Limiting Probability Measures

IRFAN ALAM

Abstract: The coordinates along any fixed direction(s) of points on the sphere $S^{n-1}(\sqrt{n})$ roughly follow a standard Gaussian distribution as n approaches infinity. We revisit this classical result from a nonstandard analysis perspective, providing a new proof by working with hyperfinite dimensional spheres. We also set up a nonstandard theory for the asymptotic behavior of integrals over varying domains in general. We obtain a new proof of the Riemann–Lebesgue lemma as a by-product of this theory. We finally show that for any function $f\colon \mathbb{R}^k \to \mathbb{R}$ with finite Gaussian moment of an order larger than one, its expectation is given by a Loeb integral over a hyperfinite dimensional sphere. Some useful inequalities between high-dimensional spherical means of f and its Gaussian mean are obtained in order to complete the above proof.

1 Introduction

Gaussian measures have been mathematically connected with the uniform surface area measures on high-dimensional spheres since at least the time of Poincaré, who observed in [16] that if n real numbers are randomly chosen under the constraint that their sum of squares equals n (this is equivalent to choosing a random vector on $S^{n-1}(\sqrt{n})$, the sphere in $\mathbb{R}^n$ of radius $\sqrt{n}$ centered at the origin), then as $n \to \infty$ the probability distribution of the first number converges to that of a standard Gaussian random variable (that is, with zero mean and covariance equaling one). Considering works on the kinetic theory of gases in Physics, this connection goes back another century. (We briefly outline this connection with Physics in Appendix A.) We will attribute this result to Poincaré for having made the connection explicit.

For any sphere S centered at the origin in a Euclidean space, there is a unique orthogonal transformation invariant probability measure $\bar{\sigma}_S$. (We will omit the subscript when the

sphere under consideration is clear from context.) For each $k \in \mathbb{N}$ and $n \in \mathbb{N}_{\geq k}$, let $\pi_k^{(n)} \colon \mathbb{R}^n \to \mathbb{R}^k$ denote the projection on to the first k coordinates under the standard basis. (We will omit the superscript when the dimension is clear from context.) For a Borel set $B \subseteq \mathbb{R}^k$, we write:

$$\bar{\sigma}_{S^{n-1}(\sqrt{n})}(B) := \bar{\sigma}_{S^{n-1}(\sqrt{n})}[S^{n-1}(\sqrt{n}) \cap (\pi_k^{(n)})^{-1}(B)]$$

In the same spirit, we identify each measurable function $f \colon \mathbb{R}^k \to \mathbb{R}$ with a function on $\mathbb{R}^n$ by composing it with the projection $\pi_k^{(n)}$. This allows us to talk about integrals of such an f over domains in $\mathbb{R}^n$ for $n \in \mathbb{N}_{\geq k}$.

We let $\mu_{(k)}$ denote the standard Gaussian measure on $\mathbb{R}^k$ (again, omitting the subscript when the dimension is clear). With these conventions, we may write Poincaré's observation succinctly in terms of the following limit:

$$(1\text{--}1) \qquad \lim_{n \to \infty} \bar{\sigma}_{S^{n-1}(\sqrt{n})}(B) = \mu(B) \text{ for all Borel sets } B \subseteq \mathbb{R}$$

By standard measure theory, it is not difficult to see that the above can be rephrased in a more general form as follows. (As discussed above, the integral on the left side of (1–2) will be understood as that of the function $f \circ \pi_k^{(n)}$ for all $n \in \mathbb{N}_{>k}$.)

Theorem 1.1 (Poincaré [16]) *For all bounded measurable functions $f \colon \mathbb{R}^k \to \mathbb{R}$, we have:*

$$(1\text{--}2) \qquad \lim_{n \to \infty} \int_{S^{n-1}(\sqrt{n})} f d\bar{\sigma} = \int_{\mathbb{R}^k} f d\mu$$

Similar ideas were later used by Lévy [9] to do infinite dimensional analysis, and then by Wiener [21] to construct Brownian motion. McKean [13] surveyed most of the relevant work from that period. Cutland and Ng explored these themes using nonstandard analysis (which provides the language of hyperfinite dimensional spheres) in [5]. They gave a new construction of the Wiener measure using the nonstandard machinery.

The current paper may be considered a sequel to [5] in some sense. Indeed one of our aims is to view the above classical result (Theorem 1.1) as a statement about Loeb integrals on hyperfinite dimensional spheres, and obtain the same result for a larger class of functions. Toward that end, we give a new nonstandard proof of Poincaré's theorem in Section 2.3. A novel feature of this proof is that it does not require any explicit integral calculations – it follows from straightforward applications of the weak law of large numbers and the definition of the uniform surface area measure on a sphere as a pushforward of a Gaussian measure. In Section 3, we also establish a nonstandard

approach of extending such results from bounded measurable functions to other classes of functions. The general framework described in Sections 2 and 3 may be thought of as an invitation to apply nonstandard analysis to other asymptotic problems in probability and measure theory. One such application is carried out in Alam [1] to generalize recent works of Sengupta [20] and Peterson–Sengupta [15] that connect Gaussian Radon transforms with limiting spherical integrals. This generalization is the topic of one of the author's ongoing projects.

We also give a classical standard proof of Theorem 1.1 in Section 4.1 – it follows by dominated convergence theorem once the integral over the sphere is "disintegrated" properly (for example, using Sengupta [20, Proposition 4.1]). As pointed out in Remark 4.1, this proof of Theorem 1.1 does not immediately generalize to work for an arbitrary μ–integrable function. The nonstandard framework of Section 3 allows one to get conditions (see Theorems 3.1 and 3.4) under which a result of the type of Theorem 1.1 for bounded measurable functions (over general domains) can be extended to unbounded functions. Though we do not use this terminology, the framework in Section 3 is similar to the framework of *graded probability spaces*, as in Hoover [7] and Keisler [8].

Aside from its application to spherical integrals, the approach of Section 4 is potentially useful in many other situations in which limits of integrals may be studied. A new proof of the Riemann-Lebesgue Lemma is provided (see Theorem 3.5) as an example of its use. Finally, in order to verify the sufficient conditions from Section 4 in the case of spherical integrals, we also prove some inequalities between spherical means and $L^p(\mathbb{R}^k, \mu)$ norms of functions on $\mathbb{R}^k$ (see Theorem 4.6 and Corollary 4.7). Thus, the main results of this paper can be divided into three types:

- Results viewing the limiting behavior of integrals over varying abstract domains as a single integral over a nonstandard domain.
- Inequalities between spherical integrals and Gaussian integrals.
- Applications of the results of the above types to systematically generalize Theorem 1.1 on limiting spherical integrals to a bigger class of functions.

1.1 Summary and motivation of our key results

Recall that for a Borel measurable function $f : \mathbb{R}^k \to \mathbb{R}$, we are interested in

$$\lim_{n \to \infty} \int_{S^{n-1}(\sqrt{n})} f(x_1, \ldots, x_k) d\bar{\sigma}(x_1, \ldots, x_n)$$

where we view f as a function on $\mathbb{R}^n$ by first projecting the input into the first k coordinates. Assuming Theorem 1.1, if f is bounded, then we know from (1–2) that this

limit is equal to the expected value of f with respect to the standard Gaussian measure μ on $\mathbb{R}^k$. Since we are assuming the limiting result (1–2) for bounded functions, we have (using $\mathbb{1}_B$ to denote the indicator function of a set B) the following for a possibly unbounded Borel measurable function $f : \mathbb{R}^k \to \mathbb{R}$:

$$(1\text{–}3) \qquad \lim_{m\to\infty} \lim_{n\to\infty} \int_{S^{n-1}(\sqrt{n})} f\mathbb{1}_{|f|\le m}d\bar{\sigma} = \lim_{m\to\infty} \int_{\mathbb{R}^k} f\mathbb{1}_{|f|\le m}d\mu = \int_{\mathbb{R}^k} fd\mu$$

However, we wanted to find $\lim_{n\to\infty} \int_{S^{n-1}(\sqrt{n})} fd\bar{\sigma}$, which (assuming that f is integrable over $S^{n-1}(\sqrt{n})$ for large $n \in \mathbb{N}$) is the same as the following:

$$\lim_{n\to\infty} \lim_{m\to\infty} \int_{S^{n-1}(\sqrt{n})} f\mathbb{1}_{|f|\le m}d\bar{\sigma}$$

Thus, in order to go from a result on bounded functions to a result on more general functions, we want to be able to switch the order of limits in (1–3). However, there is no general theory of switching double limits.

From the point of view of nonstandard analysis, the situation is simpler since the large–n behavior of any sequence is captured in the values attained by the nonstandard extension of that sequence at hyperfinite indices. For a hyperfinite $N > \mathbb{N}$, the sphere $S^{N-1}(\sqrt{N})$ inherits a finitely additive internal probability measure from the sequence $(S^{n-1}(\sqrt{n}), \bar{\sigma}_{S^{n-1}(\sqrt{n})})_{n\in\mathbb{N}}$. The Nth term in the nonstandard extension of the sequence $\left(\int_{S^{n-1}(\sqrt{n})} fd\bar{\sigma} \right)_{n\in\mathbb{N}}$ is then the *–integral of $^*f$ with respect to this internal measure. It turns out that the limiting integral for a general measurable function $f \colon \mathbb{R}^k \to \mathbb{R}$ exists (knowing that it exists and is equal to the Gaussian mean for bounded measurable functions) if $^*f$ is S–integrable over $S^{N-1}(\sqrt{N})$. In a more abstract setting, Theorem 3.1 essentially tells us that we can switch these limits if the tail double-limit $\lim_{m\to\infty} \lim_{n\to\infty} \int_{S^{n-1}(\sqrt{n})} |f| \, \mathbb{1}_{|f|>m}d\bar{\sigma}$ is zero. This condition of the tail double-limit being zero is just a standard reformulation of one of the equivalent conditions that ensure the S–integrability of $^*f$ over $S^{N-1}(\sqrt{N})$ (see (2) of Theorem 1.5).

A partial converse of the above result holds for nonnegative functions, which is covered in Theorem 3.4. Thus the set of all nonnegative functions for which the limit of spherical integrals is equal to the Gaussian integral is precisely the set of nonnegative functions for which the above tail double-limit is zero. While Theorems 3.1 and 3.4 come out of nonstandard measure theoretic considerations, we paraphrase a standard version for convenience as follows:

Theorem 1.2 *Let $(E, \mathcal{E})$ be a measure space. Let $k \in \mathbb{N}$, and for each $n \in \mathbb{N}_{>k}$ suppose $\Omega_n \subseteq E^{n'}$ for some $n' \in \mathbb{N}_{>k}$. Suppose that $\mathcal{F}_n$, the given sigma-algebra*

on Ω_n, is induced by the product sigma-algebra $\mathcal{E}_{n'}$ on $E^{n'}$. Let $(\Omega_n, \mathcal{F}_n, \nu_n)$ be a sequence of Borel probability spaces. Let $\mathbb{P}$ be a probability measure on $(E^k, \mathcal{E}_k)$ such that $\lim_{n\to\infty} \nu_n(B) = \mathbb{P}(B)$ for any $B \in \mathcal{E}_k$. Then for any function $f: E^k \to \mathbb{R}$, (1) implies (2):

(1) *The function f is integrable on (Ω_n, ν_n) for all large $n \in \mathbb{N}$; furthermore:*

$$\lim_{m\to\infty} \lim_{n\to\infty} \int_{\Omega_n \cap \{|f| \geq m\}} |f| \, d\nu_n = 0$$

(2) *The function f is $\mathbb{P}$–integrable and $\lim_{n\to\infty} \int_{\Omega_n} f d\nu_n = \int_{E^k} f d\mathbb{P}$.*

Furthermore, if f is assumed to be nonnegative, then the conditions (1) and (2) are equivalent.

The above theorem can also be interpreted more classically as a statement involving uniform integrability. While we do not focus on this aspect, it is interesting to emphasize that the nonstandard arguments using S–integrability thus encompass standard uniform integrability techniques.

In the case when Ω_n are the spheres $S^{n-1}(\sqrt{n})$, we verify the above double limit condition for all functions on $\mathbb{R}^k$ with a finite $(1 + \epsilon)$–Gaussian moment, where ϵ is any positive real number. This allows us to extend the result in Theorem 1.1 to all such functions (see Theorem 4.8). The main step in this verification is an inequality (see Theorem 4.6 and Corollary 4.7) between sufficiently high-dimensional spherical means and $L^p(\mathbb{R}^k, \mu)$ norms of functions on $\mathbb{R}^k$, which we summarize as follows:

Theorem 1.3 *For each $p \in \mathbb{R}_{>1}$, there is a constant $C_p \in \mathbb{R}_{>0}$ such that the following holds:*

$$(1\text{–}4) \qquad \int_{S^{n-1}(\sqrt{n})} |g| \, d\bar{\sigma}_n \leq C_p [\mathbb{E}_\mu(|g|^p)]^{\frac{1}{p}} \text{ for all } g \in L^p(\mathbb{R}^k, \mu) \text{ and } n \in \mathbb{N}_{>4(k+1)q}$$

where $q \in \mathbb{R}_{>0}$ is such that $1/p + 1/q = 1$.

Furthermore, we may replace the constant C_p in the above inequality by a real number as close to 1 as desired if n is taken large enough. (This large n depends only on $p \in \mathbb{R}_{>1}$ and the desired distance of the constant from 1.)

