CHANGE OF VARIABLE FOR HAUSDORFF MEASURE (from the beginning) (*)

by TIM TRAYNOR (in Windsor)(**)

1. Hausdorff Measure.

In many of the lectures of the Schools on Measure and Real Analysis, Hausdorff measure is used to discuss sets of fractional dimension. Here, instead, we wish to discuss sets of integral dimension: curves, surfaces, solids, et cetera, and see how Hausdorff measure makes it possible to discuss arc length, surface area (and surface integral) and the like in a consistent manner.

The point of view taken is that (for integral dimension) Hausdorff measure is what Lebesgue measure would be if it were only defined. Let us start with the set $K = [0,1] \times [0,1] \times \{1\} \subset \mathbb{R}^3$, a surface in $\mathbb{R}^3$. It clearly has volume 0, area 1, and length $+\infty$. The reason it is "clear" that it has area 1 is that K is a "copy" of the 2-dimensional square $[0,1] \times [0,1]$ which has Lebesgue 2-dimensional measure $\boldsymbol{\lambda}^2([0,1] \times [0,1]) = 1$. It has length $+\infty$ because it contains infinitely many disjoint copies of the interval $[0,1]$ which has 1-dimensional Lebesgue measure 1. And, of course, its volume is its 3-dimensional Lebesgue measure, which is 0.

Let $\boldsymbol{\lambda}^n$ be Lebesgue outer measure in $\mathbb{R}^n$. It is **Borel regular** in the sense that Borel sets are measurable and each $A \subset \mathbb{R}^n$ is contained in a Borel set B of the same measure. The other (outer) measures which we discuss will be also of this type. We will identify the measure on the Borel sets with the corresponding outer measure.

A basic property of Lebesgue (outer) measure $\boldsymbol{\lambda}^n$ in $\mathbb{R}^n$ — and the basis for this entire article — is that it is translation invariant:

(*) Presentato al "Workshop di Teoria della Misura e Analisi Reale", Grado (Italia), 19 settembre-2 ottobre 1993.

(**) Indirizzo dell'Autore: Department of Mathematics, University of Windsor, Windsor, Ontario, N9B 3P4 (Canada). E-mail: TT@Uwindsor.ca

Basic Theorem. *For any* $a \in \mathbb{R}^n$, $\boldsymbol{\lambda}^n(a + A) = \boldsymbol{\lambda}^n(A)$, *for all* $A \subset \mathbb{R}^n$. *Every other translation invariant measure on Borel subsets of* $\mathbb{R}^n$ *giving a finite value* c *to a unit cube is of the form* $\mu(A) = c\boldsymbol{\lambda}^n(A)$.

Hence, also every translation invariant Borel regular outer measure giving a finite value to a unit cube is a multiple of Lebesgue outer measure.

What do we mean, when we say "copy"? If $A \subset \mathbb{R}^n$, and $B \subset \mathbb{R}^m$, to say that B is a copy of A (for the purposes of length and area, and the like) surely means that there is a one-to-one function T mapping A onto B, and at the very least, that for two points $x, y \in A$ the Euclidean distance between the corresponding points Tx and Ty is the same as the distance between x and y; that is, T is an **isometry**. One might also want lines and planes to be preserved, but this is automatic.

1.1. Theorem. *A map* T *on* $\mathbb{R}^n$ *to* $\mathbb{R}^m$ *is an isometry iff it is an orthogonal (linear) transformation, followed by a translation.*

By definition, $U : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is **orthogonal** if $Ux \cdot Uy = x \cdot y$, for all $x, y \in \mathbb{R}^n$. Linearity is a consequence of this, for if we let $x = U(sa + b)$ and $y = sUa + Ub$, we find that $|x|^2 = |y|^2 = x \cdot y$, hence $x = y$ from the identity $|x - y|^2 = |x|^2 + |y|^2 - 2x \cdot y$.

Proof. Since an orthogonal transformation is linear and preserves length, it preserves distance; since translations also do so, any T of the stated form is an isometry.

Conversely, let T be an isometry and put $Ux = Tx - T0$, for all x. Then U is still an isometry, and $U0 = 0$. Thus,

$$2x \cdot y = |x|^2 + |y|^2 - |x - y|^2 = |Ux|^2 + |Uy|^2 - |Ux - Uy|^2 = 2Ux \cdot Uy,$$

so U preserves inner products and $Tx = T0 + Ux$, is of the required form. $\diamondsuit$

Consequently:

1.2. THEOREM. *Lebesgue measure is invariant under isometries of $\mathbb{R}^n$ with itself.*

Proof. If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isometry then it is the composition of a linear map with a translation, hence $\mu(A) = \boldsymbol{\lambda}^n(TA)$ defines a translation invariant measure. Thus, $\boldsymbol{\lambda}^n(TA) = c\boldsymbol{\lambda}^n(A)$. If B is a ball centered at 0, then TB is just the translate $T0 + B$. Since $\boldsymbol{\lambda}^n$ is translation invariant, $\boldsymbol{\lambda}^n(TB) = \boldsymbol{\lambda}^n(B)$, which shows $c = 1$. $\diamondsuit$

The program, then, is to produce a measure $\mathbf{h}^n$ in $\mathbb{R}^m$ which gives the measure $\boldsymbol{\lambda}^n(A)$ to each isometric copy TA of a (Borel) set $A \subset \mathbb{R}^n$. And then to investigate its behaviour under a change of variable.

Producing $\mathbf{h}^n$

(The reader who already knows something about Hausdorff measure may prefer to skip the rest of this section.)

Since such a measure also must be invariant under isometries, one makes its definition depend only on diameters of sets. In a metric space, we will use the letter d to denote both distance and diameter: $d(A) = \inf\{d(x, y) : x, y \in A\}$.

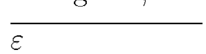
Consider the case $n = 1$, first. Lebesgue measure on the line satisfies

$$\boldsymbol{\lambda}^1(A) = \inf\{\sum_{B \in \mathcal{C}} d(B) : \mathcal{C} \text{ is countable, } \bigcup \mathcal{C} \supset A\}.$$

This is because the diameter of a finite interval is its length and each bounded set B is covered by an interval of the same diameter. Thus, we might expect the same formula

$$\mu(A) = \inf\{\sum_{B \in \mathcal{C}} d(B) : \mathcal{C} \text{ is countable, } \bigcup \mathcal{C} \supset A\}.$$

would give the right concept of length (1 dimensional measure) in higher dimensions. But, if K is $[0, 1] \times \{0, \varepsilon\}$, the union of two parallel line segments in $\mathbb{R}^2$, each of length 1, of distance ε apart,



then $\mu(K) \le d(K) = (1+\varepsilon^2)^{\frac{1}{2}}$, so for ε small, $\mu(K)$ is considerably less than the desired value of 2. Worse, if $A = [0,1] \times \{\frac{1}{n} : n \in \mathbb{N}\}$, then $\mu(A) \le d(A) = \sqrt{2}$, whereas A is a disjoint union of denumerably many disjoint isometric images I_n of the unit interval, so the measure $\mathbf{h}^1$ to be constructed would have $\mathbf{h}^1(A) = \sum_n \mathbf{h}^1(I_n) = +\infty$.

