Introduction

Here, in nearly full detail, is the Theorem that outer measure is not changed by orthogonal changes of coordinates. The last Lemma in the note is important for later use. We begin with effect of linear maps on outer measure.

Theorem *Suppose that* $T : \mathbb{R}^n \to \mathbb{R}^n$ *is linear. Then for every set* $E \subseteq \mathbb{R}^n$,

$$|TE|_e = |\det T||E|_e.$$

Proof: We will use a Lemma to bring the proof down to some calculations.

Lemma *Suppose that* $T : \mathbb{R}^n \to \mathbb{R}^n$ *is linear. Then for every interval* $I \subseteq \mathbb{R}^n$,

$$|TI| = |\det T||I|.$$

We can use measure here in place of outer measure because linear mappings are Lipschitz transformations.

We will use the Lemma first to prove the Theorem.

First we suppose that T is invertible. To estimate $|E|_e$ in terms of $|TE|_e$ we suppose that $TE \subseteq \bigcap_j J_j$. Then (by the Lemma)

$$E \subseteq \bigcap_j T^{-1} J_j \;\text{ so }\; |E|_e \leq \sum_j |T^{-1}J_j| = |\det T^{-1}| \sum_j |J_j|.$$

Hence by taking the infimum on the right, we have $|E|_e \leq |\det T|^{-1} |TE|_e$. We can now reverse the rôles of E and TE to show that $|TE|_e \leq |\det T||E|_e$.

If T is not invertible, so is T^T, and we will not use the Lemma. There is a non-zero unit vector x_o such that $T^T x_o = 0$. We show next that

$$T\mathbb{R}^n \subseteq \{y \in \mathbb{R}^n : y \perp x_o\}.$$

Given that $y = Tx$, $\langle y, x_o\rangle = \langle Tx, x_o\rangle = \langle x, T^T x_o\rangle = 0$. Thus for any $E \subseteq \mathbb{R}^n$, TE is contained in a hyperplane.

The set on the right is a hyperplane, and every hyperplane is a translate of such a set. Our next task is to show that the measure of every hyperplane is zero, and this will complete the proof in case $\det T = 0$ (why?) .

Some $x_{0m} \neq 0$, so for all $w \in \mathbb{R}^n$, $w =: w' + w_m e_m$. Then if $y \perp x_0$, $y_m = (-x_0/x_{0m}) \cdot y' =: g(y')$, so the *graph* of g in $\mathbb{R}^n$ coincides with the graph of g' as a function from *an* $\mathbb{R}^{n-1}$ into $\mathbb{R}$, and g' is continuous.

Lemma: *If* $g : \mathbb{R}^n \to \mathbb{R}$ *is continuous then the graph of* g *in* $\mathbb{R}^{n+1}$ *has measure zero in* $\mathbb{R}^{n+1}$.

Proof: It is enough to show that the part of the graph "above" the cube Q with edge-length 1 and center 0 has measure zero, for the same argument shows (except for scaling) that the same is true for the part of the graph "above" a cube with edge-length 2^k, and the entire graph is the countable union of such "parts."

Let $\epsilon > 0$ be given. Since g is uniformly continuous on Q there exists $\delta > 0$ such that for all x and y in Q, $|x - y| < \delta \Rightarrow |g(x) - g(y)| < \epsilon$. If we now dyadically subdivide Q so many times that the *diameter* of a subdivision-cube I is less than δ, we see that (with c denoting the center of I) the part of the graph of g "above" I is contained in the interval $I \times [g(c) - \epsilon, g(c) + \epsilon]$, which has measure (volume) $2\epsilon|I|$. These intervals are non-overlappinmg, so the measure of their union is the same as the sum of the measures of the cubes I, which is 1, times the length, 2ϵ, of each interval $[g(c) - \epsilon, g(c) + \epsilon]$. Since $\epsilon > 0$ is arbitrary the part of the graph "above" the cube Q with edge-length 1 and center 0 has measure zero, and the Lemma follows.

