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 $TE \subseteq \bigcup_j J_j$. Then (by the Lemma)

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

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 there is a non-zero 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.

The easy case is a hyperplane H that is perpendicular to one of the coordinate axes. We may, by a translation, assume that the hyperplane is a subspace. Given $\epsilon > 0$, we can find an interval that is the Cartesian product of a cube in $\mathbb{R}^{n-1}$ having center zero and edge length k and an interval $[a, b]$ having center zero and length $k^{-n}2^{-k}k\epsilon$ that contains the intersection of H and the cube of edge k with center 0. The union of all these intervals, for $k \geq 1$, has measure at most ϵ. Thus H is a set of measure zero. If our hyperplane H (again assumed to be a subspace) is not one of the coordinate hyperplanes, we will construct a non-singular matrix M such that $H = M\{y \in \mathbb{R}^n : y \perp e_m\}$ for some m. This will show that H is a set of measure zero.

Let us suppose that our vector x_o such that $T^T x_o = 0$ has at least two non-zero coordinates. Otherwise, x_o is a non-zero multiple of one of the standard basis vectors e_j. We then suppose that $x_{om} \neq 0$ and $x_{ok} \neq 0$, where $k \neq m$. We put $a := x_o/|x_o|$ and we put $b := a - e_m$. Then $a_m \neq 0$ and $b_m = a_m - 1 \neq 0$ since $|a_m| < 1$ (because $a_m \neq 0$, $a_k \neq 0$ and $|a| = 1$).

We define $M = I - ab^T$. To show $\det M \neq 0$ it is enough to show (because the dimension is finite) that $Mx = 0 \Rightarrow x = 0$. If we had $Mx = 0$ and $x \neq 0$ then we would have

$$0 = Mx = x - \langle b, x\rangle\, a \quad \text{so} \quad x \text{ is a nonzero multiple of } a.$$

But then

$$Ma = a - \langle b, a\rangle\, a = a - \langle a - e_m, a\rangle\, a = a(1 - \langle a - e_m, a\rangle) = a_m \cdot a \neq 0.$$

Hence M is invertible. Next we suppose that $y = Mx$ and compute $\langle y, a\rangle = \langle y, x_o\rangle / |x_o|$:

$$\langle y, a\rangle = \langle Mx, a\rangle = \langle x, M^T a\rangle = \langle x, (I - ba^T)a\rangle = \langle x, a - b\rangle = \langle x, a - (a - e_m)\rangle = x_m.$$

Thus $H = M\{y \in \mathbb{R}^n : y \perp e_m\}$ so $|H| = 0$ since Lipschitz transformations map sets of measure zero into sets of measure zero.

We need to prove the Lemma now. We begin with three special types of linear transformations. The first and third types only make sense if $n > 1$.

First type: "switch rows" T is given by a permutation matrix $P = (\, [j = pi] \,) = \sum_i e_i e_{pi}^T$, 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, and 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 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 $k < 0$. 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 $k \neq \ell$. 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, 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 $P_{1k} T P_{2\ell}$ 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 thus 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 $n > 2$). 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 k are positive. Then $S_\alpha K$ is the parallelogram with vertices $Z := (0, 0)$, $A := (w, 0)$, $B := (w + \alpha h, \, h)$, and $C := (\alpha h, \, h)$. 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 $\alpha h < w$ 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 $Z A_o C$ (with its hypotenuse and horizontal leg missing), R is the rectangle $A_o B B_o C$ (with its upper and lower edges missing) and T_2 is the triangle $A B B_o$ (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

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

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

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

using $\alpha > 0$ and the Theorem in the case of matrices of the second type. This completes the proof of this case. Again, the Theorem is true for linear operators of this type.

Completion of the proof of the Lemma If T is invertible then there exist a positive integer N and matrices E_k, $1 \leq k \leq N$, 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 $E_1 \cdots E_N = T^{-1}$.

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

We begin by setting a "standard" coordinate system in $\mathbb{R}^n$ with basis vectors $\{e_1, \ldots, e_n\}$, where, for $1 \le j \le n$, e_j is the column vector with coordinates all zero except the j-th coordinate, which is one. To our notations used for intervals we can add this: express the *points* in an interval as $a + t$, where $0 \le t_j \le b_j - a_j$ for $1 \le j \le n$. We can also write $a + \sum_j t_j e_j$.

Now suppose we decide to change our coordinate system. We choose a new origin, Z, and basis vectors ϵ_j, $1 \le j \le n$. We choose them so that they are orthonormal (with respect to the standard system!) and so that the matrix $\mathcal{O}$ with columns $\epsilon_1 \cdots \epsilon_n$ (with respect to the standard system) has determinant one – hence the same orientation as the standard system (this is not necessary at all, just a choice most people would make). We now have our own version of $\mathbb{R}^n$: we begin with an n-tuple ξ of real numbers and associate it with the vector $Z + \sum_j \xi_j \epsilon_j$. We will usually call this vector ξ. From the standard point of view, $x := Z + \sum_j \xi_j \epsilon_j$ is related to the n-tuple ξ by $x - Z = \mathcal{O}\xi$ or by $\xi = \mathcal{O}^T(x - Z)$, since $\mathcal{O}^T$ is the inverse of $\mathcal{O}$.

From our new point of view, an interval J consists of (our new) vectors ξ belonging to a Cartesian product of intervals of real numbers. Its points have the form $\alpha + \sum_j t_j \epsilon_j$, $0 \le t \le b_j - a_j$, in the standard coordinate system, so the coordinate vector would be $\xi_t = \mathcal{O}^T(\alpha - Z + \sum_j t_j \epsilon_j) = \mathcal{O}^T(\alpha - Z) + \sum_j t_j e_j$ in the new system.

Thus the interval J in the new system, with volume $v_{\mathcal{O}}(J) = \prod_j (b_j - a_j)$, is the (standard) set of points $\mathcal{O}J + Z$, where J consists of the vectors ξ_t just written down. The ξ_t are not standard vectors, despite the fact that they are n-tuples of real numbers. However, we can regard the set of *coordinates* of J in the new system as a set J' in the standard system, and its (standard) volume will be the same as $v_{\mathcal{O}}(J)$, namely the product of the lengths of the factor intervals.

Because the matrix $\mathcal{O}$ is orthogonal, with determinant one, the set $\mathcal{O}J + Z$ (which is the interval called J in the new system) is a translated *rotated* interval. Now we regard J as the set J' of its coordinates, and so $\mathcal{O}J' + Z$ can be thought of as a transformation taking place in the standard system. By a translation and our Theorem, $|\mathcal{O}J' + Z| = |\mathcal{O}J'| = |\det \mathcal{O}||J'| = |J'| = v_{\mathcal{O}}(J)$.

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

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

where $v_{\mathcal{O}}(J)$ is the product of the edge-lengths of J, as measured in the new system.

We then have

$$\mathcal{O}^T(E - Z) = S \subseteq \bigcup_j J_j \quad \text{so} \quad E \subseteq \bigcup_j (\mathcal{O}J_j + Z)$$

and thus (omitting the primes we used above)

$$|E|_e \le \sum_j |\mathcal{O}J_j + Z| = \sum_j |J_j| = \sum_j v_{\mathcal{O}}(J_j), \quad \text{so} \quad |E|_e \le |E|_e^{\mathcal{O}}.$$

Finally, we note that we can regard the system $\{\epsilon_1, \ldots, \epsilon_n\}$ as the standard system, so that the reverse inequality holds.