Hurwitz's Theorem

Richard Koch

February 19, 2015

Theorem 1 (Hurwitz; 1898) *Suppose there is a bilinear product on R^n with the property that*

$$||v \circ w|| = ||v||||w||$$

Then $n = 1, 2, 4,$ or 8.

Proof; Step 1: Pick an orthonormal basis $e_1, e_2, \ldots, e_n$ for R^n, and consider the map $v \to e_i \circ v$ from R^n to R^n. This map is a linear transformation $A_i : R^n \to R^n$. Since $||e_i \circ v|| = ||e_i||||v|| = ||v||$, it is orthogonal, so $A_i^T A_i = I$.

If $r_1, r_2, \ldots, r_n$ are real numbers, we must have

$$< \sum r_i A_i(v), \sum r_j A_j(v) > = ||(\sum_i r_i e_i) \circ v||^2 = || \sum r_i e_i||^2 ||v||^2$$

for all v and all r_i, r_j, so

$$\sum r_i r_j < v, A_i^T A_j v > = \left(\sum r_i^2 \right) ||v||^2$$

But $A_i^T A_i = I$, so this formula becomes

$$\left(\sum_i r_i^2 \right) ||v||^2 + \sum_{i<j} \left(r_i r_j < v, (A_i^T A_j + A_j^T A_i) v > \right) \; = \left(\sum r_i^2 \right) ||v||^2$$

and consequently for all v, r_i and r_j,

$$\sum_{i<j} \left(r_i r_j < v, (A_i^T A_j + A_j^T A_i) v > \right) \; = 0$$

Fix i and j and let $r_i = r_j = 1$ and all other $r_k = 0$. We conclude that for all i and j,

$$< v, (A_i^T A_j + A_j^T A_i) v > = 0$$

Let $S = A_i^T A_j + A_j^T A_i$ and notice that $S^T = S$. Since $< v + w, S(v + w) >= 0$,

$$< v, S(v) > + < w, S(v) > + < v, S(w) > + < w, S(w) > = \ 0$$

and so

$$< w, S(v) > + < v, S(w) > = \ 0$$

So $< S^T(w), v > + < v, S(w) >=< S(w), v > + < v, S(w) >= 2 < v, S(w) >= 0$. This can only happen for all v and w if $S = 0$. We conclude that

$$A_i^T A_j + A_j^T A_i = 0$$

In the end, we have n linear transformations $A_1, A_2, \ldots, A_n$ satisfying $A_i^T A_i = I$ and $A_i^T A_j + A_j^T A_i = 0$. We will now ignore the context in which these matrices arose, and prove directly from these equations that $n = 1, 2, 4$, or 8.

Proof; Step 2: If our algebra had a unit e, we could have used it as a basis element, so that for example $A_n = I$.

We can achieve that directly by defining $B_i = A_i A_n^T$. Then $B_n = A_n A_n^T = I$. Moreover $B_i B_i^T = A_i A_n^T A_n A_i^T = A_i A_i^T$. Finally

$$B_i B_j^T + B_j B_i^T = A_i A_n^T A_n A_j^T + A_j A_n^T A_n A_i^T = A_i A_j^T + A_j A_i^T = 0$$

We now return to the original "A" notation. So assume $A_i^T A_i = I$ and $A_i^T A_j + A_j^T A_i = 0$ and $A_n = I$.

Then when $i < n$ we have $A_i A_n^T + A^n A_i^T = 0$, or $A_i + A_i^T = 0$. So $A_i^T = -A_i$. But $A_i^T A_i = I$, so $A_i^2 = -I$. In addition when $i, j < n$, $A_i A_j = -A_j A_i$.

When $n = 1$, there are no such A_i. But otherwise we have $\det(A_i)^2 = \det(-I) = (-1)^n$, which can only happen when n is even.

This concludes our study of the cases $n = 1, 2, 3$.

Proof; Step 3: From now on, assume $n \geq 4$ and n is even. We can ignore everything above except the existence of matrices A_i for $1 \leq i < n$ satisfying $A_i^2 = -I$ and $A_i A_j = -A_j A_i$. We even ignore A_{n-1} because our argument requires an even number of A_i.

At this point, the character of the proof changes. Form the set of all matrices $A_1^{\delta_1} A_2^{\delta_2} \ldots A_{n-2}^{\delta_{n-2}}$ where the δ_i are either zero or one. The number of such matrices is 2^{n-2}. We will prove that these matrices are linearly independent. The matrices live within the set of all $n \times n$ matrices, which has dimension n^2. So $2^{n-2} \leq n^2$.