This difficulty is overcome by covering only with small sets.

For each $\delta > 0$, call a family $\mathcal{C}$ of sets a δ-cover of A if $\bigcup \mathcal{C} \supset A$ and $d(C) \le \delta$, for all $C \in \mathcal{C}$ and put

$$\mathbf{h}^1_\delta(A) = \inf \left\{ \sum_{C\in\mathcal{C}} d(C) : \mathcal{C} \text{ a countable } \delta\text{-cover of } A \right\}.$$

As δ decreases to 0, $\mathbf{h}^1_\delta(A)$ increases to a limit, denoted $\mathbf{h}^1(A)$, the 1-**dimensional Hausdorff measure** of A.

This set function is definable in every metric space, and of course, depends on the metric space; this, however, is not a weakness but a strength. The various outer measures $\mathbf{h}^1$ are preserved under isometries. Indeed, if f is an isometry between metric spaces then $d(C) = d(f(C))$ for all C and hence $\mathbf{h}^1(A) = \mathbf{h}^1(f(A))$. (If f is not surjective, the argument is slightly more complicated. See Lemma below.)

The process used to construct $\mathbf{h}^1$ is called Carathéodory's second construction (or Method II). (See the appendix.) It always produces an outer measure for which Borel sets are measurable. And we shall see that $\mathbf{h}^1$ is also Borel regular.

Thus, we have solved the problem for dimension 1: If X is $\mathbb{R}^m$, or any metric space, and $T : \mathbb{R} \longrightarrow X$ is an isometry, then $\mathbf{h}^1(TA) = \mathbf{h}^1(A) = \boldsymbol{\lambda}^1(A)$, for all $A \subset \mathbb{R}$.

Since the volume of a cube in $\mathbb{R}^n$ is proportional to the n^{th} power of its diameter, we guess that to solve the problem for n dimensional objects in a metric space X we should involve $d(A)^n$ instead of $d(A)$. Put

$$\begin{aligned}
\mathcal{H}^n_\delta(A) &= \inf \left\{ \sum_{C\in\mathcal{C}} d(C)^n : \mathcal{C} \text{ a countable } \delta\text{-cover of } A \right\}, \\
\mathcal{H}^n(A) &= \lim_{\delta \longrightarrow 0} \mathcal{H}^n_\delta(A) \\
\mathbf{h}^n(A) &= c_n \mathcal{H}^n(A)
\end{aligned}$$

where c_n is a positive finite constant, to be chosen. For many purposes it is convenient to choose $c_n = 1$. [Bandt, Graf, Edgar, Falconer].

Here, instead, we wish to choose c_n so that in $\mathbb{R}^n$, $\mathbf{h}^n$ will equal $\boldsymbol{\lambda}^n$ [Federer, Billingsley, Evans and Gariepy].

1.3. Lemma. *If $A \subset X_0 \subset X$, then to calculate $\mathcal{H}^n_\delta(A)$, $\mathcal{H}^n(A)$, $\mathbf{h}^n(A)$, it is enough to use covers consisting of subsets of X_0.*

Proof. Indeed, if $\mathcal{C}$ is a δ-cover of A, then so is $\{C \cap X_0 : C \in \mathcal{C}\}$ and $d(C \cap X_0) \leq d(C)$, for all $C \in \mathcal{C}$. ♢

1.4. Theorem. *The outer measures $\mathcal{H}^n$, hence also $\mathbf{h}^n$, are invariant under isometries from one metric space to another and are Borel regular. The constant c_n may be chosen so that for $A \subset \mathbb{R}^n$, $\mathbf{h}^n(A) = \boldsymbol{\lambda}^n(A)$.*

With this choice, we call $\mathbf{h}^n$ **n-dimensional Hausdorff measure**. Since $\mathbf{h}^n$ is invariant under isometries, no confusion arises from using the same letter for this measure in different spaces.

Proof. Let X and Y be metric spaces, $f : X \longrightarrow Y$ an isometry and $A \subset X$. We prove that $\mathcal{H}^n_\delta(A) = \mathcal{H}^n_\delta(f(A))$. Since $f(A) \subset f(X)$, by the previous lemma, we may assume $f(X) = Y$. Then, since $d(f(C)) = d(C)$, $\mathcal{C} \leftrightarrow \{f(C) : C \in \mathcal{C}\}$ is a one-to-one correspondence between the δ-covers of A and of $f(A)$, and the sums in the definitions remain invariant.

The closure of any set has the same diameter. Thus, in the definition of $\mathcal{H}^n(A)$, one may restrict the δ-covers to consist of closed sets. For each $k \in \mathbb{N}$, let $\mathcal{C}_k$ be a countable $\frac{1}{k}$-cover of A by closed sets with $\sum_{C \in \mathcal{C}_k} d(C)^n \leq \mathcal{H}^n_{\frac{1}{k}}(A) + \frac{1}{k}$. Then $B := \bigcap_k \bigcup \mathcal{C}_k$ is a Borel set containing A with $\mathcal{H}^n_{\frac{1}{k}}(A) \leq \mathcal{H}^n_{\frac{1}{k}}(B) \leq \mathcal{H}^n_{\frac{1}{k}}(A) + \frac{1}{k}$. In the limit, we have $\mathcal{H}^n(A) = \mathcal{H}^n(B)$. Thus, $\mathcal{H}^n$ is Borel regular.

Let Q be the unit cube. Since $\mathcal{H}^n$ is translation invariant it will, by the basic theorem, be a multiple of $\boldsymbol{\lambda}^n$ if $\mathcal{H}^n(Q) < \infty$.

Now, Q can be divided into a finite number of disjoint cubes C of diameter $\leq \delta$, and for each C, $d(C)^n = (\sqrt{n})^n \boldsymbol{\lambda}^n(C)$; hence,

$\mathcal{H}^n_\delta(Q) \leq (\sqrt{n})^n$, and in the limit $\mathcal{H}^n(Q) \leq (\sqrt{n})^n < \infty$.

Thus, $\mathcal{H}^n(A) = c\boldsymbol{\lambda}^n(A)$, $c = \mathcal{H}^n(Q) < \infty$, for all Borel sets A. Since both $\mathcal{H}^n$ and $\boldsymbol{\lambda}^n$ are Borel regular, this is true also for non-Borel sets.