1.2 List of notation recurring throughout the paper

We use the superstructure approach, as in Albeverio et al [2]. Let X be a set of urelements and let $V(X)$ be its superstructure. We fix a sufficiently saturated nonstandard

extension of $V(X)$. In this paper, we work with measures defined on a sequence of measure spaces, and want to construct a natural Loeb measure on any element in the nonstandard extension of such a sequence. One issue in doing so could be that the measure spaces might not all lie in a single iterated power set over X (in which case, we cannot think of the sequence of measure spaces as an element of $V(X)$). In particular, this would be an issue if our measure spaces were the Borel spaces $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ and X was the set of real numbers. To get around this difficulty, we take a set X that contains (copies of) $\mathbb{R}^n$ for each $n \in \mathbb{N}$.

The set of finite nonstandard real numbers will be denoted by $^*\mathbb{R}_{\text{fin}}$. Any $N \in {}^*\mathbb{N}\backslash\mathbb{N}$ will be called hyperfinite, which we express by writing $N > \mathbb{N}$. Whenever it exists, we denote the standard part of a nonstandard element x by $\mathbf{st}(x)$. For $x, y \in {}^*\mathbb{R}$, we also write $x \approx y$ to mean that $\mathbf{st}(x - y) = 0$.

To establish notation as well as frequently referenced results, we now give a brief overview of what we need from basic Loeb measure theory. Let Ω be an internal set and let $\mathcal{F}$ be an internal algebra on Ω. Given a finite, finitely additive internal measure $\mathbb{P}$ (that is, $\mathbb{P}\colon \mathcal{F} \to {}^*\mathbb{R}_{\geq 0}$ satisfies $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) < \infty$, and $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ whenever $A \cap B = \emptyset$), the map $\mathbf{st}(\mathbb{P})\colon \mathcal{F} \to \mathbb{R}_{\geq 0}$ is an ordinary finite, finitely additive measure. By saturation, it follows that $\mathbf{st}(\mathbb{P})$ satisfies the premises of Carathéodory Extension Theorem. By that theorem, it extends to a unique measure on $\sigma(\mathcal{F})$ (the smallest sigma-algebra containing $\mathcal{F}$), whose completion is called the **Loeb measure** of $\mathbb{P}$. The corresponding complete measure space $(\Omega, L(\mathcal{F}), L\mathbb{P})$ is called the **Loeb space** of $(\Omega, \mathcal{F}, \mathbb{P})$. We will use the following simplification of Ross [17, Theorem 5.1, page 105] extensively.

Proposition 1.4 *Let $(\Omega, L(\mathcal{F}), L\mathbb{P})$ be the Loeb probability space of $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose $F\colon \Omega \to {}^*\mathbb{R}$ is an internal function that is measurable in the sense that $F^{-1}(^*B) \in \mathcal{F}$ for all $B \in \mathcal{B}(\mathbb{R})$ (where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$). If $F(\omega) \in {}^*\mathbb{R}_{fin}$ for $L\mathbb{P}$–almost all $\omega \in \Omega$, then $\mathbf{st}(F)$ is Loeb measurable (that is, measurable as a map from $(\Omega, L(\mathcal{F}))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$).*

For a standard measure space $(\Omega, \mathcal{F})$, let $\mathbf{Prob}(\Omega, \mathcal{F})$ be the set of probability measures on $(\Omega, \mathcal{F})$. If $\mathcal{C} \in V(X)$ is a collection of measure spaces, then $\mathbf{Prob}(\mathcal{C})$ denotes the set of all probability measures on elements in $\mathcal{C}$. Any element in $^*\mathbf{Prob}(\mathcal{C})$ is a finitely additive internal probability on an internal measure space. For any $\mathbb{P} \in \mathbf{Prob}(\mathcal{C})$, there is an *integral operator* that takes certain functions (those in the space $L^1(\mathbb{P})$ of integrable real-valued functions on the underlying sample space of $\mathbb{P}$) to their integrals with respect to $\mathbb{P}$. Thus if $(\Omega, \mathcal{F}, \mathbb{P}) \in {}^*\mathcal{C}$ is an internal probability space, we also have the associated space $^*L^1(\Omega, \mathbb{P})$ of *–integrable functions.

For any $*$–integrable $F\colon \Omega \to {}^*\mathbb{R}$, one then has ${}^*\int_\Omega F d\mathbb{P} \in {}^*\mathbb{R}$, which we call the $*$-*integral* of F over $(\Omega, \mathbb{P})$. This $*$–integral on ${}^*L^1(\Omega)$ inherits many properties (an important one being linearity) from the ordinary integral by transfer. If F is finite almost surely with respect to the corresponding Loeb measure, then $\mathbf{st}(F)$ is Loeb measurable by Proposition 1.4. In that case, it is interesting to study the relation between the $*$–integral of F and the Loeb integral of $\mathbf{st}({}^*F)$. The following result covers this for a useful class of functions (see Ross [17, Theorem 6.2, page 110] for a proof).

Theorem 1.5 *Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is an internal probability space and $F \in {}^*L^1(\Omega)$ is such that $L\mathbb{P}(F \in {}^*\mathbb{R}_{fin}) = 1$. Then the following are equivalent:*

(1) ${}^*\int_\Omega |F|\, d\mathbb{P} \in {}^*\mathbb{R}_{fin}$, *and the following holds:*

$$\mathbf{st}\left({}^*\!\int_\Omega |F|\, d\mathbb{P}\right) = \lim_{m\to\infty} \mathbf{st}\left({}^*\!\int_\Omega |F|\, \mathbb{1}_{\{|F|\leq m\}} d\mathbb{P}\right)$$

(2) For every $M > \mathbb{N}$, we have $\mathbf{st}\left({}^*\int_\Omega |F|\, \mathbb{1}_{\{|F|>M\}} d\mathbb{P}\right) = 0$.

(3) ${}^*\int_\Omega |F|\, d\mathbb{P} \in {}^*\mathbb{R}_{fin}$, *and for any $B \in \mathcal{F}$ we have:*

$$\mathbb{P}(B) \approx 0 \Rightarrow {}^*\!\int_\Omega |F|\, \mathbb{1}_B d\mathbb{P} \approx 0$$

(4) $\mathbf{st}(F)$ is Loeb integrable, and $\mathbf{st}\left({}^*\int_\Omega |F|\, d\mathbb{P}\right) = \int_\Omega |\mathbf{st}(F)|\, dL\mathbb{P}$.

A function satisfying the equivalent conditions in Theorem 1.5 is called S-*integrable* on $(\Omega, \mathcal{F}, \mathbb{P})$. The notion of S–integrability, first developed by Anderson [3], is one of the most ubiquitous concepts in nonstandard measure theory. Given a Loeb measurable function $f\colon \Omega \to \mathbb{R}$, a natural question is when does it occur as the standard part of an internal function. An internal measurable function $F\colon \Omega \to {}^*\mathbb{R}$ is called a *lifting* of a Loeb measurable function f if $L\mathbb{P}(\mathbf{st}(F) = f) = 1$. The following theorem shows that $*$–integrable functions can be characterized as those possessing S–integrable liftings (see Ross [17, Theorem 6.4, page 111] for a proof).

Theorem 1.6 *Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is an internal probability space, and let $(\Omega, L(\mathcal{F}), L(\mathbb{P}))$ be the associated Loeb space. Suppose $f\colon \Omega \to \mathbb{R}$ is Loeb measurable. Then f is Loeb integrable if and only if it has an S–integrable lifting.*

1.3 Structure of the paper

Section 2 contains a nonstandard proof of Theorem 1.1 (carried out in Section 2.3), which is prefaced by some basic nonstandard measure theory that we will use and a discussion on spherical measures (along with their nonstandard counterparts).

Section 3 continues the theme by studying sequences of abstract measure spaces for which a result of the type of Poincaré is known. It gives conditions under which such results hold for more general functions, allowing us to express the limiting behavior of certain integrals by a Loeb integral on a single limiting measure space. An application that yields a new proof of the Riemann-Lebesgue lemma is carried out in Theorem 3.5.

In Section 4, we apply the results of Section 3 to the case of high-dimensional spheres, and obtain a generalization of the classical result on limits of spherical integrals to a large class of Gaussian integrable functions (see Theorem 4.8). Toward that end, we also obtain some useful inequalities between high-dimensional spherical means and Gaussian means (see Theorem 4.6 and Corollary 4.7).

2 A Quick nonstandard proof of Poincaré's theorem

Using the nonstandard characterization of limit points, Poincaré's theorem is essentially a statement about the Loeb measure of the fiber (in the hyperfinite-dimensional sphere $S^{N-1}(\sqrt{N})$ for $N > \mathbb{N}$) of a finite-dimensional set equaling its Gaussian measure. In a more general setting, we analyze this type of phenomenon in the next subsection. These results are routine but essential in setting up later proofs.

2.1 When a Loeb measure matches up with a standard measure on a subspace

In what follows, there will be a measure space $(E, \mathcal{E})$ such that we assume X to contain copies of E^n for all $n \in \mathbb{N}$. The corresponding product sigma-algebra on E^n will be denoted by $\mathcal{E}_n$. Recall that we will be working with a sufficiently saturated nonstandard extension of the superstructure $V(X)$ over X. Let $k \in \mathbb{N}$. For $n \in \mathbb{N}_{\geq k}$, if $\Omega \in \mathcal{E}_n$ and ν is a measure on the induced sub-sigma-algebra on Ω, then for any $B \in \mathcal{E}_k$, we denote $\nu(\Omega \cap (B \times E^{n-k}))$ by $\nu(B)$. Similarly, we can talk about integrating a measurable function $f \colon E^k \to \mathbb{R}$ over Ω by extending f canonically to E^n.

Proposition 2.1 *Let $\Omega \in {}^*V(X)$ be such that $\Omega \subseteq {}^*E^N$ for some $N \in {}^*\mathbb{N}$. Let $\mathcal{E}$ be a sigma-algebra on E, and let $\mathcal{E}_k$ denote the corresponding product sigma-algebra on E^k for each $k \in \mathbb{N}$. Let ${}^*\mathcal{E}_N$ denote the corresponding internal algebra on ${}^*E^N$ (defined by extension of the sequence $\{\mathcal{E}_k\}_{k \in \mathbb{N}}$, which is an element of $V(X)$ when viewed as a function on $\mathbb{N}$). Let $\mathcal{F}$ be the restriction of ${}^*\mathcal{E}_N$ to Ω.*

Fix $k \in \mathbb{N}$ and suppose $\mathbb{P} \in \mathbf{Prob}(E^k, \mathcal{E}_k)$. Let $\nu \in {}^*\mathbf{Prob}(\Omega, \mathcal{F})$. If $L\nu$ is the corresponding Loeb measure and if $N \geq k$, then:

$$(2\text{--}1) \qquad \int_\Omega \mathbf{st}({}^*f) \, dL\nu = \int_{E^k} f \, d\mathbb{P} \text{ for all bounded measurable } f \colon E^k \to \mathbb{R}$$

if and only if

$$(2\text{--}2) \qquad L\nu({}^*B) = \mathbb{P}(B) \text{ for all } B \in \mathcal{E}_k$$

Proof If $f \colon E^k \to \mathbb{R}$ is bounded measurable, then $\mathbf{st}({}^*f)$ is Loeb measurable on Ω by Proposition 1.4. Hence the left side of equation (2–1) is well-defined.

The forward implication is immediate by taking $f = \mathbb{1}_B$, the indicator function of $B \in \mathcal{E}_k$. For the reverse implication, assume that $L\nu({}^*B) = \mathbb{P}(B)$ for all $B \in \mathcal{E}_k$ (that is, indicator functions of measurable sets satisfy (2–1)). The set of functions satisfying (2–1) is closed under taking finite $\mathbb{R}$-linear combinations, and hence all simple functions satisfy (2–1). Fix a bounded measurable function $f \colon E^k \to \mathbb{R}$. By standard measure theory (see, for example, Folland [6, Theorem 2.10]), there is a sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple functions that converges to f uniformly on E^k.

For $\epsilon \in \mathbb{R}_{>0}$, find $n_\epsilon \in \mathbb{N}$ such that we have the following inequality:

$$|f_n(x) - f(x)| < \epsilon \text{ for all } x \in E^k \text{ and } n \in \mathbb{N}_{\geq n_\epsilon}$$

By transfer, for all $n \in \mathbb{N}_{\geq n_\epsilon}$, we get $|{}^*f_n(x) - {}^*f(x)| < \epsilon$ on ${}^*E^k$. Hence:

$$|\mathbf{st}({}^*f_n(x)) - \mathbf{st}({}^*f(x))| \leq \epsilon \text{ for all } n \in \mathbb{N}_{\geq n_\epsilon} \text{ and } x \in {}^*E^k$$

As a consequence, we get

$$\left| \int_\Omega \mathbf{st}({}^*f) \, dL\nu - \int_\Omega \mathbf{st}({}^*f_n) \, dL\nu \right| \leq \epsilon \text{ for all } n \in \mathbb{N}_{\geq n_\epsilon},$$

that is, $\qquad \left| \int_\Omega \mathbf{st}({}^*f) \, dL\nu - \int_{E^k} f_n \, d\mathbb{P} \right| \leq \epsilon \text{ for all } n \in \mathbb{N}_{\geq n_\epsilon}.$

But $\lim_{n\to\infty} \int_{E^k} f_n \, d\mathbb{P} = \int_{E^k} f \, d\mathbb{P}$, by the Dominated Convergence Theorem. Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary, this implies $\int_\Omega \mathbf{st}({}^*f) \, dL\nu = \int_{E^k} f \, d\mathbb{P}$, completing the proof. $\qquad \square$

The hypothesis in Proposition 2.1 is an abstract rendering of the premise of our central problem about limits of spherical measures. Indeed, we may think of E as $\mathbb{R}$, the space Ω as the hyperfinite dimensional sphere $S^{N-1}(\sqrt{N})$ for some $N > \mathbb{N}$, and $\mathbb{P}$ as the standard Gaussian measure μ. Then, (2–2) is the nonstandard characterization of (1–1), while (2–1) corresponds to (1–2). To strengthen this theme, in the next subsection, we will take a standard sequence of probability spaces and replace Ω by the Nth term (for any $N > \mathbb{N}$) of the nonstandard extension of that sequence. We first record some useful implications of Proposition 2.1 below.

