Boletín de la Sociedad Matemática Mexicana Vol. 29, No. 1, 1984

ON THE HURWITZ PROBLEM OVER AN ARBITRARY FIELD*

BY DANIEL B. SHAPIRO

In this article we deal with product formulas of the type

$$(x_1^2 + x_2^2 + \cdots + x_r^2)(y_1^2 + y_2^2 + \cdots + y_s^2) = z_1^2 + z_2^2 + \cdots + z_n^2,$$

where $X = (x_1, \cdots, x_r)$ and $Y = (y_1, \cdots, y_s)$ are systems of indeterminates, and each $z_i = z_i(X, Y)$ is a bilinear form in X, Y with coefficients in a field F. A triple (r, s, n) is said to be *admissible over F* if such a formula exists. We assume throughout that the characteristic of F is not 2.

Which triples (r, s, n) are admissible over F? This question seems to be very difficult, with only a few special cases settled. The purpose of this paper is to present a new proof of Adem's theorem on the admissibility of $(r, n-1, n)$. Before stating this theorem let us outline some of the earlier work done in this area.

The question of admissibility was first formulated by A. Hurwitz in 1898 in the case F is the field $\mathbb{C}$ of complex numbers. In [H1] he proved that (n, n, n) is admissible over $\mathbb{C}$ if and only if $n = 1, 2, 4$ or 8. About 20 years later the admissibility of (r, n, n) was characterized in terms of the Hurwitz-Radon function $\rho(n)$, defined as follows.

Definition. Let $n = 2^m u$ where u is odd. If $m = 4a + b$ where $0 \le b \le 3$ then $\rho(n) = 8a + 2^b$. Equivalently,

$$\rho(n) = \begin{cases} 2m+1 & \text{if} \quad m \equiv 0 \\ 2m & \text{if} \quad m \equiv 1 \\ 2m & \text{if} \quad m \equiv 2 \\ 2m+2 & \text{if} \quad m \equiv 3 \end{cases} \qquad (\text{mod } 4).$$

THEOREM (1). *(r, n, n) is admissible over F if and only if $r \le \rho(n)$.*

This result was first proved when $F = \mathbb{R}$ by Radon [R] and when $F = \mathbb{C}$ by Hurwitz [H2] (published posthumously). The "if" part requires a construction of a formula of size $(\rho(n), n, n)$. Such formulas valid over any field F have been constructed by a number of authors. The "only if" part for general F can be proved by a straightforward modification of Hurwitz's argument. Further details and references can be found in the survey article [S].

ADEM'S THEOREM (2). [A] *Suppose $(r, n-1, n)$ is admissible over F.*
(i) *If n is even then (r, n, n) is admissible over F.*
(ii) *If n is odd then $(r, n-1, n-1)$ is admissible over F.*

Combining this result with Theorem 1, we know exactly when $(r, n-1, n)$

* This research was supported in part by the NSF.

is admissible. If $n = 2k$ or $2k + 1$ for an integer k, then $(r, n - 1, n)$ is admissible over F if and only if $r \leq \rho(2k)$.

Before beginning the proof of Adem's theorem, we remark that Yuzvinsky has a result on the much harder case $(r, n - 2, n)$.

THEOREM (3). (Yuzvinsky [Y]). *If* $n \equiv 3 \pmod 4$ *then* $(4, n - 2, n)$ *is not admissible over* F.

The three theorems above suffice to characterize the admissibility of (r, s, n) over F whenever $r \leq 4$. The smallest case not settled so far is (5, 9, 12). In the case F has characteristic 0, some further conditions are known. The proofs involve some algebraic topology, and are outlined in [S]. For example, (5, 9, 12) is not admissible over F if F has characteristic 0.

Our proof of Adem's theorem uses an argument involving matrices over rings and the generalization of the cross product. We begin with three lemmas.

LEMMA (4). *Let* $X = (x_1, \cdots, x_r)$ *be a system of indeterminates over* F. *Then* (r, s, n) *is admissible over* F *if and only if there exists an* $n \times s$ *matrix* A *whose entries are linear forms in* X *with coefficients in* F, *satisfying* $A^tA = (\Sigma x_i^2)I_s$.

The proof is well known, going back to Hurwitz. Details appear in [S], for example. One can further express this A as $A = x_1A_1 + \cdots + x_rA_r$, and restate the conditions in terms of the $n \times s$ matrices A_i over F. For our purposes however, it is convenient to work directly with A as a matrix over the polynomial ring $R = F[X]$.

For the next two lemmas let R be any commutative ring with 1, and consider the free R-module $V = R^n$ as an inner product space by means of the usual dot product. Let $\{e_1, e_2, \cdots, e_n\}$ be the canonical basis of V. Then $e_i \cdot e_j = \delta_{ij}$, (Kronecher delta).