On the other hand, each set C in $\mathbb{R}^n$ is contained in a cube of side $d(C)$, so if $\mathcal{C}$ is a cover of Q, then

$$\boldsymbol{\lambda}^n(Q) \leq \sum_{C \in \mathcal{C}} d(C)^n,$$

so that $1 \leq \mathcal{H}^n(Q) = c$.

Thus, we may choose $c_n = c^{-1}$ to obtain $\mathbf{h}^n(A) = \boldsymbol{\lambda}^n(A)$ for all $A \subset \mathbb{R}^n$. $\diamondsuit$

It happens that this value $c_n = \mathcal{H}^n(Q)^{-1}$ is just the Lebesgue measure of a ball of diameter 1. (See the Appendix.) But we won't need this fact.

1.5 Proposition. *If A_1 and A_2 are subsets of $\mathbb{R}^m$ contained in orthogonal vector subspaces X_1 and X_2 of dimension n and k respectively, then $\mathbf{h}^{n+k}(A_1 \oplus A_2) = \mathbf{h}^n(A_1)\mathbf{h}^k(A_2)$.*

Proof. There is an isometry U mapping $X_1 \oplus X_2$ onto $\mathbb{R}^n \times \mathbb{R}^k$ of the form, $U(x+y) = (U_1x, U_2y)$. Since U_i restricted to X_i is an isometry and $A_i \subset X_i$, we have $\mathbf{h}^{n+k}(A_1 \oplus A_2) = \mathbf{h}^{n+k}(U_1A_1 \times U_2A_2) = \boldsymbol{\lambda}^{n+k}(U_1A_1 \times U_2A_2) = \boldsymbol{\lambda}^n(U_1A_1)\boldsymbol{\lambda}^k(U_2A_2) = \mathbf{h}^n(U_1A_1)\mathbf{h}^k(U_2A_2) = \mathbf{h}^n(A_1)\mathbf{h}^k(A_2)$. $\diamondsuit$

2. The Area Formula — Change of Variable

Fix $n \in \mathbb{N}$, and let $\mathbf{h}^n$ denote n-dimensional measure in whatever space we are in. If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we wish to find how to calculate $\mathbf{h}^n(f(A))$ in terms of $\mathbf{h}^n = \boldsymbol{\lambda}^n$ on $\mathbb{R}^n$, and to integrate over $f(A)$.

The formula $\mathbf{h}^n f(A) = \mathbf{h}^n(f(A))$ clearly defines an outer measure, though without further conditions, one cannot say much about its measurable sets.

If f is an isometry, then as we have seen, $\mathbf{h}^n f(A) = \mathbf{h}^n(A)$. If f is Lipshitz with Lipshitz constant c, then $\mathbf{h}^n f(A) \leq c^n \mathbf{h}^n(A)$. More generally:

2.1. LEMMA. *Let f and g be map a set A into (possibly different) metric spaces such that $d(f(x), f(y)) \leq cd(g(x), g(y))$, for all $x, y \in A$. Then $\mathbf{h}^n f(A) \leq c^n \mathbf{h}^n g(A)$.*

Proof. If $\mathcal{C}$ is a δ-cover of $g(A)$, then the $fg^{-1}(B)$, for $B \in \mathcal{C}$ cover $f(A)$ and have diameter $d(fg^{-1}(B)) \leq cf(B)$,

$$\mathcal{H}^n_{c\delta}(f(A)) \leq \sum_i c^n d(B).$$

Taking inf over such covers gives

$$\mathcal{H}^n_{c\delta}(fA) \leq c^n \mathcal{H}^n_\delta(g(A)).$$

Now pass to the limit and multiply by the normalizing constant. $\diamondsuit$

Consider first a linear map $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Its smallest Lipshitz constant is $\|T\|$, the operator norm. Thus, $\mathbf{h}^n T(A) \leq \|T\|^n \mathbf{h}^n(A) = \|T\|^n \boldsymbol{\lambda}^n(A) < \infty$, for all $A \subset \mathbb{R}^n$.

Let $\|T\|$ be $\mathbf{h}^n(T[Q])$ where Q is the unit cube in $\mathbb{R}^n$.

2.2. PROPOSITION. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then,*

$$\mathbf{h}^n T(A) = \|T\| \mathbf{h}^n(A),$$

for all $A \subset \mathbb{R}^n$. T is injective iff $\|T\| \neq 0$.

Proof. If T is not one-to-one, then its range is a vector subspace of dimension $k < n$. Thus, there is an isometry f mapping the range of T to a k dimensional subspace of $\mathbb{R}^n$, and as such has Lebesgue n-dimensional measure 0. Hence, $\mathbf{h}^n T(A) = \mathbf{h}^n(f(TA)) = \boldsymbol{\lambda}^n(f(TA)\}) = 0$. In particular, $\|T\| = \mathbf{h}^n T(Q) = 0$, so the formula holds in this case.

In the case T is one-to-one, $\{A : TA \text{ is Borel }\}$ is a σ-ring $\mathcal{S}$. Since T is continuous, TA is compact for each compact $A \subset \mathbb{R}^n$.

Thus, $\mathcal{S}$ contains the compacts of $\mathbb{R}^n$, hence all Borel sets of $\mathbb{R}^n$. Since all the Borel sets of $\mathbb{R}^m$ are measurable for the version of $\mathbf{h}^n$ in $\mathbb{R}^m$, this shows that the restriction of $\mathbf{h}^n T$ to the Borel sets is a measure which is translation invariant, and finite on Q. Therefore, by the basic theorem,

$$\mathbf{h}^n T(B) = \boldsymbol{|}T\boldsymbol{|}\,\mathbf{h}^n(B),$$

for B Borel. Since $\mathbf{h}^n$ is Borel regular, this will hold for non-Borel sets if $\mathbf{h}^n T$ is also Borel regular.

If A is arbitrary, there is a Borel subset $C \supset TA$ of $\mathbb{R}^m$, with $\mathbf{h}^n(C) = \mathbf{h}^n(TA)$. Then $B = T^{-1}[C]$ is Borel, $B \supset A$, and $C \supset TB \supset TA$. Thus, $\mathbf{h}T(B) = \mathbf{h}T(A)$, as required.

Finally, if T is injective, $T\mathbb{R}^n$ is isometric with $\mathbb{R}^n$. Thus, $\mathbf{h}^n(T\mathbb{R}^n) = \mathbf{h}^n(\mathbb{R}^n) = \infty$, so $\boldsymbol{|}T\boldsymbol{|} \neq 0$. $\diamondsuit$

The **scale factor** $\boldsymbol{|}T\boldsymbol{|}$ can be calculated in terms of determinants (and we will do this in section 3), but this is irrelevant to the main result.