Corollary 2.2 *In the setting of Proposition 2.1, suppose (2–1), and hence (2–2), hold. Then $L\nu(\{x \in \Omega : {}^*f(x) \in {}^*\mathbb{R}_{fin}\}) = 1$ for all measurable $f \colon E^k \to \mathbb{R}$.*

Proof If $B_n := \{x \in E^k : |f(x)| < n\}$ for $n \in \mathbb{N}$, then the required probability is

$$L\nu\left(\cup_{n \in \mathbb{N}} {}^*B_n\right) = \lim_{n \to \infty} L\nu({}^*B_n) \overset{(2-2)}{=} \lim_{n \to \infty} \mathbb{P}(B_n) = 1,$$

thus completing the proof. $\qquad\square$

Corollary 2.3 *In the setting of Proposition 2.1, suppose (2–1) holds. Then, for any $\mathbb{P}$–integrable function $f \colon E^k \to \mathbb{R}$, we have that $\mathbf{st}({}^*f)$ is $L\nu$–integrable. Furthermore, we have*

$$\int_\Omega |\mathbf{st}({}^*f)| \, dL\nu = \int_{E^k} |f| \, d\mathbb{P},$$

and

$$\int_\Omega \mathbf{st}({}^*f) \, dL\nu = \int_{E^k} f \, d\mathbb{P}.$$

Proof We see that $\mathbf{st}({}^*f)$ is Loeb measurable on Ω by Corollary 2.2 and Proposition 1.4. Also, by Corollary 2.2, $\mathbf{st}({}^*f)\mathbb{1}_{\{|{}^*f|<n\}} \uparrow \mathbf{st}({}^*f)$ $L\nu$–almost surely. Hence, we have:

$$\int_\Omega |\mathbf{st}({}^*f)| \, dL\nu = \lim_{n \to \infty} \int_\Omega \mathbf{st}({}^*\,|f|) \cdot \mathbb{1}_{\{{}^*|f| \le n\}} dL\nu$$

$$= \lim_{n \to \infty} \int_{E^k} |f| \cdot \mathbb{1}_{\{|f| \le n\}} d\mathbb{P}$$

$$= \int_{E^k} |f| \, d\mathbb{P} < \infty$$

The first line follows from the Monotone Convergence Theorem (applied on the Loeb space $(\Omega, L(\mathcal{F}), L\nu)$), the second line follows from (2–1), and the third line follows from the Monotone Convergence Theorem (applied on the probability space $(E^k, \mathcal{E}_k, \mathbb{P})$).

Now, since $\lim_{n \to \infty} \left(\mathbf{st}({}^*f) \cdot \mathbb{1}_{\{|{}^*f|<n\}}\right) = \mathbf{st}({}^*f)$ $L\nu$-almost surely (using Corollary 2.2), and since $\left|\mathbf{st}({}^*f) \cdot \mathbb{1}_{\{|{}^*f|<n\}}\right| \le |\mathbf{st}({}^*f)| \in L^1(\Omega, L\nu)$, it follows that:

$$\int_\Omega \mathbf{st}({}^*f) \, dL\nu = \lim_{n \to \infty} \int_\Omega \mathbf{st}({}^*f) \cdot \mathbb{1}_{\{|{}^*f| \le n\}} dL\nu$$

$$= \lim_{n \to \infty} \int_{E^k} f \cdot \mathbb{1}_{\{|f| \le n\}} d\mathbb{P}$$

$$= \int_{E^k} f \, d\mathbb{P}$$

The first line follows from the Dominated Convergence Theorem (applied on the Loeb space $(\Omega, L(\mathcal{F}), L\nu)$), the second line follows from (2–1), and the third line follows from the Dominated Convergence Theorem (applied on the measure space $(E^k, \mathcal{E}_k, \mathbb{P})$). This completes the proof. □

Corollary 2.4 *In the setting of Proposition 2.1, the following are equivalent:*

(1) $\int_\Omega \mathrm{st}(^*f)\, dL\nu = \int_{E^k} f\, d\mathbb{P}$ *for all bounded measurable* $f\colon E^k \to \mathbb{R}$.

(2) $L\nu(^*B) = \mathbb{P}(B)$ *for all* $B \in \mathcal{E}_k$.

(3) $L\nu(^*B) \leq \mathbb{P}(B)$ *for all* $B \in \mathcal{E}_k$.

(4) $L\nu(^*B) \geq \mathbb{P}(B)$ *for all* $B \in \mathcal{E}_k$.

Proof (1) $\Leftrightarrow$ (2) follows from Proposition 2.1. Also, (3) and (4) follow from (2) immediately. Conversely, assume (3). For any Borel set $B \subseteq E^k$, we have

$$(2\text{–}3) \qquad\qquad L\nu(^*B) \leq \mathbb{P}(B), \text{ and}$$

$$(2\text{–}4) \qquad\qquad L\nu(^*E^k \backslash {}^*B) \leq \mathbb{P}(E^k \backslash B) \Rightarrow L\nu(^*B) \geq \mathbb{P}(B).$$

Combining (2–3) and (2–4) gives (2). The proof of (4) $\Rightarrow$ (2) is similar. □

We end this subsection with the remark that if E is a Hausdorff topological space equipped with its Borel sigma-algebra, and if the probability measure $\mathbb{P}$ is Radon, then (2–1) and (2–2) are both equivalent to the Loeb measure $L\nu$ agreeing with $\mathbb{P}$ on the nonstandard extensions of all open (or all compact) subsets of E.

Proposition 2.5 *In the setting of Proposition 2.1, suppose E is a Hausdorff topological space and let $\mathcal{B}(E^k)$ be the Borel sigma-algebra on E^k. If $\mathbb{P}$ is a Radon probability measure on E^k, then the following are equivalent:*

(1) $\int_\Omega \mathrm{st}(^*f)\, dL\nu = \int_{E^k} f\, d\mathbb{P}$ *for all bounded Borel measurable* $f\colon E^k \to \mathbb{R}$.

(2) $L\nu(^*B) = \mathbb{P}(B)$ *for all* $B \in \mathcal{B}(E^k)$.

(3) $L\nu(^*B) \leq \mathbb{P}(B)$ *for all* $B \in \mathcal{B}(E^k)$.

(4) $L\nu(^*B) \geq \mathbb{P}(B)$ *for all* $B \in \mathcal{B}(E^k)$.

(5) $L\nu(^*O) = \mathbb{P}(O)$ *for all open sets* $O \subseteq E^k$.

(6) $L\nu(^*C) = \mathbb{P}(C)$ *for all compact sets* $C \subseteq E^k$.

Proof The equivalence of (1), (2), (3), and (4) has been established without any conditions on $\mathbb{P}$ in the previous corollary. Also, (2) $\Rightarrow$ (5) is immediate. To complete the proof, we will show that (5) $\Rightarrow$ (6) and (6) $\Rightarrow$ (4).

To see (5) $\Rightarrow$ (6), note that if C is a compact subset of the Hausdorff space E^k, then C is closed, so that the subset $O := E^k \backslash C$ is open. By using the fact that $^*C = {}^*E^k \backslash {}^*O$, and then applying (5) to O, we obtain the following:

$$L\nu(^*C) = 1 - L\nu(^*O) = 1 - \mathbb{P}(O) = \mathbb{P}(C)$$

We now prove (6) $\Rightarrow$ (4). To that end, take any $B \in \mathcal{B}(E^k)$. For any compact subset $C \subseteq B$, we have $^*C \subseteq {}^*B$, so that (6) implies the following:

$$L\nu(^*B) \geq L\nu(^*C) = \mathbb{P}(C) \text{ for all compact subsets } C \text{ of } B.$$

Taking supremum over all compact subsets of B and using the fact that the measure $\mathbb{P}$ is Radon, we thus obtain the desired inequality as follows:

$$L\nu(^*B) \geq \sup\{\mathbb{P}(C) : C \text{ is a compact subset of } B\} = \mathbb{P}(B) \qquad \square$$

2.2 Basic facts about surface area measures and their nonstandard counterparts

In this subsection, we review three different ways to think about the uniform surface area measure on spheres in Euclidean spaces. One aim of our review is to explain the corresponding internal probability measures on hyperfinite dimensional spheres that we obtain by transfer. We refer to Matilla [11, Chapter 3] and Sengupta [20, Section 4] for basic properties of spherical surface area measures.

For each $n \in \mathbb{N}$, we let $\mathcal{B}_n = \mathcal{B}(\mathbb{R}^n)$, the Borel sigma-algebra on $\mathbb{R}^n$, and $O(n)$ be the set of all orthogonal linear transformations of $\mathbb{R}^n$. Let $\mathcal{S}_0$ be the set of all spheres centered at the origin (in any dimension $n \in \mathbb{N}$ and of any radius $r \in \mathbb{R}_{>0}$). Since X contains copies of all Euclidean spaces, $\mathcal{S}_0$ is an element of $V(X)$. Consider the function $\dim \colon \mathcal{S}_0 \to \mathbb{N}$ that takes each sphere S to the smallest dimension $n \in \mathbb{N}$ such that $S \subseteq \mathbb{R}^n$. We are being pedantic about the "smallest dimension" since we have been identifying (during discussions on measures of sets) a subset S of $\mathbb{R}^n$ with the subset $S \times \mathbb{R}^{n'-n} \subseteq \mathbb{R}^{n'}$ for $n' \in \mathbb{N}_{>n'}$.

It is known that there is a unique rotation-preserving probability measure on any sphere centered at origin equipped with its Borel sigma-algebra. More formally:

(2–5) $\quad \forall S \in \mathcal{S}_0 \, \exists! \bar{\sigma} \in \mathbf{Prob}(S, \mathcal{B}(S)) \, \forall n \in \mathbb{N}$

$$(n = \dim(S)) \to (\forall R \in O(n) \, \forall A \in \mathcal{B}(S) \, [\bar{\sigma}(R(A)) = \bar{\sigma}(A)])$$

For any $S \in {}^*\mathcal{S}_0$ in the nonstandard universe, the transfer principle implies that the set ${}^*\mathrm{Prob}(S, {}^*\mathcal{B}(S))$ consists of a unique finitely additive internal function, say $\bar{\sigma}_S \colon {}^*\mathcal{B}(S) \to {}^*[0,1]$, that is *–rotation preserving and satisfies $\bar{\sigma}_S(S) = 1$. By the usual Loeb measure construction, we get $L\bar{\sigma}_S$ on $\mathcal{L}({}^*\mathcal{B}(S))$ (a sigma-algebra containing $\sigma({}^*\mathcal{B}(\mathcal{S}))$, which we call the *uniform Loeb surface measure* on S. As before, we will often drop the superscript S in $\bar{\sigma}_S$ when the sphere is clear from context.

In finite dimensions, we also have the notion of surface area. For the sphere $S := S^d(R)$ of radius $R \in \mathbb{R}_{>0}$, centered at the origin in $\mathbb{R}^{d+1}$, one can consider the surface area map $\sigma_S \colon \mathcal{B}(S) \to \mathbb{R}$, which satisfies the volume-of-cone formula

$$\lambda_{d+1}\left(\bigcup_{0 \leq t \leq 1} tA\right) = \frac{1}{d+1} R\sigma_S(A)$$

where λ_{d+1} is the Lebesgue measure on $\mathbb{R}^{d+1}$, and $A \in \mathcal{B}(S)$. This surface area function has the following properties:

- For any $d \in \mathbb{N}$ and any $R \in \mathbb{R}_{>0}$, we have $\sigma_{S^d(R)}(S^d(R)) = c_d \cdot R^d$, where $c_d = \sigma_{S^d(1)}(S^d(1)) = (d+1)\pi^{(d+1)/2}/\Gamma\left(\frac{d+1}{2}+1\right) = 2\pi^{(d+1)/2}/\Gamma\left(\frac{d+1}{2}\right)$.
- For any $S \in \mathcal{S}$ and any $A \in \mathcal{B}(S)$, we have $\bar{\sigma}_S(A) = \frac{\sigma_S(A)}{\sigma_S(S)}$, where $\bar{\sigma}_S$ is the rotation preserving probability measure on S, as in (2.2).

By transfer, we have the notion of *–surface area (that is applicable to hyperfinite-dimensional spheres as well) in the nonstandard universe. This could be used as an alternative way to define the uniform Loeb surface measure.

