp(-x^2)dx = \sqrt{\pi}$, and $\int_{\mathbb{R}} x^2 \exp(-x^2)dx = \frac{\sqrt{\pi}}{2}$,

$$
\begin{aligned}
\int_{\mathbb{R}^N} y_i y_j d\sigma_{\alpha,N}(y) &= \delta_{i,j} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} y_j^2 \prod_{k=1}^{N} (2\pi\alpha q_k)^{-1/2} \exp\left(-\frac{y_k^2}{2\alpha q_k}\right) dy_k \\
&= \delta_{i,j} \left(\prod_{k\neq j} \int_{\mathbb{R}} (2\pi\alpha q_k)^{-1/2} \exp\left(-\frac{y_k^2}{2\alpha q_k}\right) dy_k \right) \\
&\quad \cdot \int_{\mathbb{R}} y_j^2 (2\pi\alpha q_j)^{-1/2} \exp\left(-\frac{y_j^2}{2\alpha q_j}\right) dy_j \\
&= \delta_{i,j} \int_{\mathbb{R}} y_j^2 (2\pi\alpha q_j)^{-1/2} \exp\left(-\frac{y_j^2}{2\alpha q_j}\right) dy_j \\
&= \delta_{i,j}\alpha q_j.
\end{aligned}
$$

and for $x \in \mathbb{R}^N$, taking as known the Fourier transform of a Gaussian function,

$$
\begin{aligned}
\int_{\mathbb{R}^N} e^{-ix\cdot y} d\sigma_{\alpha,N}(y) &= \prod_{k=1}^{N} \int_{\mathbb{R}} e^{-ix_k y_k} (2\pi\alpha q_k)^{-1/2} \exp\left(-\frac{y_k^2}{2\alpha q_k}\right) dy_k \\
&= \prod_{k=1}^{N} \exp\left(-\frac{\alpha q_k x_k^2}{2}\right) \\
&= \exp\left(-\frac{\alpha}{2} \sum_{k=1}^{N} q_k x_k^2\right).
\end{aligned}
$$

Then, using $\Phi = \hat{\mu}$, the integral of the left-hand side of (2) over $\mathbb{R}^N$ with respect to $\sigma_{\alpha,N}$ is

$$
\begin{aligned}
\text{Re} \int_{\mathbb{R}^N} \Phi(\pi_N^*(y)) d\sigma_{\alpha,N}(y) &= \text{Re} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \exp(-i\langle \pi_N^*(y), x\rangle) d\mu(x) d\sigma_{\alpha,N}(y) \\
&= \text{Re} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \exp(-iy \cdot \pi_N(x)) d\sigma_{\alpha,N}(y) d\mu(x) \\
&= \text{Re} \int_{\mathbb{R}^N} \exp\left(-\frac{\alpha}{2} \sum_{k=1}^{N} q_k x_k^2\right) d\mu(x) \\
&= \int_{\mathbb{R}^N} \exp\left(-\frac{\alpha}{2} \sum_{k=1}^{N} q_k x_k^2\right) d\mu(x).
\end{aligned}
$$

8

The integral of the right-hand side of (2) over $\mathbb{R}^N$ with respect to $\sigma_{\alpha,N}$ is

$$
\begin{aligned}
\int_{\mathbb{R}^N} (1 - \epsilon - 2\delta^{-2} \|\pi_N^*(y)\|_m^2) d\sigma_{\alpha,N}(y) &= 1 - \epsilon - 2\delta^{-2} \int_{\mathbb{R}^N} \sum_{k=1}^N k^{2m} y_k^2 d\sigma_{\alpha,N}(y) \\
&= 1 - \epsilon - 2\delta^{-2} \sum_{k=1}^N k^{2m} \alpha q_k^2 \\
&= 1 - \epsilon - 2\delta^{-2} \alpha \sum_{k=1}^N k^{-2}.
\end{aligned}
$$

Combining these, (2) is

$$
\int_{\mathbb{R}^{\mathbb{N}}} \exp\left(-\frac{\alpha}{2} \sum_{k=1}^N q_k x_k^2 \right) d\mu(x) \geq 1 - \epsilon - 2\delta^{-2} \alpha \sum_{k=1}^N k^{-2}.
$$

Taking $N \to \infty$,

$$
\int_{\mathbb{R}^{\mathbb{N}}} \exp\left(-\frac{\alpha}{2} \sum_{k=1}^\infty q_k x_k^2 \right) d\mu(x) \geq 1 - \epsilon - 2\delta^{-2} \alpha \cdot \zeta(2). \tag{3}
$$

But

$$
\begin{aligned}
\lim_{\alpha \to 0^+} \exp\left(-\frac{\alpha}{2} \sum_{k=1}^N q_k x_k^2 \right) &= \lim_{\alpha \to 0^+} \exp\left(-\frac{\alpha}{2} \sum_{k=1}^N k^{2(-m-1)} x_k^2 \right) \\
&= \begin{cases} 1 & x \in \mathfrak{S}_{-m-1} \\ 0 & x \notin \mathfrak{S}_{-m-1}, \end{cases}
\end{aligned}
$$

so taking $\alpha \to 0^+$, (3) yields

$$
\int_{\mathbb{R}^{\mathbb{N}}} \chi_{\mathfrak{S}_{-m-1}}(x) d\mu(x) \geq 1 - \epsilon,
$$

i.e. $\mu(\mathfrak{S}_{-m-1}) \geq 1 - \epsilon$, and $\mu(\mathfrak{S}^*) \geq \mu(\mathfrak{S}_{-m-1})$. That is, we have proved that for any $\epsilon > 0$ there is some $m \in \mathbb{Z}$ such that

$$
\mu(\mathfrak{S}^*) \geq \mu(\mathfrak{S}_{-m-1}) \geq 1 - \epsilon,
$$

which shows that $\mu(\mathfrak{S}^*) = 1$, i.e. that μ is a probability measure. $\square$

7 References

There don't seem to be many detailed presentations of the Bochner-Minlos theorem, or Minlos's theorem, in the literature. Cf. Sazonov's theorem. Doing some searching through books, the following look like they have enough details so that they are not worse than nothing, which is more than can always be said:

- Helge Holden, Bernt Øksendal, Jan Ubøe and Tusheng Zhang, *Stochastic Partial Differential Equations: A Modeling, White Noise Functional Approach*, second ed., Appendix A, pp. 257ff.

- Takeyuki Hida, *Brownian Motion*, pp. 116ff., §3.2

- I. M. Gelfand and N. Ya. Vilenkin, *Generalized Functions. Volume 4: Applications of Harmonic Analysis*

- V. I. Bogachev, *Measure Theory*, vol. II, p. 124, Theorem 7.13.7

- A. V. Skorokhod, *Basic Principles and Applications of Probability Theory*, p. 51, §2.4.4

- N. Bourbaki, *Integration II: Chapters 7–9*, p. IX.100, §6, Theorem 4