For a map $f : V \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$, put $\mathbf{Jf}(\mathbf{x}) = \boldsymbol{|}\mathbf{f}'(\mathbf{x})\boldsymbol{|}$, whenever the (total) derivative $f'(x)$ exists. We will call $\mathbf{Jf}$ the **Jacobian** of f, because it turns out in the case $n = m$ to be the absolute value of the Jacobian determinant. Notice, by the way that $f'(x)$ is injective iff $\mathbf{Jf}(\mathbf{x}) \neq \mathbf{0}$.

Let $L(\mathbb{R}^n, \mathbb{R}^m)$ denote the normed space of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

2.3. Lemma. *Let $T_0 \in L(\mathbb{R}^n, \mathbb{R}^m)$.*

(1) If T_0 is injective, then for each $r > 1$, there exists $\delta > 0$ such that for all $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ with $\|T - T_0\| < \delta$,

$$\frac{1}{r}|T_0 z| \leq |Tz| \leq r|T_0 z|$$

for all $z \in \mathbb{R}^n$, so that $\frac{1}{r}\boldsymbol{|}T_0\boldsymbol{|} \leq \boldsymbol{|}T\boldsymbol{|} \leq r\boldsymbol{|}T_0\boldsymbol{|}$.

(2) If T_0 is not injective (i.e. $\boldsymbol{|}T_0\boldsymbol{|} = 0$), then $\|T - T_0\| < \delta$ implies

$$\boldsymbol{|}T\boldsymbol{|} \leq \delta(\|T_0\| + \delta)^{n-1}.$$

(3) Hence, the scale factor $\boldsymbol{|}T\boldsymbol{|}$ is a continuous function of T.

Proof. (1) Since T_0 is continuous, $|T_0x|$ takes a minimum value $c = c_{T_0}$ on the (compact) sphere $\{x \in \mathbb{R}^n : |x| = 1\}$, and since T_0 is injective, $c \neq 0$. Thus, for all $z \in \mathbb{R}^n$, $|T_0 z| \geq c|z|$. For any other $T \in L(\mathbb{R}^n, \mathbb{R}^m)$,

$$|(T - T_0)z| \leq \|T - T_0\||z| \leq c^{-1}\|T - T_0\||T_0 z|.$$

Thus,

$$\begin{aligned} |Tz| \leq |T_0 z| + |Tz - T_0 z| &\leq (1 + c^{-1}\|T - T_0\|)|T_0 z| \leq r|T_0 z|, \\ |Tz| \geq |T_0 z| - |Tz - T_0 z| &\geq (1 - c^{-1}\|T - T_0\|)|T_0 z| \geq r^{-1}|T_0 z|, \end{aligned}$$

for $\|T - T_0\| < \delta$, small enough. By Lemma 2.1, $\boldsymbol{|}T\boldsymbol{|} = \mathbf{h}^n T(Q) \leq r^n \mathbf{h}^n T_0(Q) = r^n \boldsymbol{|}T_0\boldsymbol{|}$ and similarly $\boldsymbol{|}T_0\boldsymbol{|} \leq r^n \boldsymbol{|}T\boldsymbol{|}$.

(2) If T_0 is not injective, let v_1 be a unit vector with $T_0 v_1 = 0$. Extend this to an orthonormal basis $\{v_1, \ldots, v_n\}$. Let Q_1 be the unit cube determined by this basis. And let Q_2 be the $n-1$ dimensional cube determined by $\{v_2, \ldots, v_n\}$.

Suppose $\|T - T_0\| < \delta$. Then

$$|Tv_1| = |(T - T_0)v_1| < \delta,$$

and $\|T\| \leq \|T_0\| + \delta$.

Let P be the projection onto the subspace orthogonal to Tv_1. Then $\|PT\| \leq \|T\| \leq \|T_0\| + \delta$. And TQ_1 is contained in $[0,1]Tv_1 \oplus PTQ_2$, so that by Proposition 1.5, $\mathbf{h}^n(TQ_1) \leq \mathbf{h}^1([0,1]Tv_1)\mathbf{h}^{n-1}(PTQ_2) = |Tv|\|PT\|^n < \delta(\|T_0\| + \delta)^n$.

(3) follows immediately from (1) and (2). $\diamondsuit$

2.4. Lemma. *Let f be differentiable at $x \in \mathbb{R}^n$, with $f'(x)$ injective. Then, for each $r > 1$, there exist $\delta > 0$ and $k \in \mathbb{N}$ such that for $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ with $\|T - f'(x)\| < \delta$,*

$$|f(x) - f(y)| \leq r|Tx - Ty|, \qquad |Tx - Ty| \leq r|f(x) - f(y)|$$

for all x, y with $|x - y| < \frac{1}{k}$.

Proof. Let $r > 1$ and choose r_0 with $r > r_0 > 1$. Apply the previous lemma to r_0 and $T_0 = f'(x)$, obtaining a $\delta > 0$ with

$$\frac{1}{r_0}|T_0 z| \le |Tz| \le r_0 |T_0 z|$$

for $\|T - f'(x)\| < \delta$. In particular, such T is also injective. Thus, there is $c_T > 0$ with $|Tz| \ge c_T|z|$ for all $z \in \mathbb{R}^n$. Fix $\varepsilon > 0$. By definition of differentiable, there exists $k \in \mathbb{N}$ such that $|y - x| < \frac{1}{k}$ implies

$$|f(y) - f(x) - f'(x)(y - x)| \le \varepsilon c_T |y - x|,$$

$$\begin{aligned} |f(y) - f(x)| &\le |f'(x)(y - x)| + \varepsilon c_T |y - x| \\ &\le r_0 |T(y - x)| + \varepsilon |T(y - x)| \\ &= (r_0 + \varepsilon)|Ty - Tx|, \end{aligned}$$

and similarly, $|f(y) - f(x)| \ge (\frac{1}{r_0} - \varepsilon)|Ty - Tx|$. The conclusion follows, by choosing ε so that $r_0 + \varepsilon < r$ and $\frac{1}{r_0} - \varepsilon > \frac{1}{r}$.

2.5. LEMMA. *Let $f : V \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be continuously differentiable, V an open set and $r > 1$. Then $V_0 := \{x \in V :$ $f'(x)$ is injective $\}$ is a disjoint union of a sequence Borel sets B_i for which there exists $T = T_i$ such that for all $x, y \in B_i$*

$$|Tz| \le r|f'(x)z|, \qquad |f'(x)z| \le r|Tz|, \qquad \textit{for all } z \in \mathbb{R}^n \qquad (*)$$

and

$$|f(x) - f(y)| \le r|Tx - Ty|, \qquad |Tx - Ty| \le r|f(x) - f(y)|. \quad (**)$$

Proof. Since f is continuously differentiable, and the set of injective T is open, V_0 is open, hence Borel. [1]

Fix a countable dense subset $\mathcal{T}$ of $L(\mathbb{R}^n, \mathbb{R}^m)$. For each $x \in V_0$, we may choose by the previous two lemmas, a $T \in \mathcal{T}$ and $k \in \mathbb{N}$ satisfying $(*)$ and satisfying $(**)$ for y such that $|x - y| < \frac{1}{k}$. Turn

1 Actually, V_0 is a Borel set even if f is just continuous [Fed, p 211].

this around: for fixed T, k let $E(T,k)$ be the set of those x for which these inequalities hold. Then the $E(T,k)$ cover V_0.