Yet another way to arrive at the uniform surface area measure on a sphere is by looking at an appropriate pushforward of a Gaussian measure. If μ is the standard Gaussian measure on $\mathbb{R}^n$ (here $n \in \mathbb{N}$), and S^{n-1} is the unit sphere in $\mathbb{R}^n$, then the rotation invariance of μ implies that $\mu \circ g^{-1}$ is a rotation invariant probability measure on S^{n-1} (and hence is the same as $\bar{\sigma}$), where

$$g : \mathbb{R}^n \backslash \{0\} \to S^{n-1} \text{ defined by } g(x) = \frac{x}{\|x\|}.$$

For spheres centered at origin but having radius $R \in \mathbb{R}_{>0}$, we can use the pushforward through the map Rg (this is scalar multiple by R). For instance, for $N > \mathbb{N}$, if $\bar{\sigma}$ is the internal uniform surface area measure on $S^{N-1}(\sqrt{N})$ and $\mu_{(N)}$ is the internal Gaussian measure on ${}^*\mathbb{R}^N$ with mean $\mathbf{0}$ and covariance identity, then for any set $B \in {}^*\mathcal{B}(S^{N-1}(\sqrt{N}))$, we have:

(2–6) $$\bar{\sigma}(B) = \mu_{(N)}\left(\left\{x \in {}^*\mathbb{R}^N : \frac{\sqrt{N}x}{\|x\|} \in B\right\}\right)$$

This characterization of the uniform surface area measure yields the classical result of Poincaré (Theorem 1.1) without doing any computations. We show that in the next subsection.

2.3 A nonstandard proof of Poincaré's theorem

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $(X_n)_{n\in\mathbb{N}}$ is a sequence of iid $\mathcal{N}(0,1)$ random variables (that is, the X_i are independent Gaussian random variables with mean 0 and variance 1). In that case, $(X_n{}^2 - 1)_{n\in\mathbb{N}}$ is an iid sequence of random variables with mean zero and finite variance (in fact, the variance is equal to one). Hence the weak law of large numbers implies the following:

$$(2\text{--}7) \qquad \lim_{n\to\infty} \mathbb{P}\left(\left|\frac{(X_1{}^2 - 1) + \ldots + (X_n{}^2 - 1)}{n}\right| > \epsilon\right) = 0 \text{ for all } \epsilon \in \mathbb{R}_{>0}$$

Each X_i (where $i \in \mathbb{N}$), as a function from Ω to $\mathbb{R}$, has a nonstandard extension ${}^*X_i$, which, by transfer, is a ${}^*\mathcal{N}(0,1)$ random variable, that is, ${}^*\mathbb{P} \circ {}^*X_i{}^{-1}$ is the same as the internal measure ${}^*\mu_{(1)}$ (the nonstandard extension of the standard Gaussian measure $\mu_{(1)}$ on $\mathbb{R}$).

Consider the function $X\colon \mathbb{N} \times \Omega \to \mathbb{R}$ defined by:

$$(2\text{--}8) \qquad\qquad X(n, \omega) := X_n(\omega) \text{ for all } n \in \mathbb{N}, \text{ and } \omega \in \Omega$$

Considering the nonstandard extension of X, we see that

$$^*X(i, \omega) = {}^*X_i(\omega) \text{ for all } i \in \mathbb{N}, \text{ and } \omega \in \Omega.$$

Furthermore, this allows us to naturally talk about the Nth element of the original sequence of random variables for any $N \in {}^*\mathbb{N}$ (and all those elements will be independent and internally Gaussian distributed with mean 0 and variance 1). In the sequel, we will often be loose with notation, and use X_i as both a standard and a nonstandard random variable (when it is considered as a nonstandard random variable, it is understood to be given by the nonstandard extension of the map $X\colon \mathbb{N} \times \Omega \to \mathbb{R}$), with the usage being clear from context.

For the rest of this section, fix $N > \mathbb{N}$. Let $\bar{\sigma}$ be the internal uniform surface area measure on $S^{N-1}(\sqrt{N})$. Let $Y = (X_1)^2 + \ldots + (X_N)^2$.

Lemma 2.6 *There exists an infinitesimal $\xi > 0$ such that*

$$(2\text{--}9) \qquad\qquad {}^*\mathbb{P}\left(\left|\frac{Y}{N} - 1\right| > \xi\right) \approx 0.$$

Proof Consider $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < 1$. Then we have

$$^*\mathbb{P}\left(\left|\frac{Y}{N} - 1\right| > \epsilon\right) = {}^*\mathbb{P}\left(\left|\frac{(X_1{}^2 - 1) + \ldots + (X_N{}^2 - 1)}{N}\right| > \epsilon\right)$$

where the right side is infinitesimal by the nonstandard characterization of limits applied to (2–7). The lemma now follows by underflow applied to the following internal set:

$$\left\{\epsilon \in {}^*\mathbb{R}_{>0} : {}^*\mathbb{P}\left(\left|\frac{Y}{N} - 1\right| > \epsilon\right) < \epsilon\right\} \qquad \square$$

For a set $S \subseteq \mathbb{R}^k$ and a real number $\alpha \in \mathbb{R}$, the set αS is the set of all scalar products (of elements of S) by α. That is,

$$\alpha S := \{y \in \mathbb{R}^k : y = \alpha x \text{ for some } \alpha \in A\}.$$

For $S \subseteq \mathbb{R}^k$ and $A \subseteq \mathbb{R}$, the set AS is defined as the set of all scalar products of elements of S with elements in A. That is,

$$(2\text{–}10) \qquad\qquad AS := \cup_{\alpha \in A} \alpha S.$$

Scalar products (with elements of $^*\mathbb{R}$ or with internal subsets of $^*\mathbb{R}$) are analogously defined in the nonstandard universe by transfer. We note the following elementary fact about small scalings of compact sets that will be useful in the sequel.

Lemma 2.7 *Let C be a compact subset of $\mathbb{R}^k$. Then we have:*

$$(2\text{–}11) \qquad\qquad \bigcap_{n \in \mathbb{N}_{>1}} \left[1 - \frac{1}{n}, 1 + \frac{1}{n}\right] C = C$$

Proof Let the left side of (2–11) be called $\widetilde{C}$ for brevity. It is clear that $C \subseteq \widetilde{C}$. To show the inclusion from the other side, consider $x \in \widetilde{C}$. Thus, for each $n \in \mathbb{N}_{>1}$ there exist $\alpha_n \in \mathbb{R}$ and $y_n \in C$ such that $x = \alpha_n y_n$. By the sequential compactness of C, find a subsequence $(n_k)_{k \in \mathbb{N}}$ such that $\lim_{k \to \infty} y_{n_k}$ exists as an element of C. Say, $\lim_{k \to \infty} y_{n_k} = y \in C$. Note that, by construction we have $\lim_{k \to \infty} \alpha_{n_k} = 1$. By continuity of the scalar product map we thus have the following,

$$(2\text{–}12) \qquad x = \lim_{k \to \infty} \alpha_{n_k} y_{n_k} = \left(\lim_{k \to \infty} \alpha_{n_k}\right)\left(\lim_{k \to \infty} y_{n_k}\right) = y \in C$$

completing the proof. $\qquad \square$

We now prove Poincaré's theorem that we restate here for convenience:

Theorem 1.1 For all bounded measurable functions $f\colon \mathbb{R}^k \to \mathbb{R}$, we have:

$$\lim_{n\to\infty} \int_{S^{n-1}(\sqrt{n})} f d\bar{\sigma} = \int_{\mathbb{R}^k} f d\mu$$

Proof Let B be a Borel subset of $\mathbb{R}^k$ and let $\mathbf{X} = (X_1,\ldots,X_N)$ be as defined in (2–8). For $k \in \mathbb{N}$, let $\mathbf{X}_{(k)}$ be the projection $(X_1,\ldots,X_k)$ onto $^*\mathbb{R}^k$. Using (2–6) and taking standard parts on both sides yields the following:

(2–13)
$$L\bar{\sigma}(\{(x_1\ldots,x_N) \in S^{N-1}(\sqrt{N}) : (x_1,\ldots,x_k) \in {}^*B\})$$
$$= L^*\mathbb{P}\left(\frac{\sqrt{N}\mathbf{X}_{(k)}}{\sqrt{Y}} \in {}^*B\right) = L^*\mathbb{P}\left(\mathbf{X}_{(k)} \in \sqrt{\frac{Y}{N}}{}^*B\right)$$

Using Lemma 2.6, the last expression is less than or equal to

$$L^*\mathbb{P}\left(\mathbf{X}_{(k)} \in \bigcap_{n=2}^m {}^*\left[1-\frac{1}{n}, 1+\frac{1}{n}\right]B\right) = L^*\mathbb{P}\left(\mathbf{X}_{(k)} \in {}^*\left(\bigcap_{n=2}^m\left[1-\frac{1}{n}, 1+\frac{1}{n}\right]B\right)\right)$$

for all $m \in \mathbb{N}$.

Taking limits as $m \to \infty$, we obtain:

$$L\bar{\sigma}(\{(x_1\ldots,x_N) \in S^{N-1}(\sqrt{N}) : (x_1,\ldots,x_k) \in {}^*B\})$$

(2–14)
$$\leq \lim_{m\to\infty} L^*\mathbb{P}\left(\mathbf{X}_{(k)} \in {}^*\left[\bigcap_{n=2}^m\left[1-\frac{1}{n}, 1+\frac{1}{n}\right]B\right]\right)$$
$$= \lim_{m\to\infty} \mathbb{P}\left(\mathbf{X}_{(k)} \in \bigcap_{n=2}^m\left[1-\frac{1}{n}, 1+\frac{1}{n}\right]B\right)$$
$$= \mathbb{P}\left(\mathbf{X}_{(k)} \in \bigcap_{n\in\mathbb{N}_{>1}}\left[1-\frac{1}{n}, 1+\frac{1}{n}\right]B\right)$$

By (2–14) and Lemma 2.7, we have the inequality

$$L\bar{\sigma}(\{(x_1\ldots,x_N) \in S^{N-1}(\sqrt{N}) : (x_1,\ldots,x_k) \in {}^*C\}) \leq \mathbb{P}\left(\mathbf{X}_{(k)} \in C\right) = \mu_{(k)}(C)$$

for all compact subsets $C \subseteq \mathbb{R}^k$.

Since $N > \mathbb{N}$ is arbitrary and $\mu_{(k)}$ is a Radon measure, Proposition 2.5 and the nonstandard characterization of limits complete the proof. $\square$

3 On the limiting behavior of a sequence of probability spaces

Toward the proof of Poincaré's theorem in the previous section, we showed that for an arbitrary $N > \mathbb{N}$ the surface area measure over $S^{N-1}(\sqrt{N})$ (which may be thought of as the Nth element of the sequence of spheres $(S^{n-1}(\sqrt{n}))_{n \in \mathbb{N}}$) assigns the same measure (up to infinitesimals) to fibers of finite dimensional sets as the Gaussian measures of such sets (in their respective ambient Euclidean spaces). This idea is explored in more abstract settings in the current section in order to generalize to limiting results for integrals of unbounded functions.

3.1 Integrating finite dimensional functions along nice sequences of probability spaces

Let $\{(\Omega_n, \mathcal{F}_n, \nu_n)\}_{n \in \mathbb{N}}$ be a sequence of probability spaces. Viewing the sequence as a function on $\mathbb{N}$, we get an internal probability space $(\Omega_N, \mathcal{F}_N, \nu_N)$ for each $N > \mathbb{N}$. Note that we have been dropping the * when it is clear from context that the index N is hyperfinite. Philosophically, the Loeb space $(\Omega_N, L(\mathcal{F}_N), L\nu_N)$ for $N > \mathbb{N}$ should capture the long-term behavior of the sequence $\{(\Omega_n, \mathcal{F}_n, \nu_n)\}_{n \in \mathbb{N}}$ of probability spaces. We will often omit the sigma-algebra when there is no chance of confusion. Drawing inspiration from Theorem 1.5(4), we obtain the following theorem in this regard.

Theorem 3.1 *Let $(E, \mathcal{E})$ be a measure space. Let $k \in \mathbb{N}$, and for each $n \in \mathbb{N}_{>k}$ suppose $\Omega_n \subseteq E^{n'}$ for some $n' \in \mathbb{N}_{>k}$. Suppose that $\mathcal{F}_n$, the given sigma-algebra on Ω_n, is induced by the product sigma-algebra $\mathcal{E}_{n'}$ on $E^{n'}$. Let $(\Omega_n, \mathcal{F}_n, \nu_n)$ be a sequence of Borel probability spaces. Let $f : E^k \to \mathbb{R}$ satisfy*

$$(3\text{--}1) \qquad \lim_{m \to \infty} \lim_{n \to \infty} \int_{\Omega_n \cap \{|f| \geq m\}} |f| \, d\nu_n = 0$$

Then, f is integrable over Ω_n for large n, so that the sequence $\alpha_{f,n} := \int_{\Omega_n} f d\nu_n$ is well-defined for large n. Furthermore, for any $N > \mathbb{N}$, the function $\mathbf{st}(^*f)$ is Loeb integrable over $(\Omega_N, L(\mathcal{F}_N), L\nu_N)$ and satisfies

$$\mathbf{st}(\alpha_{f,N}) = \int_{\Omega_N} \mathbf{st}(^*f) dL\nu_N$$

Remark 3.2 Bounded measurable functions trivially satisfy the hypothesis in (3–1).