LEMMA (5). *Let* $V^{n-1} = V \times V \times \cdots \times V$ *be the direct product. There is a unique map* $p: V^{n-1} \to V$ *satisfying*

(1) p *is* $(n - 1)$*-linear and alternating*;

(2) $v_j \cdot p(v_1, v_2, \cdots, v_{n-1}) = 0$, *for any* $v_1, \cdots, v_{n-1} \in V$;

(3) $p(e_1, e_2, \cdots, e_{n-1}) = e_n$;

(4) $p(v_1, \cdots, v_{n-1}) \cdot p(v_1, \cdots, v_{n-1}) = \det((v_i \cdot v_j))$, *for any* $v_1, \cdots, v_{n-1} \in V$.

Proof. This is the standard generalization of the usual cross product $p: R^3 \times R^3 \to R^3$ in the case $n = 3$. We sketch how p arises in terms of exterior algebra. The dot product on $V = R^n$ induces a dot product on $\Lambda^p V$ by requiring: $(u_1 \wedge \cdots \wedge u_p) \cdot (v_1 \wedge \cdots \wedge v_p) = \det((u_i \cdot v_j))$, the determinant of the Gram matrix. Using the basis vector $e = e_1 \wedge e_2 \wedge \cdots \wedge e_n$ of $\Lambda^n V$, we define the (Hodge) star operator $*: \Lambda^{n-p} V \to \Lambda^p V$ by requiring $\lambda \wedge \mu = ((*\lambda) \cdot \mu)e$, whenever $\lambda \in \Lambda^{n-p} V$ and $\mu \in \Lambda^p V$. Our map p is then defined as $p(v_1, v_2, \cdots, v_{n-1}) = *(v_1 \wedge v_2 \wedge \cdots \wedge v_{n-1})$. The properties (1), (2), (3) follow easily and (4) is deduced from the formula: $(*\alpha) \cdot (*\beta) = \alpha \cdot \beta$ whenever $\alpha, \beta \in \Lambda^{n-p} V$.

A concise exposition of these ideas can be found in [F] Chap. II. For more information see [B]. □

For concreteness let us describe p in terms of determinants. Given the column vectors $v_1, v_2, \cdots, v_{n-1} \in V = R^n$, let A be the $n \times (n-1)$ matrix having these columns. Let A_j be the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the j-th row, and set $d_j = (-1)^{n-j}\det(A_j)$. Then

$$p(v_1, v_2, \cdots, v_{n-1}) = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} \in V.$$

This formula can be derived quickly from the exterior algebra description given above. Alternatively one can take this formula as the definition of p and deduce the properties listed in Lemma 6. To do this, we first use expansion by minors to see that $p(v_1, \cdots, v_{n-1})\cdot w = \det(A \mid w)$, where $(A \mid w)$ denotes the $n \times n$ matrix obtained by adjoining the column w to A. Properties (1), (2), (3) are now easy, but (4) requires more work. That formula follows from the "Lagrange identity" (sometimes called the "Cauchy-Binet formula") given in [B] §8, no. 2 or [F] Chap. II, exer. 6. One can also deduce (4) by considering the determinant of $(A \mid u)^t(A \mid u)$ when $u = p(v_1, \cdots, v_{n-1})$.

LEMMA (6). *Let $v_1, \cdots, v_{n-1} \in V = R^n$ and let these be the columns of the $n \times (n-1)$ matrix A. Let $p(v_1, v_2, \cdots, v_{n-1}) = (d_1, d_2, \cdots, d_n)^t$ as above. If $A^tA = aI_{n-1}$ for some $a \in R$, then $d_j^2 \equiv 0(\bmod a^{n-2})$, for each $j = 1, 2, \cdots, n-1$.*

Proof. We need only consider the case $j = 1$. Express A in blocks as: $A = \left(\frac{w^t}{A_1}\right)$ where $w \in R^{n-1}$ and A_1 has size $(n-1) \times (n-1)$. We know $d_1 = \pm \det(A_1)$. Also $aI_{n-1} = A^tA = ww^t + A_1{}^tA_1$, so that we have $d_1{}^2 = \det(A_1{}^tA_1) = \det(aI_{n-1} - ww^t)$. Now the square matrix ww^t has characteristic polynomial $\chi(t) = \det(tI_{n-1} - ww^t) = t^{n-2}(t-c)$ where $c = w^tw \in R$. This equality can be seen over fields using eigenvalues, and then generalized to rings in the standard way. A more direct proof is provided by the clever argument of Schmid [Sch]. Therefore $d_1{}^2 = \chi(a) = a^{n-2}(a-c)$ in the ring R. □

We are now in a position to prove Adem's Theorem. Suppose $(r, n-1, n)$ is admissible over F. By Lemma 4 we have an $n \times (n-1)$ matrix A, whose entries in $F[X] = F[x_1, \cdots, x_r]$ are linear forms, satisfying $A^tA = aI_{n-1}$, where $a = x_1{}^2 + \cdots + x_r{}^2$. If $r = 1$ the conclusion is trivial, so let us assume $r > 1$. Then the element $a \in F[X]$ is squarefree. (In fact if $r \geq 3$ then a is irreducible.)