Each $E(T,k)$ is a Borel set. Indeed, $x \in E(T,k)$ iff the inequalities hold for all z in a countable dense subset of $\mathbb{R}^n$ and for all y in a countable dense subset of $\{y : |x-y| < \frac{1}{k}\}$. Moreover, $E(T,k)$ is a countable union of Borel sets $E(T,k,j)$ of diameter $< \frac{1}{k}$. The required countable set of B_i comes from disjointifying the countable family $\{V_0 \cap E(T,k,j) : T \in \mathcal{T}, k, j\}$. $\diamondsuit$

2.6. Main Result. *Let f be a 1-1 continously differentiable map of an open set V of $\mathbb{R}^n$ to $\mathbb{R}^m$. Then for all Borel subsets A of V,*

$$\int_A \mathbf{J}f(x)\, \boldsymbol{\lambda}^n(dx) = \mathbf{h}^n(f[A]).$$

and hence,

$$\int_A g(f(x))\mathbf{J}f(x)\, \boldsymbol{\lambda}^n(dx) = \int_{f[A]} g(y)\, \mathbf{h}^n(dy).$$

for all Borel maps $g : f[V] \longrightarrow \mathbb{R}$ for which one side exists.

On the left side the integration is with respect to Lebesgue measure $\boldsymbol{\lambda}^n = \mathbf{h}^n$; on the right side, with respect to Hausdorff n-dimensional measure in $\mathbb{R}^m$.

Proof. The second equation (the change of variable formula) comes follows from the first (the area formula) by the usual argument envolving approximation by simple functions, so we omit it.

Suppose first that $A \subset V_0 = \{x : f'(x) \text{ is injective }\}$. Fix $r > 1$ and choose sets B_i and linear maps T_i as in the previous lemma. Put $A_i = A \cap B_i$. From $(*)$ and Lemma 2.1,

$$\mathbf{h}^n(T_iQ) \leq r^n \mathbf{h}^n(f'(x)Q), \qquad \mathbf{h}^n(f'(x)Q) \leq r^n \mathbf{h}^n(T_iQ)$$

that is,

$$\| T_i \| \leq r^n \mathbf{J}f(x), \qquad \mathbf{J}f(x) \leq r^n \| T_i \|$$

for $x \in A_i$. Integrating over A_i yields

$$\frac{1}{r^n}\mathbf{h}T(A_i) \leq \int_{A_i} \mathbf{J}f \, d\mathbf{h}^n \leq r^n \mathbf{h}T(A_i).$$

Let us work on the right side. From (∗∗) and 2.1 again,

$$\mathbf{h}^n T_i(A_i) \le r^n \mathbf{h}^n f(A_i).$$

Thus,

$$\int_{A_i} \mathbf{J} f \, d\mathbf{h}^n \le r^{2n} \mathbf{h}^n f(A_i),$$

and after summation over i,

$$\int_A \mathbf{J} f \, d\mathbf{h}^n \le r^{2n} \mathbf{h}^n f(A).$$

Now A no longer depends on r, so we may let $r \longrightarrow 1$, to obtain

$$\int_A \mathbf{J} f \, d\mathbf{h}^n \le \mathbf{h}^n f(A).$$

The reverse inequality is similar.

We now consider the case $A \subset \{x : f'(x) \text{ is not 1–1 }\}$. Since $\mathbf{J} f = 0$ on A, $\int_A \mathbf{J} f \, d\mathbf{h}^n = 0$. Thus, we are to show that $\mathbf{h}^n f(A)$ is also 0.

Since $f'(x)$ is continuous in x and $\mathbb{R}^n$ is σ-compact, we may assume $\boldsymbol{\lambda}^n(A) < \infty$ and $\|f'(x)\|$ is bounded by some $K \in \mathbb{R}$. Factor f as $P \circ f_1$ where

$$f_1 : \mathbb{R}^n \longrightarrow \mathbb{R}^m \times \mathbb{R}^n : x \mapsto (fx, \delta x)$$

and $P(y, z) = y$. Then f_1 is injective and continuously differentiable with

$$f_1'(x)v = (f'(x)v, \delta v), \text{ for all } v \in \mathbb{R}^n.$$

We see that $f_1'(x)$ is injective. The difference between $T = f_1'(x)$ and the linear map $T_0 : v \mapsto (f'(x)v, 0)$ is $v \mapsto (0, \delta v)$, which has norm δ, and $\|T_0\| = \|f'(x)\| \le K$, so

$$\mathbf{J} f_1(x) \le \delta (K + \delta)^{n-1},$$

by Lemma 2.3(2). The previous case applies to f_1, so

$$\mathbf{h}^n f(A) \le \mathbf{h}^n f_1(A) = \int_A \mathbf{J} f_1 \, d\mathbf{h}^n \le \delta (K + \delta)^{n-1},$$

and hence $\mathbf{h}^n f(A) = 0$. ♢

REMARKS.

(1) This result was proved without actual evaluation of $\mathbf{J}f$. In applications, one may need actual values. This is the topic of the next section.

(2) Federer[Fed] and [EvGa] state this theorem, not for C^1 functions on an open set, but for Lipschitz functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. By Rademacher's Theorem [Fed, 3.1.6] such an f is differentiable at $\boldsymbol{\lambda}^n$-almost all points.

(3) The hypothesis that f be one-to-one may be removed (and is so in [Fed] and [EvGa]) if one counts the multiplicity of f, that is, the number of times f takes a given value y in A. In this regard, note that the proof divided A into a sequence of disjoint pieces, on each of which f was one-to-one.

(4) The measurability assumptions on A and g can also be relaxed.

3. Evaluating the scale factor and Jacobian

First notice that if $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, linear, $\| ST \| = \boldsymbol{\lambda}^n(STQ) = \| S \| \boldsymbol{\lambda}^n(TQ) = \| S \| \| T \|$, where Q as before denotes a unit cube.

Case $n = m$.

Here we will show that $\| T \| = |\det T|$, so that $\mathbf{J}f(x)$ is the absolute value of the usual Jacobian determinant.

QUICK PROOF. T may be factored as $U_1 D U_2$ where U_1, U_2 are orthogonal and D is diagonal with entries $d_1, \ldots, d_n \geq 0$ on the diagonal of its standard matrix. (See A.7.) Then, since $\| D \| = \boldsymbol{\lambda}^n([0, d_1] \times \ldots \times [0, d_n]) = d_1 \ldots d_n = \det D$, $\| T \| = \| D \| = |\det D| = |\det T|$. ♢

Elementary proof. If T is not invertible, we already know $\| T \| = 0 = |\det T|$.