For even n, this inequality is true for $n = 2, 4, 6, 8$, and no other n. So the bulk of the theorem follows from this independence statement.

If there is a dependent relation between the $A_1^{\delta_1} A_2^{\delta_2} \ldots A_{n-2}^{\delta_{n-2}}$, pick a relation

$$\sum \lambda_{\delta_1,\ldots,\delta_{n-2}} A_1^{\delta_1} A_2^{\delta_2} \ldots A_{n-2}^{\delta_{n-2}}$$

with as few non-zero coefficients as possible.

We are going to employ two tricks. The first is to multiply the terms of our dependence relation by one term $A_1^{\delta_1} A_2^{\delta_2} \ldots A_{n-2}^{\delta_{n-2}}$ on the right. The second is to multiply the terms of our dependence relation by some fixed A_i on both the left and the right.

What happens to the terms of our dependence relation when we do one of these tricks. Let us look at an example. Consider

$$(A_1 A_3 A_4)(A_2 A_3)$$

We can simplify using the rule $A_i A_j = -A_j A_i$ to get the terms in the correct order. We can simplify using the rule $A_i^2 = -I$ to get rid of duplicated terms. In the above example,

$$(A_1 A_3 A_4)(A_2 A_3) = -A_1 A_3 A_2 A_4 A_3 = A_1 A_2 A_3 A_4 A_3 = -A_1 A_2 A_3 A_3 A_4 = A_1 A_2 A_4$$

Ignore signs for a moment and concentrate on the terms. Let δ indicate an $n-2$ tuple

$$(\delta_1, \delta_2, \ldots, \delta_{n-2}) \in Z_2 \times Z_2 \times \ldots \times Z_2$$

If δ is such a vector, the term A^δ indicates the corresponding $A_1^{\delta_1} \ldots A_{n-2}^{\delta_{n-2}}$. If τ is another $n-2$-tuple, the generalization of the previous example and a little thought shows that $A^\delta A^\tau = \pm A^{\delta+\tau}$. Since $\delta \to \delta + \tau$ is a one-to-one and onto map from $Z_2 \times Z_2 \times \ldots \times Z_2$ to itself, multiplying on the right by some fixed A^δ produces a dependence relation with the same number of terms, and the same coefficients up to signs, although the actual terms which occur will change.

The same argument works if we multiply a relation on the left and right by the same A_i, except that this time we'll have the same terms, and coefficients are the same up to signs. If some signs change while others remain the same, then we can add the original to the new version and get a dependence relation with fewer terms, which is the desired contradiction. We get nothing if no signs change, or if all signs change.

Employ the first trick, where δ is an $(n-2)$-tuple which represents one of the terms of the minimal dependence relation. In the new relation, this term changes to $\delta = (0, 0, \ldots, 0)$ and so one of the terms of our dependence relation is I. From now on, we assume this.

Now multiply by A_i on both the left and the right. The term I will become $A_i I A_i = -I$, so the sign of its coefficient will change. Consequently the signs of all nonzero terms must change, or else we could find a dependence relation with fewer nonzero terms.

Consider terms with just one A_j. The result depends on whether $i = j$ or $i \neq j$.

$$A_i \to A_i A_i A_i = -A_i$$

and

$$A_j \to A_i A_j A_i = -A_i A_i A_j = A_j$$

Since $n - 2$ is even, we can find $j \neq i$. We conclude that no nonzero terms with one A can occur.

Consider an expression with two terms, $A_i A_j$. If we multiply this on both sides by A_i, we get

$$A_i A_i A_j A_i = -A_j A_i = A_i A_j$$

Since this term did not change sign, it cannot occur in our dependence relation. So no terms with two As can occur.

Consider a term with three terms $A_i A_j A_k$. Since $n - 2$ is even, there must be another index m unequal to i, j, k. Then

$$A_m(A_i A_j A_k)A_m = -A_i A_m A_j A_k A_m = A_i A_j A_m A_k A_m = -A_i A_j A_k A_m A_m = A_i A_j A_k$$

so this term cannot occur. So no terms with three As can occur.

It is now clear what happens in general. If there are an even number of A_is in a term, we can multiply on the left and right by one of the A_i in the term and get the same term without a sign change, which cannot happen. If there are an odd number of A_i in a term, we can find a A_m not in the term, multiply by it on the left and right, and not change the sign of the term.