Let $v_1, v_2, \cdots, v_{n-1} \in F[X]^n$ be the columns of A and define $u = p(v_1, v_2, \cdots, v_{n-1}) = (d_1, d_2, \cdots, d_n)^t$ as in the Lemmas. Then $u\cdot u = \det((v_i\cdot v_j)) = \det(A^tA) = \det(aI_{n-1}) = a^{n-1}$. The description $d_j = \pm\det(A_j)$ implies that $d_j \in$

$F[X]$ is a homogeneous polynomial of degree $n-1$. Also Lemma 6 says $d_j^2 \equiv 0 (\mathrm{mod}\ a^{n-2})$.

Case (i): $n = 2k$ is even. Since $F[X]$ is a unique factorization domain and $a \in F[X]$ is squarefree, we conclude $d_j \equiv 0(\mathrm{mod}\ a^{k-1})$. Then the vector $\hat{u} = a^{-k+1}u$ still lies in $F[X]^n$. Computing degrees we see that the entries of $\hat{u}$ are linear forms in $F[X]$. Also $v_j \cdot \hat{u} = 0$, so that $A^t\hat{u} = 0$, and $\hat{u} \cdot \hat{u} = a^{-2k+2}(u \cdot u) = a$. Then the $n \times n$ matrix $(A \mid \hat{u})$ satisfies the conditions of Lemma 4, showing that (r, n, n) is admissible over F.

Case (ii): $n = 2k + 1$ is odd. Then $d_j \equiv 0(\mathrm{mod}\ a^k)$ so that the vector $\hat{u} = a^{-k}u$ still lies in $F[X]^n$. Computing degrees we see that $\hat{u}$ has entries of degree 0, that is: $\hat{u} \in F^n$. Also $A^t\hat{u} = 0$ and $\hat{u} \cdot \hat{u} = 1$. We can choose a new orthonormal basis of F^n having $\hat{u}$ as the last element. Rewriting everything relative to the new basis we have the same situation as before, but now with $\hat{u} = e_n = (0, \cdots, 0, 1)^t$. Since $A^t\hat{u} = 0$, the bottom row of A is zero: $A = \left(\frac{B}{0}\right)$ where B has size $(n-1) \times (n-1)$. Then B satisfies the conditions of Lemma 4, showing that $(r, n-1, n-1)$ is admissible over F. □

THE OHIO STATE UNIVERSITY
COLUMBUS, OHIO, U.S.A.

REFERENCES

[A] J. ADEM, *On the Hurwitz problem over an arbitrary field* I, II, Bol. Soc. Mat. Mexicana **25** (1980) 29–51 and **26** (1981) 29–41.

[B] N. BOURBAKI, Eléments de Mathématiques, Algèbre, Chap. 3, Algèbre Multilineaire, Hermann, Paris, 1958.

[F] H. FLANDERS, Differential Forms, Academic Press, New York, 1963.

[H1] A. HURWITZ, *Über die Komposition der quadratischen Formen von beliebig vielen Variabeln*, Nachr. Akad. Wiss., Göttingen, Math. Phys. Kl. (1898) 309–316. Reprinted in Math. Werke II, 565–571.

[H2] ———, *Über die Komposition der quadratischen Formen*, Math. Ann. **88** (1923) 1–25. Reprinted in Math. Werke II, 641–666.

[R] J. RADON, *Lineare Scharen orthogonaler Matrizen*, Abh. Math. Sem. Univ. Hamburg 1 (1922) 1–14.

[S] D. B. SHAPIRO, *Products of Sums of Squares*, Expo. Math. **2** (1984) 235–261.

[Sch] J. SCHMID, *A remark on characteristic polynomials*, Amer. Math. Monthly **77** (1970) 998–999. Reprinted in: Selected Papers on Algebra, Math. Assoc. of Amer., 1977, p. 332.

[Y] S. YUZVINSKY, *On the Hopf condition over an arbitrary field*, Bol. Soc. Mat. Mexicana **28** (1983) 1–7.