Every invertible linear transformation T on $\mathbb{R}^n$ can be factored as a product (composition) $T = E_k \dots E_1$, where the E_i are "elementary" linear tranformations of the forms:

E_i interchanges two coordinates;
E_i multiplies a co-ordinate by -1;
E_i multiplies the first co-ordinate by $c > 0$;
E_i adds the second co-ordinate to the first.

Since $T \mapsto \|T\|$ and $T \mapsto |\det T|$ are both multiplicative, it is enough to establish the equality for T of one of these types. The first two are isometries (scale factor 1) and have determinant 1 and -1; and if T is of the multiplication type: $(x_1, x_2, \dots, x_n) \mapsto (cx_1, x_2, \dots, x_n)$, $c > 0$, then $\det T = c$ and $\|T\| = \boldsymbol{\lambda}^n(TQ) = \boldsymbol{\lambda}^n([0, c] \times [0, 1] \times \dots \times [0, 1]) = c$. Finally, if $T : (x_1, x_2, \dots, x_n) \mapsto (x_1 + x_2, x_2, \dots, x_n)$, let $S : (x_1, x_2, \dots, x_n) \mapsto (x_1, -x_2, \dots, x_n)$. Then $STST$ is the identity and $\|S\| = 1$, so $1 = \|S\|\|T\|\|S\|\|T\| = \|T\|^2$. Hence $\|T\| = 1 = \det T$. $\diamondsuit$

Case $n = 1$

Here $T : \mathbb{R}^1 \longrightarrow \mathbb{R}^m$ is of the form $Tx = xv$, $v \in \mathbb{R}^m$ and $TQ = T[0, 1]$ is isometric with $[0, |v|]$, so $\|T\| = \boldsymbol{\lambda}^1([0, |v|]) = |v|$. Thus, $\mathbf{h}^1(f(A)) = \int_A |f'(x)|\, dx$, extending the usual formula for arc length of a 1-1 differentiable curve.

Case $n = m - 1$

Let $\{e_1, \dots, e_n\}$ be the standard basis in $\mathbb{R}^n$ and $\{e'_1, \dots, e'_m\}$ the standard basis in $\mathbb{R}^m$. View the elements of $\mathbb{R}^n$ and of $\mathbb{R}^m$ as column vectors. Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be linear with standard matrix $[T] = (a_{ij}) = [a_1|a_2|\dots|a_n]$, $a_j = Te_j$. The map $x \in \mathbb{R}^m \mapsto \det[x|a_1|\dots|a_n]$ is linear, so there exists a vector $v \in \mathbb{R}^m$ such that $v \cdot x = \det[x|a_1|\dots|a_n]$. For each i, $v_i = v \cdot e'_i = \det[e'_i|a_1|\dots|a_n]$ is $(-1)^{1+i}$times the determinant obtained by deleting the j^{th} row. We may denote v by $a_1 \times \dots \times a_n$, since for $n = 2$, this is the ordinary cross product of two vectors in R^3.

3.1. THEOREM. *If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}$ is linear, then* $\| T \| = |Te_1 \times \ldots \times Te_n|$.

If v is not 0, let u be a unit vector in the same direction; otherwise, let u be any unit vector orthogonal to the $a_1, \ldots, a_n$. Then, $|v| = |v \cdot u|$ the absolute value of the determinant of $[u|a_1| \ldots a_n]$. This latter is the matrix of a linear transformation on $S : \mathbb{R}^m \longrightarrow \mathbb{R}^m$

$$Se'_1 = u_1, Se'_i = a_{i-1}, \text{ for } i = 2, \ldots, m.$$

Let C be the unit cube of $\mathbb{R}^m$. Then

$$SC = [0,1]u \oplus TQ,$$

so

$$\begin{aligned} \| S \| &= \mathbf{h}^m([0,1]u \oplus TQ) \\ &= \mathbf{h}^1([0,1]u)\mathbf{h}^n(TQ) \\ &= \mathbf{h}^n(TQ) = \| T \|, \end{aligned}$$

and by the case $n = m$, $\| S \| = |\det S| = |v|$, as promised.

Thus to obtain $\mathbf{J}f(x)$, one finds the matrix of partials

$$[f'(x)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \ldots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \ldots & \frac{\partial f_m}{\partial x_n} \end{bmatrix},$$

and finds the square root of the sum of the squares of the submatrices of size $n \times n$.

The general case $n \leq m$

The last paragraph of the preceding case is really the general case.

3.2. THEOREM. *If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear then* $\| T \|^2 = \det(T^*T)$ *and is the sum of the squares of the $n \times n$ sub-determinants of the matrix of T.*

Here $T^* : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ is the ajoint map: for $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $Tx \cdot y = x \cdot T^* y$. The matrix of T^* is the transpose of the matrix of T.

Proof. According to the Polar Decomposition Theorem (A.4), T may be factored as US where S is a symmetric transformation on $\mathbb{R}^n$ and U is an orthogonal map on $\mathbb{R}^n$ to $\mathbb{R}^m$. Thus U^*U is the identity on $\mathbb{R}^n$ and $\|T\|^2 = \|S\|^2 = (\det S)^2 = (\det S^*)(\det S) = \det(S^*S) = \det(S^*U^*US) = \det(T^*T)$.

Let $C(m,n)$ be the set of all increasing maps τ from $\{1,\ldots,m\}$ to $\{1,\ldots,n\}$ and $\mathbf{S}_n$ be the set of permutations σ of $\{1,\ldots,n\}$. Every injective map on $\{1,\ldots,n\}$ to $\{1,\ldots,m\}$ may be uniquely factored as $\tau\sigma$.