Proof For a fixed $\epsilon \in \mathbb{R}_{>0}$ there exists $\ell_\epsilon \in \mathbb{N}$ such that the following holds: for any $m \geq \ell_\epsilon$ there is an $n_{\epsilon,m} \in \mathbb{N}$ such that for all $n \geq n_{\epsilon,m}$ we have

$$(3\text{--}2) \qquad \int_{\Omega_n \cap \{|f| \geq m\}} |f| \, d\nu_n < \epsilon$$

In particular, f is integrable on Ω_n for all $n > n_{\epsilon,\ell_\epsilon}$, with the integral of the absolute value being at most $(\ell_\epsilon + \epsilon)$. Further, for any $M, N > \mathbb{N}$, transfer yields

$$^*\!\!\int_{\Omega_N} |^*f| \, \mathbb{1}_{\{|^*f|>M\}} d\nu_N \leq {}^*\!\!\int_{\Omega_N} |^*f| \, \mathbb{1}_{\{|^*f|>\ell_\epsilon\}} d\nu_N < \epsilon \text{ for all } \epsilon \in \mathbb{R}_{>0}$$

Given $N > \mathbb{N}$, $^*f$ is S–integrable on Ω_N by Theorem 1.5(2).

Now, $\alpha_{f,N}$ is the *–integral of $^*f$ over (Ω_N, ν_N) by transfer. Note that

$$f = f_+ - f_-$$

where $f_+ := \max\{f, 0\}$ and $f_- := \max\{-f, 0\}$. By transfer, we then have:

$$(3\text{--}3) \qquad \alpha_{f,N} = \alpha_{f_+,N} - \alpha_{f_-,N}$$

Since $^*f$ is S–integrable on (Ω_N, ν_N), so are $^*f_+$ and $^*f_-$ (this is because $|^*f_+|$ and $|^*f_-|$ are at most equal to $|^*f|$). Since $^*f_+$ and $^*f_-$ are nonnegative functions, Theorem 1.5(4) implies:

$$(3\text{--}4) \qquad \begin{aligned} \alpha_{f_+,N} &= \int_{\Omega_N} \mathbf{st}(^*f_+) \, dL\nu_N, \text{ and} \\ \alpha_{f_-,N} &= \int_{\Omega_N} \mathbf{st}(^*f_-) \, dL\nu_N. \end{aligned}$$

Using this in (3–3) and then using the fact that $\mathbf{st}(^*f)$ is Loeb integrable completes the proof. $\square$

Corollary 3.3 *Let $(E, \mathcal{E})$ be a measure space. Let $k \in \mathbb{N}$, and for each $n \in \mathbb{N}_{>k}$ suppose $\Omega_n \subseteq E^{n'}$ for some $n' \in \mathbb{N}_{>k}$. Suppose that $\mathcal{F}_n$, the given sigma-algebra on Ω_n, is induced by the product sigma-algebra $\mathcal{E}_{n'}$ on $E^{n'}$. Let $(\Omega_n, \mathcal{F}_n, \nu_n)$ be a sequence of Borel probability spaces. Let $\mathbb{P}$ be a probability measure on $(E^k, \mathcal{E}_k)$ such that $L\nu_N(^*B) = \mathbb{P}(B)$ for any $B \in \mathcal{E}_k$ and $N > \mathbb{N}$.*

(i) *If $f \colon E^k \to \mathbb{R}$ is measurable then*

$$L\nu_N(\{x \in \Omega_N : {}^*f(x) \in {}^*\mathbb{R}_{fin}\}) = 1 \text{ for all } N > \mathbb{N}.$$

(ii) *If* $f\colon E^k \to \mathbb{R}$ *is bounded and measurable then*

$$\lim_{n\to\infty} \int_{\Omega_n} f d\nu_n = \int_{E^k} f d\mathbb{P} = \int_{\Omega_N} \mathbf{st}(^*f)\, dL\nu_N \text{ for all } N > \mathbb{N}.$$

(iii) *If* $f\colon E^k \to \mathbb{R}$ *is* $\mathbb{P}$*–integrable then we have that* $\mathbf{st}(^*f)$ *is* $L\nu_N$*–integrable for all* $N > \mathbb{N}$. *Furthermore, for any* $N > \mathbb{N}$ *we have:*

$$\int_{E^k} f d\mathbb{P} = \int_{\Omega_N} \mathbf{st}(^*f)\, dL\nu_N, \text{ and } \int_{E^k} |f|\, d\mathbb{P} = \int_{\Omega_N} |\mathbf{st}(^*f)|\, dL\nu_N$$

Proof (i) follows from Corollary 2.2. (ii) follows from Theorem 3.1, Corollary 2.3 and the nonstandard characterization of limits. Finally, (iii) follows from Corollary 2.3, completing the proof. □

Note that Corollary 3.3(iii) allows us to express the expected value of a $\mathbb{P}$–integrable function $f\colon E^k \to \mathbb{R}$ as the Loeb integral of $\mathbf{st}(^*f)$ over Ω_N for all hyperfinite N. However, this does not necessarily imply that the sequence $\alpha_{f,n} := \int_{\Omega_n} f d\nu_n$ converges to $\int_{E^k} f d\mathbb{P}$, as $\alpha_{f,N}$ may not be infinitesimally close to the Loeb integral of $\mathbf{st}(^*f)$ over Ω_N in general. To see a counterexample, consider $(E, \mathcal{E}) = (\mathbb{N}_0, \mathcal{P}(\mathbb{N}))$ (where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$), with $\Omega_n := \{0, n\}$ for each $n \in \mathbb{N}$. Define $\mathbb{P} := \mathbb{1}_{\{0\}}$, the probability measure concentrated at 0. Define $\nu_n(\{0\}) = 1 - 1/n$ and $\nu_n(\{n\}) = 1/n$. Then for any $N > \mathbb{N}$, the Loeb measure $L\nu_N$ assigns full mass to $\{0\}$. Thus the hypotheses of Corollary 3.3 are satisfied. Consider the measurable function $f\colon \mathbb{N}_0 \to \mathbb{R}$ defined by $f(n) := n$ for all $n \in \mathbb{N}$. It is clear that $\alpha_{f,N}$ equals 1 while the Loeb integral of $\mathbf{st}(^*f)$ equals 0.

In view of Theorem 1.5, the correct criterion needed for $\alpha_{f,N}$ to be infinitesimally close to the Loeb integral of $\mathbf{st}(^*f)$ over Ω_N for nonnegative functions f is the S–integrability of $^*f$ over Ω_N. This also means that the sufficient criterion (3–1) in Theorem 3.1 is necessary if we restrict to nonnegative functions. We record and prove these observations in the following theorem.

Theorem 3.4 *In the setting of Corollary 3.3, the following are equivalent for a non-negative function* $f\colon E^k \to \mathbb{R}_{\geq 0}$:

(1) f *is* $\mathbb{P}$*–integrable and* $\lim_{n\to\infty} \int_{\Omega_n} f d\nu_n = \int_{E^k} f d\mathbb{P}$.

(2) *The nonstandard extension* $^*f$ *is* S*–integrable on* Ω_N *for all* $N > \mathbb{N}$.

(3) *The function* f *is integrable on* (Ω_n, ν_n) *for all large* $n \in \mathbb{N}$, *and furthermore:*

$$\lim_{m\to\infty} \lim_{n\to\infty} \int_{\Omega_n \cap \{f \geq m\}} f d\nu_n = 0$$

Proof $(1) \Rightarrow (2)$ Assume that f is $\mathbb{P}$–integrable and $\lim_{n \to \infty} \int_{\Omega_n} f d\nu_n = \int_{E^k} f d\mathbb{P}$. Using the nonstandard characterization of limits, Corollary 3.3(iii), and Theorem 1.5(4) (making use of the fact that $f = |f|$ since f is assumed to be nonnegative), it follows that ${}^*f$ is S–integrable on Ω_N for any $N > \mathbb{N}$.

$(2) \Rightarrow (3)$ Now assume that ${}^*f$ is S–integrable on Ω_N for all $N > \mathbb{N}$. As a consequence (using either Theorem 1.5(2) or Theorem 1.5(3)), we have that ${}^*f \mathbb{1}_{\{|{}^*f| \geq m\}}$ is S–integrable on Ω_N for any $N > \mathbb{N}$ and $m \in \mathbb{N}$. Fix $N_0 > \mathbb{N}$ such that the following is true. (Existence of such an N_0 is guaranteed by the nonstandard characterization of limit superior.)

$$\limsup_{n \to \infty} \int_{\Omega_n \cap \{|f| \geq m\}} |f| \, d\nu_n = \mathbf{st} \left({}^*\!\!\int_{\Omega_{N_0}} {}^* |f| \, \mathbb{1}_{\{{}^*|f| \geq m\}} d\nu_{N_0} \right)$$

By Theorem 1.5(4), we get:

$$\limsup_{n \to \infty} \int_{\Omega_n \cap \{|f| \geq m\}} |f| \, d\nu_n = \int_{\Omega_{N_0}} \mathbf{st} \left({}^* |f| \, \mathbb{1}_{\{{}^*|f| \geq m\}} \right) dL\nu_{N_0}$$

Therefore:

$$(3\text{–}5) \qquad \lim_{m \to \infty} \limsup_{n \to \infty} \int_{\Omega_n \cap \{|f| \geq m\}} |f| \, d\nu_n = \lim_{m \to \infty} \int_{\Omega_{N_0}} \mathbf{st} \left({}^* |f| \, \mathbb{1}_{\{{}^*|f| \geq m\}} \right) dL\nu_{N_0}$$

Since ${}^*f$ is S–integrable on Ω_{N_0}, it follows that $\mathbf{st}({}^*f)$ is Loeb integrable on Ω_{N_0}. Hence the limit on the right side of (3–5) is zero, as desired.

$(3) \Rightarrow (1)$ This follows from Theorem 3.1, Corollary 3.3(iii), and Theorem 1.5(4). $\qquad \square$

3.2 Application to a proof of the Riemann-Lebesgue Lemma

The theory of limiting integrals built over the last two subsections may theoretically be applied to a lot of situations in which the probability spaces are changing. While we will cover its application to spherical integrals in the next section, we include here a new proof of the famous Riemann–Lebesgue lemma as an illustration of the versatility of this theory. We paraphrase the Riemann–Lebesgue lemma below (see, for example, Rudin [19, 5.14, page 103]).

Theorem 3.5 (Riemann–Lebesgue Lemma) *Let λ be the Lebesgue measure on the interval $T := [-\pi, \pi]$. If $f \in L^1(T, \lambda)$ then we have:*

$$\lim_{n \to \infty} \int_T f(x) \cos(nx) d\lambda(x) = 0 \text{ and } \lim_{n \to \infty} \int_T f(x) \sin(nx) d\lambda(x) = 0$$

Proof For each $n \in \mathbb{N}$ define $g_n \colon T \to \mathbb{R}$ by $g_n(x) = (1 - \cos(nx))/2\pi$. The functions g_n are probability densities on $[-\pi, \pi]$. For each $n \in \mathbb{N}$ let $\mathbb{P}_n$ denote the probability measure on T with the density g_n. By integrating the densities for $n \in \mathbb{N}$ we find that the corresponding probability distribution functions are given by:

$$G_n(x) := \mathbb{P}_n\{(-\infty, x]\} = \frac{1}{2\pi}\left(x - \frac{\sin(x)}{n}\right) \text{ for all } x \in T$$

As $n \to \infty$, the sequence G_n converges pointwise to the distribution function of the uniform (normalized) Lebesgue measure $\mathbb{P}$ on $[-\pi, \pi]$. Thus $\mathbb{P}_n \xrightarrow{weak} \mathbb{P}$; that is,

$$\lim_{n \to \infty} \int_T f d\mathbb{P}_n = \int_T f d\mathbb{P} \text{ for all bounded continuous } f \colon T \to \mathbb{R}.$$

By an equivalent criterion for weak convergence, we obtain:

(3–6) $$\liminf_{n \to \infty} \mathbb{P}_n(U) \geq \mathbb{P}(U) \text{ for all open subsets } U \subseteq T$$

By the nonstandard characterization of limit inferiors, this is equivalent to:

(3–7) $$\mathbb{P}_N(^*U) \geq \mathbb{P}(U) \text{ for all open subsets } U \subseteq T \text{ and } N > \mathbb{N}$$

Since the density function g_n for $\mathbb{P}_n$ is pointwise bounded above by the density function for $\mathbb{P}$, by transfer we also obtain the other side of the above inequality. That is, we obtain:

(3–8) $$\mathbb{P}_N(^*U) \leq \mathbb{P}(U) \text{ for all open subsets } U \subseteq T \text{ and } N > \mathbb{N}$$

Combining (3–7) and (3–8), we obtain:

(3–9) $$\mathbb{P}_N(^*U) = \mathbb{P}(U) \text{ for all open subsets } U \subseteq T \text{ and } N > \mathbb{N}$$

By Proposition 2.5, we obtain:

(3–10) $$\int_{^*T} \mathbf{st}(^*f) dL\mathbb{P}_N = \int_T f d\mathbb{P} \text{ for all bounded measurable } f \colon T \to \mathbb{R} \text{ and } N > \mathbb{N}$$

For any $f \in L^1(T, \lambda)$, we use the facts that $|g_n| \leq \frac{1}{\pi}$ and $f \in L^1(T, \lambda)$ to get:

(3–11) $$\lim_{m \to \infty} \lim_{n \to \infty} \int_T |f| \, \mathbb{1}_{|f| > m} d\mathbb{P}_n(x) \leq \frac{1}{\pi} \lim_{m \to \infty} \int_T |f| \, \mathbb{1}_{|f| > m} d\lambda(x) = 0$$

Using (3–10) and (3–11) in Theorem 3.4 (with $(T, \mathbb{P}_n)$ playing the role of (Ω_n, ν_n) in that theorem), we obtain, for each $f \in L^1(T, \lambda) = L^1(T, \mathbb{P})$,

$$\lim_{n\to\infty} \int_T f(x)d\mathbb{P}_n(x) = \int_T f(x)d\mathbb{P}(x),$$

so

$$\lim_{n\to\infty} \int_T \left(\frac{f(x)}{2\pi} - \frac{f(x)\cos(nx)}{2\pi} \right) d\lambda(x) = \int_T \frac{f(x)}{2\pi} d\lambda(x),$$

therefore

$$\lim_{n\to\infty} \int_T f(x)\cos(nx)d\lambda(x) = 0.$$

The proof for $\sin(nx)$ goes exactly the same way if we replace the f_n by the probability density functions $g_n(x) = (1 - \sin(nx))/2\pi$ for $x \in T$. $\qquad\square$

3.3   What happens if the finite dimensional function is not nice in the limiting space?

In general, for a function $f\colon E^k \to \mathbb{R}$ (not necessarily satisfying the conditions in Theorem 3.4), the following result allows us to still approximate its integral by a suitably modified sequence of integrals over (Ω_n, ν_n). Note that this result is in the spirit of Littlewood's three principles from measure theory (see [10, page 26]) – approximating a potentially ill-behaved integrable function by well-behaved bounded functions.