Now we see that for every E, $TE \subseteq T\mathbb{R}^n \subseteq H$, where H is a hyperplane, and we have shown that every hyperplane has measure zero. Thus the Theorem and the Lemma hold when $\det T = 0$.

It remains to show that the Lemma holds when T is invertible. We begin with three special types of linear transformations. The first and third types require that $n > 1$.

First type: "switch around a bunch of rows" T is given by a permutation matrix $P = (\,[j = pi]\,)$, where p is a permutation of the integers $\{1, \ldots, n\}$, namely a map of that set to itself that is one-to-one and onto. The action of P on a column vector x is permutation of the coordinates of x. That is, the ith coordinate of Px is x_{pi}. Then, if I is an interval, PI is the interval obtained by permuting the factor intervals of I, so $v(PI) = v(I)$. Since the determinant of P is ± 1, not only is the Lemma true in this case, so is the Theorem, by the argument using the Lemma to prove the Theorem.

Second type: "multiply a row by a non-zero scalar" T is given by a matrix of the form $I + (k-1)e_\ell e_\ell^T$, where $k \neq 0$. The matrix is the identity matrix, modified by replacing the ℓth diagonal entry by k. Here if I is an interval, TI is the interval obtained by replacing the ℓth factor interval by $[ka_\ell,\, kb_\ell]$ if $k > 0$ and by $[kb_\ell,\, ka_\ell]$ if $k < 0$ (the other factor intervals are unchanged). Thus $v(TI) = |k|v(I)$ and $|k| = \det T$, so not only is the Lemma true in this case, so is the Theorem, by the argument using the Lemma to prove the Theorem.

Third type: "add a multiple of one row to a different row" T is given by a matrix of the form $I + \alpha e_k e_\ell^T$, where $k \neq \ell$, known as a *shear*. A linear operator of this form acts on a matrix A as follows: the k-th row of TA is the sum of the k-th row of A and α times the ℓ-th row of A. That is, T adds α times row ℓ to row k. The other rows of TA agree with the corresponding rows of A. This case is more complex than the others, for now, if I is an interval and $\alpha \neq 0$, TI is not an interval. We will express TI as the union of certain sets and translate some of them. The union of the new sets will be I, except possibly for part of the boundary of I, and the union will be an non-overlapping union. This will allow us to say that $|TI| = |I|$. Since $\det T = 1$ for such matrices, this will complete the proof of this case.

We may assume that $k = 1$ and $\ell = 2$ (we can work with $P_{1k}TP_{2\ell}$ and as we will see later, this is enough). Let us write $I = K \times L$ where $K = [a_1,\, b_1] \times [a_2,\, b_2]$ and L is the Cartesian product of the other factor intervals, if $n > 2$. Then

$$TI = S_\alpha K \times L, \quad \text{where} \quad S_\alpha = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}.$$

We know that $|TI| = |S_\alpha K||L|$, where the measures on the right are of subsets in $\mathbb{R}^2$ and $\mathbb{R}^{n-2}$ (here L is present only if $n > 2$). The proof of this has been moved to the end of this note. We will work with $\alpha > 0$ first (the other case is similar). By translation invariance we may assume that $K = [0,\, w] \times [0,\, h]$ where h and w are positive. Then $S_\alpha K$ is the parallelogram with vertices $Z := (0,\, 0)$, $A := (w,\, 0)$, $B := (w + \alpha h,\, h)$, and $C := (\alpha h,\, h)$, listed in counter-clockwise order. We will work with the interior of this parallelogram, and call it P. Since Lipschitz transformations map sets of measure zero into sets of measure zero, $|P| = |S_\alpha K|$. If $\alpha h \leq w$ we define points $A_o := (\alpha h,\, 0)$ and $B_o := (w,\, h)$. Then $P = T_1 \cup R \cup T_2$, where T_1 is the right triangle ZA_oC (with its hypotenuse and horizontal leg missing), R is the rectangle A_oBB_oC (with its upper and lower edges missing) and T_2 is the triangle ABB_o (with its hypotenuse and horizontal leg missing). We have

$$K^o = [(T_2 - we_1) \cup T_1 \cup R \cup h_1] \setminus (v_2 - we_1),$$

where h_1 is the hypotenuse of T_1 and v_2 is the vertical leg of T_2. The union is disjoint and h_1 and v_2 are sets of measure zero because they are subsets of sets of measure zero. This gives

$$|K| = |K^o| = |T_2 - we_1| + |T_1| + |R| = |T_2| + |T_1| + |R| = |P| = |S_\alpha K|.$$