Now, let the matrix of T be $A = (a_{ij})$. For each $\tau \in C(m,n)$, let A_τ be the matrix obtained from A by selecting the rows $\tau(1),\ldots,\tau(n)$. Then, since a determinant is linear in each of its columns, and vanishes if two columns coincide,

$$\det T^*T = \det \begin{bmatrix} \sum_{k=1}^m a_{k1}a_{k1} & \cdots & \sum_{k=1}^m a_{k1}a_{kn} \\ \vdots & & \vdots \\ \sum_{k=1}^m a_{kn}a_{k1} & \cdots & \sum_{k=1}^m a_{kn}a_{kn} \end{bmatrix}$$

$$= \sum_{k_1=1}^m \sum_{k_2=1}^m \cdots \sum_{k_n=1}^m a_{k_1 1} \cdots a_{k_n n} \det \begin{bmatrix} a_{k_1 1} & \cdots & a_{k_n 1} \\ \vdots & & \vdots \\ a_{k_1 n} & \cdots & a_{k_n 1} \end{bmatrix}$$

$$= \sum_{\tau \in C(m,n)} \sum_{\sigma \in \mathbf{S}_n} a_{\tau\sigma(1),1} \cdots a_{\tau\sigma(n),n} \det \begin{bmatrix} a_{\tau\sigma(1),1} & \cdots & a_{\tau\sigma(n),1} \\ \vdots & & \vdots \\ a_{\tau\sigma(1),n} & \cdots & a_{\tau\sigma(n),1} \end{bmatrix}$$

$$= \sum_{\tau \in C(m,n)} \sum_{\sigma \in \mathbf{S}_n} a_{\tau\sigma(1),1} \cdots a_{\tau\sigma(n),n} \operatorname{sgn}(\sigma) \det \begin{bmatrix} a_{\tau(1),1} & \cdots & a_{\tau(n),1} \\ \vdots & & \vdots \\ a_{\tau(1),n} & \cdots & a_{\tau(n),1} \end{bmatrix}$$

$$= \sum_\tau \det(A_\tau)\det(A_\tau),$$

as required. $\Diamond$

The $\tau \in C(m,n)$ determine the $\binom{m}{n}$ subspaces $E_\tau = \operatorname{Span}\{e'_{\tau(1)}, \ldots, e'_{\tau(n)}\}$, spanned by subsets of the standard basis of $\mathbb{R}^m$. Let P_τ be the corresponding projection. Then the A_τ of the previous proof is just the matrix of $P_\tau T$ with respect to the bases $\{e_1, \ldots, e_n\}$ and $\{e'_{\tau(1)}, \ldots, e'_{\tau(1)}\}$.

3.3. Corollary. *If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear, then $\|T\|^2 = \sum_{\tau \in C(m,n)} \|P_\tau T\|^2$; that is $\mathbf{h}^n T(A)^2 = \sum_\tau \mathbf{h}^n P_\tau T(A)^2$.*

Proof. Since E_τ is isometric with $\mathbb{R}^n$ under the map U_τ which sends $e'_{\tau(i)}$ to e_i, $\|P_\tau T\| = \|U_\tau P_\tau T\| = |\det(U_\tau P_\tau T)| = |\det A_\tau|$. $\diamondsuit$

Appendix

An **outer measure** (sometimes called simply "measure") in a space X is a countably subadditive function φ on all subsets of X to $[0, +\infty]$ with $\varphi(\emptyset) = 0$. $A \subset X$ is called **φ-measurable** if $\mu(T) = \mu(T \cap A) + \mu(T \cap A^c)$, for all $T \subset X$. The set $\mathcal{M}_\varphi$ of all φ-measurable sets is a σ-algebra of subsets of X on which φ is σ-additive.

If τ is any function defined on a family $\mathcal{A}$ of subsets of a space X with values in $[0, \infty]$, and

$$\tau^*(A) = \inf\{\sum_{C \in \mathcal{C}} \tau(C) : \mathcal{C} \subset \mathcal{A},\ \mathcal{C} \text{ is countable, and } \quad \bigcup \mathcal{C} \supset A\},$$

for all $A \subset X$, then τ^* is an outer measure, the **Carathéodory (Method I) outer measure generated** by τ and $\mathcal{C}$. ($\tau^*(\emptyset)$ is automatically 0, since empty sums are considered 0.)

A.1. Carathéodory's criterion. *If μ is an outer measure in a metric space X, all Borel sets are μ-measurable iff $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever A and B are of positive distance apart: $d(A, B) = \inf\{d(x, y) : x \in A,\ y \in B\} > 0$.*

Such a μ is called a **metric outer measure**. *Proof.* Let G be open, $T \subset X$. We have to show

$$\mu(T) \geq \mu(T \cap G) + \mu(T \setminus G)$$

Let $A = T \cap G$, and for each $i \in \mathbb{N}$ put $A_i = \{x \in A : d(x, G^c) > \frac{1}{i}\}$. Then $A_i \nearrow A$ and $d(A_i, T \setminus G) > 0$, so

$$\mu(T) \geq \mu(A_i \cup (T \setminus G)) = \mu(A_i) + \mu(T \setminus G).$$

Also, $d(A_i, A \setminus A_{i+1}) > 0$, so the result will follow from:

A.2. Lemma. *If μ is an outer measure $A_i \nearrow A$ and $\mu(E) = \mu(E \cap A_i) + \mu(E \setminus A_{i+1})$, whenever $E \subset A_i \cup (A \setminus A_{i+1})$, $i \in \mathbb{N}$, then $\mu(A_i) \longrightarrow \mu(A)$.*

Proof of the Lemma. Since $\mu(A) \geq \lim_i \mu(A_i)$, we may assume the right side is finite.

Write $B_1 = A_1$, and $B_{i+1} = A_{i+1} \setminus A_i$, for all $i \in \mathbb{N}$. Then $\bigcup_{k \leq m} B_{2k-1} \subset A_{2m-3} \cup (A \setminus A_{2m-2})$, so $\mu(\bigcup_{k \leq m} B_{2k-1}) = \mu(\bigcup_{k \leq m-1} B_{2k-1}) + \mu(B_{2m-1})$ and by induction

$$\mu(\bigcup_{k \leq m} B_{2k-1}) = \sum_{k \leq m} \mu(B_{2k-1}).$$

Since $\lim_i \mu(A_i)$ majorizes this, the series $\sum_k \mu(B_{2k-1})$ converges. Similarly, the series of even terms converges, so the entire series $\sum_i \mu(B_i)$ converges. Thus

$$\mu(A) \leq \mu(A_j) + \sum_{i > j} \mu(B_j) \longrightarrow \lim_i \mu(A_i),$$

as required. $\diamondsuit$

Carathéodory's second construction. "Method II", (of which the construction of Hausdorff measure is a special case) goes as follows. Given τ a function on a family $\mathcal{A}$ of subsets of X, a metric space, for each $\delta > 0$, let μ_δ be the Carathéodory outer measure generated by the restricion of τ to $\mathcal{A}_\delta = \{C \in \mathcal{A} : d(C) \leq \delta\}$. Then for all $A \subset X$, $\delta \geq \delta'$ implies $\mu_\delta(A) \leq \mu_{\delta'}(A)$ and $\mu(A) = \lim_{\delta \longrightarrow 0} \mu_\delta(A) = \sup_{\delta > 0} \mu_\delta(A)$ defines an outer measure.