Lemma 3.6 *In the setting of Corollary 3.3, let $f\colon E^k \to \mathbb{R}$ be $\mathbb{P}$–integrable. Given any $\epsilon, \delta, \theta \in \mathbb{R}_{>0}$ there exist an $n_0 \in \mathbb{N}$ and functions $g_n\colon \Omega_n \to \mathbb{R}$ for all $n \in \mathbb{N}_{\geq n_0}$ such that the following hold:*

  (i) *$|g_n|$ is bounded by n for all $n \in \mathbb{N}_{\geq n_0}$*

  (ii) *$\nu_n \left(|g_n - f| > \delta \right) < \epsilon$ for all $n \in \mathbb{N}_{\geq n_0}$*

  (iii) *$\left| \int_{\Omega_n} g_n d\nu_n - \int_{E^k} f d\mathbb{P} \right| < \theta$ for all $n \in \mathbb{N}_{\geq n_0}$*

Proof By Corollary 3.3(iii), we know that

$$\int_{E^k} |f| \, d\mu = \int_{\Omega_N} \mathbf{st}(^* |f|)dL\nu_N \text{ for all } N > \mathbb{N}.$$

Thus, for any $N > \mathbb{N}$, the map $\mathbf{st}(^*f)$ is Loeb integrable on Ω_N, and hence has an S–integrable lifting $G_N\colon \Omega_N \to {}^*\mathbb{R}$ by Theorem 1.6. In particular,

(3–12) $$L\nu_N \left(\mathbf{st}(G_N) = \mathbf{st}(^*f) \right) = 1, \text{ and}$$

(3–13) $$\mathbf{st} \left({}^* \!\! \int_{\Omega_N} G_N d\nu_N \right) = \int_{\Omega_N} \mathbf{st}(G_N)dL\nu_N = \int_{\Omega_N} \mathbf{st}(^*f)dL\nu_N = \int_{E^k} f d\mathbb{P}.$$

Equation (3–12) follows from the definition of lifting. The first equality in (3–13) follows from Theorem 1.5(4) applied to the nonnegative S–integrable functions $(G_N)_+ := \max\{G_N, 0\}$ and $(G_N)_- := \max\{-G_N, 0\}$. The second equality in (3–13) follows from equation (3–12), while the last equality in (3–13) follows from Corollary 3.3(iii).

Without loss of generality, we can assume that $|G_N| \leq N$ for all $N > \mathbb{N}$ (as we may replace G_N by the function $G_N \mathbb{1}_{|G_N| \leq N}$, which still satisfies (3–12) and (3–13)). Thus, for the given $\epsilon, \delta, \theta \in \mathbb{R}_{>0}$, the following internal set contains $^*\mathbb{N} \backslash \mathbb{N}$:

$$\mathcal{G}_{\epsilon, \delta, \theta} := \left\{ n \in {}^*\mathbb{N} : \exists G_n \in {}^*L^1(\Omega_n, \nu_n) \text{ such that } |G_n| \leq n, \right.$$

$$\left. {}^*\nu_n\left(|G_n - {}^*f| > \delta\right) < \epsilon, \text{ and } \left| {}^*\!\!\int_{{}^*\Omega_n} G_n d^*\nu_n - \int_{E^k} f d\mathbb{P} \right| < \theta \right\}$$

By underflow, we find $n_0 \in \mathbb{N}$ such that $\mathbb{N}_{\geq n_0} \subseteq \mathcal{G}_{\epsilon, \delta, \theta}$. Now fix an $n \in \mathbb{N}_{\geq n_0}$. In the nonstandard universe, the following statement is true:

$$\exists G_n \in {}^*L^1(\Omega_n, \nu_n)$$

$$\left((|G_n| \leq n) \wedge \left({}^*\nu_n\left(|G_n - {}^*f| > \delta\right) < \epsilon \right) \right) \wedge \left(\left| {}^*\!\!\int_{{}^*\Omega_n} G_n d^*\nu_n - \int_{E^k} f d\mathbb{P} \right| < \theta \right) \right)$$

Transfer of this sentence yields a $g_n \in L^1(\Omega_n, \nu_n)$ with the desired properties. $\square$

We can strengthen Lemma 3.6 as follows, by requiring the functions to have the same domain E^k.

Theorem 3.7 *In the setting of Corollary 3.3, let $f: E^k \to \mathbb{R}$ be $\mathbb{P}$–integrable. Given any $\epsilon, \delta, \theta \in \mathbb{R}_{>0}$ there exist an $n_0 \in \mathbb{N}$ and functions $g_n: E^k \to \mathbb{R}$ for all $n \in \mathbb{N}_{\geq n_0}$ such that the following hold:*

(i) *$|g_n|$ is bounded by n for all $n \in \mathbb{N}_{\geq n_0}$*

(ii) *$\nu_n\left(|g_n - f| > \delta\right) < \epsilon$ for all $n \in \mathbb{N}_{\geq n_0}$*

(iii) *$\left| \int_{\Omega_n} g_n d\nu_n - \int_{E^k} f d\mathbb{P} \right| < \theta$ for all $n \in \mathbb{N}_{\geq n_0}$*

Proof For $n \in \mathbb{N}_{\geq k}$, define $\nu_n': \mathcal{E}_k \to [0, 1]$ by $\nu_n'(B) = \nu_n((B \times E^{n-k}) \cap \Omega_n)$. For any bounded measurable $g: E^k \to \mathbb{R}$, expressing g as a uniform limit of simple functions yields:

$$(3\text{–}14) \qquad\qquad \int_{\Omega_n} g d\nu_n = \int_{E^k} g d\nu_n'$$

Let $(g_n)_{n \in \mathbb{N}}$ be a sequence of functions obtained by applying Lemma 3.6 to the sequence $(E^k, \nu'_n)_{n \in \mathbb{N}}$ of probability spaces. Then (i), (ii) and (iii) follow from the corresponding results in Lemma 3.6 together with (3–14). $\qquad \square$

4 Generalizing Poincaré's theorem

4.1 Revisiting a standard proof of Poincaré's theorem

For the rest of the paper, we let S_n denote the sphere $S^{n-1}(\sqrt{n})$ and $\bar{\sigma}_n$ denote $\bar{\sigma}_{S_n}$, for all $n \in \mathbb{N}$. Fix $k \in \mathbb{N}$ and let μ denote the standard k-dimensional Gaussian measure. Let $B_k(a)$ denote the open ball of radius a in $\mathbb{R}^k$. For a set $B \in \mathcal{B}(\mathbb{R}^k)$ and any $n \in \mathbb{N}_{\geq k}$, we define $\bar{\sigma}_n(B)$ to be the value of $\bar{\sigma}_n(\{x \in S_n : \pi_k(x) \in B\}) = \bar{\sigma}_n\left((B \times \mathbb{R}^{n-k}) \cap S_n\right)$, where π_k is the projection onto $\mathbb{R}^k$. Similarly, a function $f \colon \mathbb{R}^k \to \mathbb{R}$ is canonically extended to $\mathbb{R}^n$ by using '$f(x, y)$' to denote $f(x)$ for all $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^{n-k}$.

In an attempt to generalize Theorem 1.1, we first look at another proof of the same result using classical analysis. This proof requires directly evaluating the spherical integrals and using dominated convergence theorem (compare with the less computational proof of Theorem 1.1 in Section 2.3). We restate Theorem 1.1 below for convenience.

Theorem 1.1 For all bounded measurable functions $f \colon \mathbb{R}^k \to \mathbb{R}$, we have:

$$\lim_{n \to \infty} \int_{S^{n-1}(\sqrt{n})} f d\bar{\sigma} = \int_{\mathbb{R}^k} f d\mu$$

Proof Let λ denote the Lebesgue measure on $\mathbb{R}^k$. By Sengupta's disintegration formula (see [20, Proposition 4.1]), we have the following chain of equalities for any bounded measurable $f \colon \mathbb{R}^k \to \mathbb{R}$.

$$\int\limits_{S^{n-1}(\sqrt{n})} f d\bar\sigma_n$$

$$= \frac{1}{\sigma(S_n)} \int\limits_{x\in B_k(\sqrt{n})} \int\limits_{y\in S^{n-k-1}(\sqrt{n-\|x\|^2})} f(x,y) d\sigma(y) \frac{\sqrt{n}}{\sqrt{n-\|x\|^2}} d\lambda(x)$$

$$(4\text{–}1) \quad = \frac{1}{\sigma(S_n)} \int_{\mathbb{R}^k} \sigma\left(S^{n-k-1}\left(\sqrt{n-\|x\|^2}\right)\right) \cdot \frac{\mathbb{1}_{B_k(\sqrt{n})}(x) f(x) \sqrt{n}}{\sqrt{n-\|x\|^2}} d\lambda(x)$$

$$= \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{\frac{n}{2}} \cdot (\sqrt{n})^{n-1}} \int_{\mathbb{R}^k} \frac{2\pi^{\frac{n-k}{2}}(n-\|x\|^2)^{\frac{n-k-1}{2}}}{\Gamma\left(\frac{n-k}{2}\right)} \cdot \frac{\mathbb{1}_{B_k(\sqrt{n})}(x) f(x) \sqrt{n}}{\sqrt{n-\|x\|^2}} d\lambda(x)$$

$$= a_{n,k} b_{n,k} \int_{\mathbb{R}^k} \frac{1}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2}} \frac{\mathbb{1}_{B_k(\sqrt{n})}(x) f(x)}{\left(1 - \frac{\|x\|^2}{n}\right)^{\frac{k+2}{2}}} d\lambda(x)$$

where $a_{n,k} = \dfrac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right) \cdot \left(\frac{n-k}{2}\right)^{k/2}}$ and $b_{n,k} = \left(1 - \frac{k}{n}\right)^{k/2}$.

Note that $\lim_{n\to\infty} a_{n,k} = \lim_{n\to\infty} b_{n,k} = 1$ for all $k \in \mathbb{N}$ (the first limit following from Stirling's formula; see Rudin [18, equation 103, page 194]).

Modulo constants, for large values of n the integrand in (4–1) is bounded by $|f(x)| e^{-\|x\|^2/4}$, which is integrable on $\mathbb{R}^k$ since f is assumed to be bounded. Thus by the Dominated Convergence Theorem, the integral in (4–1) converges to $\int_{\mathbb{R}^k} f d\mu$ as $n \to \infty$, as desired. $\qquad\square$

Remark 4.1 Due to the factor of $\left(1 - \|x\|^2/n\right)^{(k+2)/2}$ in the denominator of (4–1), the Dominated Convergence Theorem does not directly work when we work with an unbounded function f, as there is no reason for $|f(x)| e^{-\|x\|^2/4}$ to be Lebesgue integrable in general. Indeed for a general Gaussian integrable f, we can bound $|f(x)| \left(1 - \|x\|^2/n\right)^{n/2}$ by $|f(x)| e^{-\|x\|^2/2}$, but there is still no obvious way to bound the whole integrand in (4–1) by a Lebesgue integrable function due to that extra factor in the denominator.