If $\alpha h \leq w$ the segment CZ is no longer contained in K. The height H at which the segment leaves K allows us to write K as the union of some congruent intervals K' plus (possibly) one interval K'' of shorter height that may occur if h is not an integer multiple of H. In either case we can re-use the first argument. For the action of S_α on each K' is a translation of its action on the first one. The action on K'' (after a suitable translation) produces a segment of the CZ type that stays within K'', since $H'' < H$.

Although the argument for $\alpha < 0$ is similar to the one just given, the identity

$$S_{-\alpha} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} S_\alpha \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{allows us to write} \quad |S_{-\alpha}K| = |DS_\alpha DK| = |S_\alpha DK| = |DK| = |K|$$

using $\alpha > 0$ and the Theorem in the case of matrices of the second type. This completes the proof of this case.

Completion of the proof of the Lemma If T is invertible then there exist a positive integer N and matrices E_k, $1 \le k \le N$, such that $E_1 \cdots E_N T = I$, where each of the E_k is a matrix of one of our three types (Elementary Row Operation matrices). If now J is an interval, then

$$|J| = |E_1 \cdots E_N T J| = |\det E_1||E_2 \cdots E_N T J| = \cdots = \left(\prod_{k=1}^{N} |\det E_k| \right) |TJ| = |\det T^{-1}||TJ|$$

since $E_1 \cdots E_N = T^{-1}$. This procedure takes care of the cases we did not consider in discussing matrices of the third type, including those when k and ℓ are not 1 and 2, and the case when they are 2 and 1.

Application: outer measure is unchanged by a rotation of the axes

The problem here arises because we sometimes insist on regarding a set as the "primitive" object, and describe it with more than one system of coordinates, for example a certain point in the city of Roseville, Minnesota.

If we are given an orthogonal matrix $\mathcal{O}$ with determinant one (also known as a rotation) we can assign to points $x = \sum_{i=1}^{n} x_i e_i$ new coordinates ξ_i by $\xi = \mathcal{O}^T x$. In the new coordinate system, which has basis vectors $\epsilon_i = \mathcal{O} e_i$, we construct intervals J that have edges parallel to the new coordinate axes. Then $\mathcal{O}J$ is an interval in our usual sense, and the volume (in the new coordinate system) of J is the same as the volume of $\mathcal{O}J$ in our initial coordinate system, because the lengths of the edges are preserved by the transformation $\mathcal{O}$. Our problem is to show that the measure of J, a rotated interval as seen in our original coordinate system, is the same when calculated in original coordinates as it is in the new coordinates, where the measure is simply the "volume," and that this holds as well for the outer measure of an arbitrary set E.

If $E \subseteq \mathbb{R}^n$ the coordinates of its elements in the new system comprise the set $\mathcal{O}^T E$. Then if (with respect to the new coordinates) we cover the set $\mathcal{O}^T E$ by families $\mathcal{J}$ of intervals J_j (in the new coordinate system), we will define the outer measure of E (as seen in the new coordinates) by

$$|E|_e^{\mathcal{O}} := \inf_{\mathcal{J} \text{ covers } \mathcal{O}^T E} \sum_{J_j \in \mathcal{J}} v_{\mathcal{O}}(J_j)$$

where $v_{\mathcal{O}}(J)$ is the product of the edge-lengths of J, as measured in the new system. Since $\mathcal{O}$ is orthogonal, $\mathcal{O}J$ is an interval in the original coordinates, so $v_{\mathcal{O}}(J) = |\mathcal{O}J|$, the measure being taken in the original coordinates.