A.3. Theorem. *Carathéodory's second construction produces only metric outer measures.*

Proof. Indeed, if A and B have distance $> \delta$ apart and a countable $\mathcal{C} \subset \mathcal{A}_\delta$ covers of $A \cup B$, then $\mathcal{C}_A = \{C \in \mathcal{C} : C \cap A \neq \emptyset\}$ and $\mathcal{C}_B = \{C \in \mathcal{C} : C \cap B \neq \emptyset\}$ are covers of A and B respectively, which have no C in common. Thus,

$$\sum_{C \in \mathcal{C}} \tau(C) \geq \sum_{C \in \mathcal{C}_A} \tau(C) + \sum_{C \in \mathcal{C}_B} \tau(C) \geq \mu_\delta(A) + \mu_\delta(B).$$

Taking infimum over all such covers $\mathcal{C}$ gives $\mu_\delta(A \cup B) \geq \mu_\delta(A) + \mu_\delta(B)$. Thus equality holds, and in the limit $\mu(A \cup B) = \mu(A) + \mu(B)$.

Determination of the normalizing constant

If u is a unit vector of $\mathbb{R}^n$, $A \subset \mathbb{R}^n$, the **Steiner symmetrization** of A with respect to the hyperplane $u^\perp = \{x : x \cdot u = 0\}$ is $S_u A = \{x + tu : |t| \leq \frac{1}{2}\mathbf{h}^1(A \cap (x + \mathbb{R}u)\})$. Here $x + \mathbb{R}u$ is the line through x in the direction of v.

A.4. Lemma.

(1) $d(S_u A) \leq d(A)$

(2) If A is $\boldsymbol{\lambda}^n$-measurable then so is $S_u(A)$ with $\boldsymbol{\lambda}^n(S_u A) = \boldsymbol{\lambda}^n(A)$.

(3) S_u is symmetric with respect to $u^\perp$.

This depends upon Fubini's Theorem in the form of Cavalieri's Principle. We omit the details. (See [Fed, 2.10.30], [Bil, p.211], [EvGa, p.67].)

A.5. Isodiametric inequality. For each $A \subset \mathbb{R}^n$ of diameter $2r$, $\boldsymbol{\lambda}^n(A) \leq \boldsymbol{\lambda}^n(B(0,r))$.

That is, even though A may not be contained in a ball of the same diameter, its measure is at most the measure of such a ball.

Proof. We may assume A is closed, hence measurable. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$ and $A' = S_{e_n} \cdots S_{e_2} S_{e_1} A$.

Then A' is symmetric about the origin, has diameter $\leq d(A) = 2r$, and measure $\boldsymbol{\lambda}^n(A') = \boldsymbol{\lambda}^n(A)$. The symmetry implies A' is contained in the closed ball about 0 radius r so $\boldsymbol{\lambda}^n(A) = \boldsymbol{\lambda}^n(A') \leq \boldsymbol{\lambda}^n(B(0,r))$, as required. $\diamondsuit$

A.6. Theorem. *For all $A \subset \mathbb{R}^n$, $\boldsymbol{\lambda}^n(A) = c_n \mathcal{H}^n(A)$, where c_n is the Lebesgue measure of a ball diameter 1.*

Proof. By 1.4, there exists a $c_n > 0$ with this property; it is a matter of determining its value. Let k be the measure of an open ball of diameter 1. Then the measure of every other ball B is $\boldsymbol{\lambda}^n(B) = kd(B)^n$; and by the isodiametric inequality, every set C has $\boldsymbol{\lambda}^n(C) \leq kd(C)^n$.

Let U be an open set, $\delta > 0$. By the Vitali covering theorem there exists a sequence of disjoint balls B_i of diameter $< \delta$ with $\mathcal{H}^n_\delta(U \setminus \bigcup_i B_i) \leq c_n^{-1} \boldsymbol{\lambda}^n(U \setminus \bigcup_i B_i) = 0$. Thus

$$k\mathcal{H}^n_\delta(U) = k\mathcal{H}^n_\delta(\bigcup_i B_i) \leq k \sum_i d(B_i)^n = \sum_i \boldsymbol{\lambda}^n(B_i) = \boldsymbol{\lambda}^n(U),$$

so also $k\mathcal{H}^n(U) \leq \boldsymbol{\lambda}^n(U)$.

On the other hand, if $U \subset \bigcup_i C_i$, $\boldsymbol{\lambda}^n(U) \leq \sum_i \boldsymbol{\lambda}^n(C_i) \leq$ $\leq \sum_i kd(C_i)^k$, which shows that $k\mathcal{H}^n(U) \geq \boldsymbol{\lambda}^n(U)$. Thus, $c_n = k$.$\diamondsuit$

A.7. Polar decomposition. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, then there exist orthogonal linear $U : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and symmetric positive $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with $T = US$, hence also orthogonal U_1, U_2 and D positive diagonal with $T = U_1 D U_2$.*

Diagonal means $D(x_1, \ldots, x_n) = (d_1 x_1, \ldots, d_n x_n)$.

Proof. The operator T^*T is positive: $T^*Tx \cdot x \geq 0$, for all x. Thus, there exist $\alpha_1, \ldots, \alpha_n \geq 0$, and an orthonomal basis $u_1, \ldots, u_n$ with $T^*Tu_i = \alpha_i u_i$. Put $d_i = \sqrt{\alpha_i}$. Since $Tu_i \cdot Tu_j = T^*Tu_i \cdot u_j = \alpha u_i \cdot u_j$, the Tu_i are orthogonal of norm d_i. Put $v_i = d_i^{-1} Tu_i$, for those i with $d_i > 0$, and extend this to an orthonormal basis $\{v_1, \ldots, v_m\}$ of $\mathbb{R}^m$. Define a linear $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, by setting $Su_i = d_i u_i$; then $U : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by setting $Uu_i = v_i$, for $i = 1, \ldots, n$. Then $USu_i = U(d_i u_i) = d_i v_i = Tu_i$, so $US = T$. Finally, a further orthogonal change of basis $Pe_i = u_i$, gives us the $P^{-1}SP$ diagonal yielding the last statement: $T = UP(P^{-1}SP)P^{-1}$.

References

[Ba91b] BANDT C., *Deterministic fractals and fractal measures*, Rendiconti dell'Istit. di Mat., Trieste **23** (1991), 1-40.

[Bil] BILLINGSLEY, *Probability and measure*, Wiley (1979).

[Ed90] EDGAR G.A., *Measure, Topology, and Fractal Geometry*, Springer-Verlag, ISBN 0-387-97272-2, New York (1990).

[EvGa] EVANS L.C. and GARIEPY R.F., *Measure Theory and the Fine Properties of Functions*, CRC Press. Boca Raton. ISBN0-8493-7157-0 (1992).

[Fa85] FALCONER K.J., *The Geometry of Fractal Sets*, ISBN 0-521-33705-4, Cambridge University Press (1985).

[Fed] FEDERER H., *Geometric Measure Theory*, Springer-Verlag, Berlin (1969).

[Gr91] GRAF S., *Random Fractals*, Rendiconti dell'Istit. di Mat., Trieste **23** (1991), 81-144.