Corollary 4.2 *For $k \in \mathbb{N}$ and $N > \mathbb{N}$, almost all points on S_N have finite first k coordinates. That is,*

$$L\bar\sigma_N(\{(x_1,\ldots,x_N) \in S^{N-1}(\sqrt{N}) : x_1,\ldots,x_k \in {}^*\mathbb{R}_{fin}\}) = 1.$$

Proof Fix k and N as above. If $m \in \mathbb{N}$, we have $L\bar{\sigma}_N({}^*(-m, m)^k) = \mu((-m, m)^k)$ by Theorem 1.1. Letting $m \to \infty$ on both sides completes the proof. □

Corollary 4.3 *For any $t \in \mathbb{R}_{>1}$, we have:*

$$\lim_{n \to \infty} \int_{\{x \in \mathbb{R}^k : \frac{n}{t} < \|x\|^2 < n\}} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{4}} d\lambda(x) = 0$$

Proof Let $t \in \mathbb{R}_{>1}$ and $N > \mathbb{N}$. As a consequence of Corollary 4.2, we obtain:

$$\bar{\sigma}_N \left(\left\{ x \in S^{N-1}(\sqrt{N}) : \frac{N}{t} < \|\pi_k(x)\|^2 < N \right\} \right) \approx 0$$

The nonstandard characterization of limits and equation (4–1) thus yield the following.

$$(4\text{–}2) \qquad \lim_{n \to \infty} \int_{\{x \in \mathbb{R}^k : \frac{n}{t} < \|x\|^2 < n\}} \frac{1}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2}} \frac{1}{\left(1 - \frac{\|x\|^2}{n}\right)^{\frac{k+2}{2}}} d\lambda(x) = 0$$

For all $n \in \mathbb{N}_{\geq 2(k+2)}$, the sequence in the statement of the Corollary is bounded above by (a constant times) the sequence in (4–2), thus completing the proof. □

Remark 4.4 We can also prove Corollary 4.3 directly by noting that

$$\left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{4}} \mathbb{1}_{\|x\|^2 \leq n} \leq e^{-\frac{\|x\|^2}{4}}$$

where the right side is Lebesgue integrable over $\mathbb{R}^k$. The proof presented above is still valuable because it exposes a connection between these integrals and surface area measures.

4.2 A useful inequality between spherical and Gaussian measures

In this subsection, we derive an inequality comparing the L^1 norm (over the sphere $S^{n-1}(\sqrt{n})$) of a function defined on $\mathbb{R}^k$ and its pth moment (for any $p \in \mathbb{R}_{>1}$) with respect to the standard Gaussian measure on $\mathbb{R}^k$.

With the foresight provided by the philosophy of spherical integrals being close to a Gaussian integral, we expect these spherical integrals to be asymptotically bounded by the $L^p(\mathbb{R}^k, \mu)$–norms as the dimensions increase. Theorem 4.6 shows that depending on the value of $p \in \mathbb{R}_{>1}$, there is a dimension (namely $4(k+1)q$) beyond which this does happen. Before we prove that theorem, we need to generalize Sengupta's disintegration formula to work for any nonnegative function.

Theorem 4.5 *Let N and k be positive integers with $k < N$. Suppose f is either a bounded measurable or a nonnegative measurable function on $S^{N-1}(a)$, the sphere in $\mathbb{R}^N = \mathbb{R}^k \times \mathbb{R}^{N-k}$ of radius a and with center 0. Then, with σ denoting surface measure (non-normalized) on spheres,*

$$(4\text{--}3) \qquad \int_{z \in S^{N-1}(a)} f(z) d\sigma(z) = \int_{x \in B_k(a)} \left(\int_{y \in S^{N-k-1}(a_x)} f(x, y) d\sigma(y) \right) \frac{a}{a_x} dx$$

for any $a \in \mathbb{R}_{>0}$, where $a_x = \sqrt{a^2 - \|x\|^2}$. The above equality means that either both sides are finite and equal, or both sides are infinite.

Proof If f is bounded measurable, then this is just Sengupta's disintegration formula (see [20, Proposition 4.1]). Otherwise, if f is nonnegative then apply Sengupta's disintegration formula to the bounded measurable functions $f_m := f \cdot \mathbb{1}_{f \leq m}$ for each $m \in \mathbb{N}$, and then use the Monotone Convergence Theorem on both sides to obtain (4–3). $\qquad \square$

Theorem 4.6 *For each $p \in \mathbb{R}_{>1}$, there is a constant $C_p \in \mathbb{R}_{>0}$ such that*

$$(4\text{--}4) \qquad \int_{S^{n-1}(\sqrt{n})} |g| \, d\bar{\sigma}_n \leq C_p [\mathbb{E}_\mu(|g|^p)]^{\frac{1}{p}} \text{ for all } g \in L^p(\mathbb{R}^k, \mu) \text{ and } n \in \mathbb{N}_{>4(k+2)q}$$

where $q \in \mathbb{R}_{>0}$ is such that $1/p + 1/q = 1$.

Proof Fix $g \in L^p(\mathbb{R}^k, \mu)$, where $p \in \mathbb{R}_{>1}$. Also, let $t \in \mathbb{N}_{>1}$. Using Theorem 4.5 instead of [20, Proposition 4.1], we can follow the same steps leading up to (4–1) to see that $\int_{S^{n-1}(\sqrt{n})} |g| \, d\bar{\sigma}_n$ is equal to

$$(4\text{--}5) \qquad \int_{\|x\|^2 \leq \frac{n}{t}} \frac{a_{n,k} b_{n,k} |g(x)|}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n} \right)^{\frac{n-k-2}{2}} d\lambda(x)$$

$$+ \int_{\frac{n}{t} < \|x\|^2 \leq n} \frac{a_{n,k} b_{n,k} |g(x)|}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n} \right)^{\frac{n}{2}} \frac{1}{\left(1 - \frac{\|x\|^2}{n} \right)^{\frac{k+2}{2}}} d\lambda(x)$$

where $a_{n,k} = \dfrac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-k}{2}\right) \cdot \left(\frac{n-k}{2}\right)^{k/2}}$ and $b_{n,k} = \left(1 - \frac{k}{n} \right)^{k/2}$ are the same constants that appear in (4–1).

Note that:

$$\left(1 - \frac{\|x\|^2}{n} \right)^{-\frac{k+2}{2}} \leq \left(\frac{t}{t-1} \right)^{\frac{k+2}{2}} \text{ whenever } \|x\|^2 \leq \frac{n}{t}$$

Also, $\left(1 - \|x\|^2/n\right)^{n/2} \mathbb{1}_{\|x\|^2 \le n}$ is at most equal to $e^{-\|x\|^2/2}$ for all $x \in \mathbb{R}^k$. Noting that $b_{n,k} < 1$ for all $n \in \mathbb{N}_{>k}$, the first summand in (4–5) is at most

$$\left(\frac{t}{t-1}\right)^{\frac{k+2}{2}} \frac{a_{n,k}}{(2\pi)^{\frac{k}{2}}} \int_{\mathbb{R}^k} |g(x)| \, e^{-\frac{\|x\|^2}{2}} \, d\lambda(x)$$

for all $n \in \mathbb{N}_{>k}$. Writing this integral as a Gaussian expected value, and then using Jensen's inequality, we have

(4–6) $$I_1 \le a_{n,k} \left(\frac{t}{t-1}\right)^{\frac{k+2}{2}} \|g\|_{L^p(\mathbb{R}^k, \mu)} \text{ for all } n \in \mathbb{N}_{>k}$$

where I_1 is the first summand in (4–5), and $\|g\|_{L^p(\mathbb{R}^k, \mu)} = \left(\mathbb{E}_\mu(|g|^p)\right)^{1/p}$.

Let $q \in \mathbb{R}_{>1}$ be such that $1/p + 1/q = 1$. Then we can write the second summand in (4–5) as follows:

(4–7) $$a_{n,k} b_{n,k} \int_{\frac{n}{t} < \|x\|^2 \le n} \frac{|g(x)|}{(\sqrt{2\pi})^{\frac{k}{p}}} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2p}} \cdot \frac{1}{(\sqrt{2\pi})^{\frac{k}{q}}} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2q} - \frac{k+2}{2}} d\lambda(x)$$

Note that $b_{n,k} < 1$ for all $n \in \mathbb{N}_{>k}$. By Hölder's inequality applied to the functions

$$x \mapsto \frac{|g(x)|}{(\sqrt{2\pi})^{\frac{k}{p}}} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2p}} \qquad \text{and} \qquad x \mapsto \frac{1}{(\sqrt{2\pi})^{\frac{k}{q}}} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2q} - \frac{k+2}{2}}$$

(on the domain $\{x \in \mathbb{R}^k : n/t < \|x\|^2 < n\}$ equipped with its Lebesgue measure), the expression in (4–7) is at most equal to the following:[1]

$$a_{n,k} \left(\int_{\substack{x \in \mathbb{R}^k \\ \frac{n}{t} < \|x\|^2 \le n}} |g(x)|^p \cdot \frac{1}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{2}} d\lambda(x) \right)^{\frac{1}{p}}$$

$$\times \left(\int_{\substack{x \in \mathbb{R}^k \\ \frac{n}{t} < \|x\|^2 \le n}} \frac{1}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n}\right)^{\left(\frac{n}{2q} - \frac{k+2}{2}\right) \cdot q} d\lambda(x) \right)^{\frac{1}{q}}$$

---

[1]An anonymous referee has pointed out that one could also apply Hölder's inequality to the functions $x \mapsto |g(x)|$ and $x \mapsto \left(1 - \|x\|^2/n\right)^{-(k+2)/2}$ on the same domain but with the measure given by $d\nu(x) = \frac{1}{(\sqrt{2\pi})^k} \left(1 - \|x\|^2/n\right)^{n/2} d\lambda(x)$.

The first term in this product is at most $a_{n,k}\left(\mathbb{E}_\mu(|g|^p)\right)^{1/p}$. Also, the integrand in the second term in this product is at most $\left(1 - \|x\|^2/n\right)^{n/4}$ for all $n \in \mathbb{N}_{>2(k+2)q}$. To summarize, if I_2 is the second summand in (4–5), then we have

$$(4\text{--}8) \qquad I_2 \le a_{n,k}\|g\|_{L^p(\mathbb{R}^k,\mu)} \cdot \theta_{n,t} \text{ for all } n \in \mathbb{N}_{>(k+2)q}$$

where:

$$(4\text{--}9) \qquad \theta_{n,t} = \left(\int_{\substack{x\in\mathbb{R}^k \\ \frac{n}{t}<\|x\|^2\le n}} \frac{1}{(\sqrt{2\pi})^k} \left(1 - \frac{\|x\|^2}{n}\right)^{\frac{n}{4}} d\lambda(x) \right)^{\frac{1}{q}}$$

Combining (4–6) and (4–8), we get

$$(4\text{--}10) \qquad \int_{S^{n-1}(\sqrt{n})} |g|\, d\bar\sigma_n \le a_{n,k} \left[\left(\frac{t}{t-1}\right)^{\frac{k+2}{2}} + \theta_{n,t} \right] \|g\|_{L^p(\mathbb{R}^k,\mu)}$$

for all $n \in \mathbb{N}_{>4(k+2)q}$ and $t \in \mathbb{N}_{>1}$.

Here $a_{n,k} = \dfrac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\cdot\left(\frac{n-k}{2}\right)^{k/2}}$ and $\theta_{n,t}$ is as in (4–9). Note that $\lim_{n\to\infty} a_{n,k} = 1$, and by Corollary 4.3 $\lim_{n\to\infty} \theta_{n,t} = 0$ for all $t \in \mathbb{N}$. Thus, for any $t \in \mathbb{N}$, the coefficient of $\|g\|_{L^p(\mathbb{R}^k,\mu)}$ in (4–10) is uniformly bounded above, by (say) C_p. This completes the proof of the theorem. $\qquad\qquad\square$

Focusing on the coefficient in (4–10), we note that given $\epsilon \in \mathbb{R}_{>0}$ we can choose $t \in \mathbb{N}_{>1}$ large enough for which the following inequality holds:

$$\left(\frac{t}{t-1}\right)^{\frac{k+2}{2}} < 1 + \frac{\epsilon}{2}$$

For this t, using Corollary 4.3, we can choose an $n_p \in \mathbb{N}$ large enough such that $\theta_{n,t} < \frac{\epsilon}{2}$ for all $n \in \mathbb{N}_{>n_p}$. Since $\lim_{n\to\infty} a_{n,k} = 1$, we can also ensure that the n_p we choose is large enough such that $a_{n,k} < 1 + \epsilon$ for all $n \in \mathbb{N}_{>n_p}$. Combining all of this, (4–10) yields the following useful corollary: we are able to bound the ratio of the spherical integral and the Gaussian L^p norm by a constant as close to 1 as we want, with the price of having to go to a potentially higher dimension to observe this phenomenon.

Corollary 4.7 *For each $p \in \mathbb{R}_{>1}$ and $\epsilon \in \mathbb{R}_{>0}$, there is an $n_p \in \mathbb{N}$ such that:*

$$(4\text{--}11) \qquad \int_{S^{n-1}(\sqrt{n})} |g|\, d\bar\sigma_n \le (1+\epsilon)[\mathbb{E}_\mu(|g|^p)]^{\frac{1}{p}} \text{ for all } g \in L^p(\mathbb{R}^k,\mu) \text{ and } n \in \mathbb{N}_{>n_p}$$

Using Theorem 4.6, the condition (3–1) of Theorem 3.1 is easily verified for all functions in $L^p(\mathbb{R}^k, \mu)$, where $p \in \mathbb{R}^k$. Using that theorem and Theorem 1.1, we obtain our main limiting result for spherical integrals.