Thus

$$|E|_e^{\mathcal{O}} = \inf_{\mathcal{J} \text{ covers } \mathcal{O}^T E} \sum_{J_j \in \mathcal{J}} |\mathcal{O}J_j| = \inf_{\mathcal{O}\mathcal{J} \text{ covers } E} \sum_{\mathcal{O}J_j \in \mathcal{O}\mathcal{J}} |\mathcal{O}J_j| = |E|_e$$

because the intervals $\mathcal{O}J_j$ cover E. Any collection of intervals that covers E in original coordinates can be converted a collection of intervals that covers $\mathcal{O}^T E$ in new coordinates, and vice versa. We multiply each $J \in \mathcal{J}$ by $\mathcal{O}$, getting $\mathcal{O}\mathcal{J}$, if we start on the left-hand side, and multiply each $I \in \mathcal{I}$ by $\mathcal{O}^T$, getting $\mathcal{O}^T \mathcal{I}$, if we start on the right-hand side.

The deferred proof

Lemma: *If $E \subseteq \mathbb{R}^n$ and L is an interval in $\mathbb{R}^m$ and $|L| > 0$, then $|E \times L|_e = |E|_e |L|_e$, the dimensions of the outer measures being $n+m$, n and m, respectively.*

Proof: First case: assume that E is bounded. Suppose intervals $I_k \subseteq \mathbb{R}^n$ cover E, with $\sum |I_k| < |E|_e + \epsilon$. Then the intervals $I_k \times L$ cover $E \times L$ so $|E \times L|_e \le (|E|_e + \epsilon)|L|_e$ and it follows that $|E \times L|_e \le |E|_e |L|_e$.

If E is not bounded, we let $E_k \uparrow E$, each E_k being bounded. Then $|E_k \times L|_e \to |E \times L|_e$ and $|E_k|_e \to |E|_e$, so $|E \times L|_e \le |E|_e |L|_e$. We digress for a Corollary:

Corollary: *If $E_1 \subseteq \mathbb{R}^n$ and $E_2 \subseteq \mathbb{R}^m$ and one of the sets is a null set then $E_1 \times E_2$ is a null set in $\mathbb{R}^{n+m}$.*

Proof: Suppose E_1 is a null set and E_2 is bounded. We cover E_1 as before,and cover E_2 by intervals J_ℓ whose volumes add to a finite sum. Then the intervals $I_k \times J_\ell$ cover $E_1 \times E_2$ so $|E_1 \times E_2|_e \le \sum_{k,\ell} |I_k||J_\ell| < C\epsilon$, and the Corollary holds for bounded E_2. The argument used before to pass to the unbounded case works, and the Corollary follows.

For the other inequality we revert to working on bounded E first. Then $E \times L$ is bounded, and (given $\epsilon > 0$) we can find an open $G \supseteq E \times L$ such that $|G|_e < |E \times L|_e + \epsilon$. We can write $z \in \mathbb{R}^{n+m}$ as $z = (x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. For each fixed $x \in E$, $\mathrm{dist}(z, G^c)$ is a continuous positive function of $y \in L$, a compact set. Its minimum value r_x is positive, and $B_{r_x}(x) \times L \subseteq G$. The union of the balls $B_{r_x}(x)$ as x runs thru E is an open set G' in $\mathbb{R}^n$ and $G' \times L \subseteq G$. We can write G' as a countable union of non-overlapping cubes Q_k. Then $G' \times L$ is a countable union of non-overlapping intervals $Q_k \times L$. This gives

$$|E \times L|_e + \epsilon > |G| \geq |G' \times L| = \sum_k |Q_k||L| = |G'||L| \geq |E|_e|L|; \quad \text{it follows that} \quad |E \times L|_e \geq |E|_e|L|.$$

If E is not bounded, the argument we used before to pass to the unbounded case works again. It follows that $|E \times L|_e = |E|_e|L|_e$ for all $E \subseteq \mathbb{R}^n$.