In the end, only the I term can occur with nonzero coefficient, but the resulting sum does not equal zero.

Proof; Step 5 To finish the argument, we need only rule out $n = 6$.

Think of the A_j as acting on C^n rather than R^n; the matrices themselves remain unchanged. Since $A_1^2 = -I$, the eigenvalues of A_1 are $\pm i$. We can decompose $C^n = C^+ \oplus C^-$ where A_1 is i on the first space and $-i$ on the second. Indeed, this direct sum contains all of C^n because $v = \frac{v - iA_1(v)}{2} + \frac{v + iA_1(v)}{2}$.

Next we claim that if $j > 1$ then $A_j(C^+) \subset C^-$ and $A_j(C^-) \subset C^+$. Indeed, $A_1 A_j = -A_j A_1$. So $A_1 v = iv$ implies

$$A_1 A_j v = -A_j A_1 v = -A_j(iv) = (-i)A_j v$$

and $A_1 v = -iv$ implies similarly that $A_1 A_j v = iA_j v$.

Since $A_j^2 = -I$, A_j is one-to-one and onto: $C^+ \to C^-$ and $C^- \to C^+$. It follows that C^+ and C^- have the same dimension over C. So if $n = 6$, then C^+ and C^- both have complex dimension 3.

Consider the map $E = A_2A_3$ and $F = A_2A_4$ defined on C^n. Since each A_i for $i > 1$ interchanges C^+ and C^-, these maps leaves these spaces invariant. In particular, they both map C^+ to itself. The maps are both isomorphisms. Moreover, the maps anticommute, for

$$(A_2A_3)(A_2A_4) = -A_3A_2A_2A_4 = A_3A_4 = -A_4A_3 = A_4A_2A_2A_3 = -(A_2A_4)(A_2A_3)$$

Compute the determinants of E and F as maps from C^+ to itself. We have $\det(EF) = \det(FE)$ and

$$\det(EF) = \det(-FE) = \det(-I)\det(FE)$$

We conclude that $\det)(-I) = 1$, but since C^+ has dimension 3, this determinant is -1. This contradiction rules out $n = 6$. QED.

Remark: We will not prove it here, but an easy consequence of this result classifies extended versions of the cross product.

Definition 1 *A* cross product *on R^n is a product $v, w \to v \times w$ such that*

- *The product is bilinear*
- *The element $v \times w$ is perpendicular to v and w*
- *The element $v \times v$ is zero*
- *If v and w are perpendicular and have length one, then $||v \times w|| = 1$*

Remark: These axioms imply that

$$||v \times w||^2 = ||v||^2||w||^2 - (v \cdot w)^2$$

Theorem 2 *On R^1, the only possible cross product is $v \times w = 0$ for all v, w. This definition fails in all higher dimensions.*

Theorem 3 *There is a non-trivial cross product on R^n if and only if $n = 3$ or 7.*

Sketch of the proof: We mimic the definition of the quaternions. Think of R^{n+1} as all $< r, v >$ where $r \in R$ and $v \in R^n$. Assuming a cross product exists on R^n, define a product on R^{n+1} by

$$< r, v >< s, w >=< rs - v \cdot w, rw + sv + v \times w >$$

Check that

$$|| < r, v >< s, w > ||^2 = || < r, v > ||^2|| < s, w > ||^2$$

and apply the Hurwitz theorem.

Remark A final note. The above proof constructed a series of matrices $A_1, \ldots, A_n$ satisfying $A_i^2 = -I$ and $A_i A_j = -A_j A_i$. More generally, suppose the A_i are abstract symbols satisfying these rules. Then the A_i generate an associative algebra containing 1 and all products of the A_i, called the *Clifford Algebra*. An additive basis for this algebra is 1 and all $A_{i_1} \cdot A_{i_2} \cdot \ldots \cdot A_{i_k}$ with $1 \leq i_1 < \ldots < i_k \leq n$. So the dimension is 2^n.

Starting at $n = 0$, the first few Clifford algebras are R, C, H, $H \oplus H$. All Clifford algebras are semisimple, and thus sums of full matrix algebras over R, C, or H. The algebras satisfy the following remarkable periodicity result:

Theorem 4 *The Clifford algebra associated with* $n+8$ *is isomorphic to the set of all* 16×16 *matrices with entries in the Clifford algebra associated with* n

This theorem is closely related to the Bott periodicity theorem.

The Clifford algebras are associative. Nevertheless, Hurwitz' proof shows that they are related to the octonions via the theory of normed division algebras.