Theorem 4.8 *If μ is the standard Gaussian measure on $\mathbb{R}^k$ and $f \in L^p(\mathbb{R}^k, \mu)$ for some $p \in \mathbb{R}_{>1}$, then the nonstandard extension $^*f$ is S–integrable on $S^{N-1}(\sqrt{N})$ for all $N > \mathbb{N}$. As a consequence, the function f is integrable on $(S^{n-1}(\sqrt{n}), \bar{\sigma}_n)$ for all large $n \in \mathbb{N}$, and*

$$\lim_{m \to \infty} \lim_{n \to \infty} \int_{S^{n-1}(\sqrt{n}) \cap \{|f| \geq m\}} |f| \, d\bar{\sigma}_n = 0.$$

Furthermore, the spherical integrals of f satisfy the following limiting behavior:

$$\lim_{n \to \infty} \int_{S^{n-1}(\sqrt{n})} f d\bar{\sigma}_n = \int_{\mathbb{R}^k} f d\mu$$

This limit of spherical integrals can be written as a single spherical integral (over an infinite sphere) $\int_{S^{N-1}(\sqrt{N})} \mathbf{st}(^*f) dL\bar{\sigma}_N$ for any hyperfinite N.

A The kinetic theory of gases and spherical surface measures

This appendix is devoted to the physical motivation behind viewing a high-dimensional spherical integral as a Gaussian mean. We will give an outline of the usual derivation of the Maxwell–Boltzmann distribution (originally discovered by Maxwell in [12] and improved by Boltzmann in [4]), and explain its connection with the problem on limiting spherical integrals studied in this paper. We recommend Chapter 5 of Pauli and Enz [14] (which we also roughly follow for our outline) for more details on the underlying physics.

We work under the assumption that a statistically large number (which we shall denote by N) of particles of a monatomic gas are moving randomly in a container of a given volume. Each particle has a mass m. We further assume that the velocity of a given particle behaves like a random vector following an isotropic continuous probability density function $f \colon \mathbb{R}^3 \to \mathbb{R}$, where the isotropicity of f just means the following:

(A–1) $\exists g \colon \mathbb{R} \to \mathbb{R}$ such that $f(v_1, v_2, v_3) = g(v_1{}^2 + v_2{}^2 + v_3{}^2)$ for all $v_1, v_2, v_3 \in \mathbb{R}$

Newtonian mechanics can be used to postulate that the pressure on any wall of the container is directly proportional to the mean squared speed of the gas particles. Combining this with the ideal gas law, it then follows that the average kinetic energy of the

particles should be directly proportional to the temperature T of the system. This is typically described by the following equation, where $\vec{v}_i$ is the velocity of the ith particle, and k is a constant called the *Boltzmann constant*. Note that the factor of $3/2$ appears in the following in order to make sure that our k agrees with the traditional value of the Boltzmann constant:

$$(\text{A–2}) \qquad \sum_{i=1}^{N} \frac{1}{2}m\|\vec{v}_i\|^2 = \frac{3}{2}kTN, \text{ that is, } \frac{\sum_{i=1}^{N}\|\vec{v}_i\|^2}{N} = \frac{3kT}{m}$$

We also assume that the three components of the velocity vector of a given particle are independent and identically distributed, with a continuous density function $h\colon \mathbb{R} \to \mathbb{R}$ satisfying the following condition:

$$(\text{A–3}) \qquad f(v_1, v_2, v_3) = h(v_1)h(v_2)h(v_3) \text{ for all } v_1, v_2, v_3 \in \mathbb{R}.$$

We define new functions $\psi\colon \mathbb{R} \to \mathbb{R}$ and $\phi\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ by the following formulae:

$$\psi(v_i) := \log(h(v_i)) \text{ for all } v_i \in \mathbb{R}$$
$$\phi(v^2) := \log(g(v^2)) \text{ for all } v \in \mathbb{R}$$

Then ϕ and ψ satisfy the following functional equation:

$$(\text{A–4}) \qquad \phi(v_1^2 + v_2^2 + v_3^2) = \psi(v_1) + \psi(v_2) + \psi(v_3)$$

Assuming that ϕ and ψ are sufficiently differentiable, it can be shown that (A–4) can be satisfied only if ϕ is linear. After some simplifications, we obtain

$$(\text{A–5}) \qquad f(v_1, v_2, v_3) = g(v_1^2 + v_2^2 + v_3^2) = Ce^{-\alpha(v_1^2+v_2^2+v_3^2)}$$

for some constants $C, \alpha \in \mathbb{R}_{>0}$.

The constant C is obtained to be $(\alpha/\pi)^{3/2}$ by integrating both sides of (A–5) and noting that the integral of f is equal to 1 as f is a probability density function. We then compute the expected value of the square of the speed $v_1^2 + v_2^2 + v_3^2$, and equate it with $3kT/m$ (which comes from (A–2), using our underlying hypothesis of N being statistically large so that the mean of the individual particles' squared speed should be very close to the theoretical expected value – more precisely, one can let $N \to \infty$ and use the Strong Law of Large Numbers). From that, we find $3/(2\alpha) = 3kT/m$, so that $\alpha = m/(2kT)$. We thus obtain the famous Maxwell–Boltzmann distribution for velocity:

$$(\text{A–6}) \qquad f(v_1, v_2, v_3) = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-\frac{m}{2kT}(v_1^2+v_2^2+v_3^2)} \text{ for all } v_1, v_2, v_3 \in \mathbb{R}$$

From the above formula, Maxwell and Boltzmann proceeded to derive probability distributions of other important functions (such as speed) of velocity. These distributions are heavily used in statistical mechanics and thermodynamics.

The problem of statistically estimating the behavior of a function of the velocity of a random gas particle can be reinterpreted in a useful way with the notion of surface area measures on Euclidean spheres. For simplicity of terms we let $N_0 := 3N$, and renormalize the constants in equation (A–2) (by assuming that $kT = m$). Writing $\vec{v}_i = (v_{i,x}, v_{i,y}, v_{i,z}) \in \mathbb{R}^3$, we then get:

$$(A\text{–}7) \qquad \sum_{i=1}^{N} \left(v_{i,x}{}^2 + v_{i,y}{}^2 + v_{i,z}{}^2 \right) = N_0$$

Hence $(\vec{v}_1, \ldots, \vec{v}_N)$ is a vector in $\mathbb{R}^{N_0}$ of norm $\sqrt{N_0}$. In other words, $(\vec{v}_1, \ldots, \vec{v}_N)$ is an element of $S^{N_0-1}(\sqrt{N_0})$. Since we do not have any information about the motion of these particles other than what is contained in equation (A–2), it is reasonable to assume that the value of $(\vec{v}_1, \ldots, \vec{v}_N)$ at a given time is a "random point" of $S^{N_0-1}(\sqrt{N_0})$. The surface area measure $\bar{\sigma}_S$ for a sphere S serves as a notion of a uniform probability measure on S. Thus we can make the observation regarding $(\vec{v}_1, \ldots, \vec{v}_N)$ being a random point of $S^{N_0-1}(\sqrt{N_0})$ more precise by postulating that the probability that $(\vec{v}_1, \ldots, \vec{v}_N)$ lies in a Borel set $B \subseteq S^{N_0-1}(\sqrt{N_0})$ is given by $\bar{\sigma}_{S^{N_0-1}(\sqrt{N_0})}(B)$.

Since we are working under the assumption that the number of particles is very large, the probability that the first component of the velocity of the first particle, and hence of a random particle (by symmetry), is in a Borel set $B_1 \subseteq \mathbb{R}^1$, should be given by $\lim_{N_0 \to \infty} \bar{\sigma}((B_1 \times \mathbb{R}^{N_0-1}) \cap S^{N_0-1}(\sqrt{N_0}))$. Also, the expected or mean value of the first component of its velocity should be given by the following integral:

$$\lim_{N_0 \to \infty} \int_{S^{N_0-1}(\sqrt{N_0})} v_{1,x} d\bar{\sigma}(v_{1,x}, v_{1,y}, v_{1,z}, \ldots, v_{N,x}, v_{N,y}, v_{N,z})$$

Similarly, the expected value of speed would be given by the limit of the integrals of $\sqrt{v_{1,x}^2 + v_{1,y}^2 + v_{1,z}^2}$. In fact, the limit of integrals of any finite-dimensional function on these spheres can be interpreted as the expected value of some function of velocities of randomly chosen particles in our gaseous system.

If there were a way to directly compute these limits, then we would be able to evaluate various probabilities associated with values taken by the velocity components, as well as recover the expected values of many functions of velocities of the particles. Furthermore, such a derivation would have the benefit of being less circular as we would not be making any assumptions on the nature (or even existence) of the density f that was derived in (A–6).

Thus the problem of generalizing Theorem 1.1 to the largest class of functions possible is intimately connected to, and has implications on our understanding of the kinetic theory of gases. Furthermore, the fact that there already exist distributions for functions of velocity such as speed (which, being equal to $\sqrt{v_1{}^2 + v_2{}^2 + v_3{}^2}$, is clearly not a bounded function) suggests that (1.1) should, in principle, be generalizable to at least some unbounded functions, which in turn makes the problem of finding all such functions naturally appealing.

Mathematically, (1–2) tells us that the Gaussian measure μ is well-equipped to measure the limiting expected value of any bounded measurable function of a given collection of coordinates. In some sense, it retains all probabilistic information of the manner in which such functions behave over these spheres in the large-N limit. From this point of view as well, it becomes a natural question to find out for which actual functions does it retain all such information.

Nonstandard analysis gives access to hyperfinite natural numbers which provide a natural model for statistically large number of particles. The probability that the velocity of a random particle lies in some set could actually be thought of as the uniform surface area of the portion of a hyperfinite-dimensional sphere corresponding to this set.

Acknowledgments

References

[1] **I Alam**, *Limiting spherical integrals of bounded continuous functions*, (2019); `arXiv: 1702.01556`

[2] **S Albeverio**, **R Høegh-Krohn**, **J E Fenstad**, **T Lindstrøm**, *Nonstandard methods in stochastic analysis and mathematical physics*, volume 122 of *Pure and Applied Mathematics*, Academic Press, Inc., Orlando, FL (1986)

[3] **R M Anderson**, *A non-standard representation for Brownian motion and Itô integration*, Israel J. Math. 25 (1976) 15–46; https://doi.org/10.1007/BF02756559

[4] **L Boltzmann**, *Studien über das Gleichgewicht der lebenden Kraft*, Wissenschafiliche Abhandlungen 1 (1868) 49–96; https://doi.org/10.1017/cbo9781139381420.006

[5] **N Cutland**, **S-A Ng**, *The Wiener sphere and Wiener measure*, Ann. Probab. 21 (1993) 1–13; https://doi.org/10.1214/aop/1176989390

[6] **G B Folland**, *Real analysis*, second edition, Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York (1999)

[7] **D N Hoover**, *Relations on probability spaces and arrays of random variables*, Preprint, Institute for Advanced Study, Princeton, NJ 2 (1979); https://www.stat.berkeley.edu/ al-dous/Research/hoover.pdf

[8] **H J Keisler**, *Hyperfinite model theory*, from: "Logic Colloquium 76 (Oxford, 1976)" (1977) 5–110. Studies in Logic and Found. Math., Vol. 87; https://doi.org/10.1016/S0049-237X(09)70421-5

[9] **P Lévy**, *Leçons d'analyse fonctionnelle*, Gauthier-Villars Paris (1922)

[10] **J E Littlewood**, *Lectures on the Theory of Functions*, Oxford University Press (1944)

[11] **P Mattila**, *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge (1995); https://doi.org/10.1017/CBO9780511623813

[12] **J C Maxwell**, *V. Illustrations of the dynamical theory of gases.—Part I. On the motions and collisions of perfectly elastic spheres*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 19 (1860) 19–32; https://doi.org/10.1080/14786446008642818

[13] **H P McKean**, *Geometry of differential space*, Ann. Probability 1 (1973) 197–206; https://doi.org/10.1214/aop/1176996973

[14] **W Pauli**, **C Enz**, *Thermodynamics and the Kinetic Theory of Gases*, Dover books on physics, Dover Publications (2000)

[15] **A Peterson**, **A N Sengupta**, *The Gaussian limit for high-dimensional spherical means*, J. Funct. Anal. 276 (2019) 815–866; https://doi.org/10.1016/j.jfa.2018.06.020

[16] **H Poincaré**, *Calcul des probabilités*, Gauthier-Villars (1912)

[17] **D A Ross**, *Loeb measure and probability*, from: "Nonstandard analysis (Edinburgh, 1996)", NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 493, Kluwer Acad. Publ., Dordrecht (1997) 91–120; https://doi.org/10.1007/978-94-011-5544-1_4

[18] **W Rudin**, *Principles of mathematical analysis*, third edition, McGraw-Hill Book Co., New York-Auckland-Düsseldorf (1976)

[19] **W Rudin**, *Real and complex analysis*, third edition, McGraw-Hill Book Co., New York (1987)

[20] **A N Sengupta**, *The Gaussian Radon transform as a limit of spherical transforms*, J. Funct. Anal. 271 (2016) 3242–3268; https://doi.org/10.1016/j.jfa.2016.07.002

[21] **N Wiener**, *Differential-Space*, Journal of Mathematics and Physics 2 (1923) 131–174; https://doi.org/10.1002/sapm192321131

Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA.

irfanalamisi@gmail.com

www.math.lsu.edu/~ialam1