FIELDS DEFINED BY POLYNOMIALS

LOWELL A. HINRICHS, IVAN NIVEN AND C. L. VANDEN EYNDEN

1. Introduction. First we consider the following question, where F is any field. For what pairs P and Q of polynomials in two variables with coefficients in F do the definitions

$$(\mathrm{I}) \qquad a \oplus b = P(a, b), \qquad a \odot b = Q(a, b),$$

for all a and b in F yield a field $(F, \oplus, \odot)$? It turns out that the answer is different for infinite fields than for finite fields, as shown in §§ 2 and 3.

Next let R be the field of real numbers. For what quadruples P_1, P_2, Q_1, Q_2 of real polynomials in four variables is $(R \times R, \oplus, \odot)$ a field, when we set

$$(\mathrm{II}) \qquad \begin{aligned} (a, b) \oplus (c, d) &= (P_1(a, b, c, d), P_2(a, b, c, d)), \\ (a, b) \odot (c, d) &= (Q_1(a, b, c, d), Q_2(a, b, c, d)), \end{aligned}$$

where (x, y) denotes an ordered pair of real numbers? This question is partially answered in §§ 4 and 5, and in § 6 it is shown that the polynomials may be of arbitrarily high degree. In § 7 it is proved that if definitions (II) do give a field, it must be isomorphic to the field of complex numbers.

2. The one-dimensional case.

THEOREM 1. *Let F be an infinite field. The system $(F, \oplus, \odot)$ in* (I) *is a field if and only if*

$$(1) \qquad \begin{aligned} P(a, b) &= a \oplus b = a + b + \gamma \\ Q(a, b) &= a \odot b = \gamma\sigma(a + b) + \sigma ab + \gamma^2\sigma - \gamma, \end{aligned}$$

where $\gamma \in F$, $\sigma \in F$ and $\sigma \neq 0$. When these conditions are satisfied the field $(F, \oplus, \odot)$ is isomorphic to F, thus $(F, \oplus, \odot) \cong (F, +, \cdot)$.

Proof. We first assume that $(F, \oplus, \odot)$ is a field and show that the polynomials P and Q have the prescribed form. By associativity we have $P(P(a, b), c) = P(a, P(b, c))$ identically in a, b, c. Now if P is of degree n in a, the degrees of the left and right sides of this identity in a are n^2 and n respectively. Since F is infinite it follows

Received November 12, 1962, and in revised form June 14, 1963. The first author listed was supported by NSF contract G 19859, the other two authors by NSF contract G 19016.

that $n^2 = n$ and hence $n = 1$. We conclude that $P(a, b)$ is linear in a and b, and the same holds for $Q(a, b)$.

Using this linearity and also the commutative properties, we can write

$$a \oplus b = \alpha(a + b) + \beta ab + \lambda ,$$
$$a \odot b = \rho(a + b) + \sigma ab + \tau .$$

Now $\beta = 0$, for if $\beta \neq 0$ we would have

$$(-\alpha/\beta) \oplus b = -\alpha^2/\beta + \lambda ,$$

and the right member is independent of b.

Suppose first that the additive and multiplicative identities are 0 and 1. Then the equations

$$a \oplus 0 = a , \qquad a \odot 0 = 0 , \qquad a \odot 1 = a$$

show that $\alpha = 1$ and $\lambda = 0$, that $\rho = 0$ and $\tau = 0$, and that $\sigma = 1$. Thus we have

$$a \oplus b = a + b , \qquad a \odot b = a \cdot b ,$$

so that $\oplus$ and $\odot$ are simply the ordinary operations.

But now suppose that z and u denote the additive and multiplicative identities of the field $(F, \oplus, \odot)$. Then the mapping

$$a \to f(a) = (u - z)a + z$$

gives $f(0) = z$ and $f(1) = u$. Since f is a one-to-one mapping of F onto F, the operations $\oplus'$ and $\odot'$ defined by

$$x \oplus' y = f^{-1}(f(x) \oplus f(y)) , \tag{2}$$
$$x \odot' y = f^{-1}(f(x) \odot f(y)) ,$$

yield a field $(F, \oplus', \odot')$ which is isomorphic under f to $(F, \oplus, \odot)$. But it is easily checked that $\oplus'$ and $\odot'$ are again polynomial operations in the sense of (I). Furthermore note that

$$x \oplus' 0 = x \odot' 1 = x , \qquad x \odot' 0 = 0 ,$$

and so by the argument of the preceding paragraph we conclude that $\oplus'$ and $\odot'$ are just $+$ and $\cdot$. Now if we substitute $x = f^{-1}(a)$ and $y = f^{-1}(b)$ into equations (2) and apply f to both sides we get

$$a \oplus b = f(f^{-1}(a) + f^{-1}(b)) = a + b - z , \tag{3}$$
$$a \odot b = f(f^{-1}(a) \cdot f^{-1}(b)) = (a - z)(b - z)(u - z)^{-1} + z .$$

Writing γ for $-z$ and σ for $(u - z)^{-1}$ we see that equations (3) are the same as (1).

Conversely, given any elements γ and $\sigma \neq 0$ of F we see that the operations defined by equations (1) give a field isomorphic to $(F, +, \cdot)$, because the mapping f^{-1} is an isomorphism:

$$f^{-1}(a \oplus b) = f^{-1}(a) + f^{-1}(b) ,$$
$$f^{-1}(a \odot b) = f^{-1}(a) \cdot f^{-1}(b) .$$

3. **Finite fields.** The restriction of Theorem 1 to infinite fields was necessary because in the proof use was made of the fact that polynomials agreeing on infinite sets must be identical. Now for a finite field F of order $q = p^n$ we see that a system $(F, \oplus, \odot)$ in (I) is a field with

$$P(a, b) = a \oplus b = a^q + b^q , \qquad Q(a, b) = a \odot b = a^q b^q .$$

But these are artificial definitions since $a^q = a$ identically in a in the finite field. However, Theorem 1 fails in a genuine sense for all cases except $q = 2, 3, 4$, as can be seen as follows.

Let g be any permutation on F leaving 0 and 1 invariant. Now g is a polynomial function because we can construct a polynomial to agree with g over the q elements of the field. Similarly the operations $\oplus$ and $\odot$ defined by

$$a \oplus b = g^{-1}(g(a) + g(b)) ,$$
$$a \odot b = g^{-1}(g(a) \cdot g(b)) , \tag{4}$$

are polynomial functions. If Theorem 1 were true for the finite field F then equations (4) would be of the form (1) for some γ and σ. But from (4) we see that $a \oplus 0 = a$ and $a \odot 1 = a$, so that 0 and 1 are the additive and multiplicative identities of $(F, \oplus, \odot)$. Hence in (1) we see that $\gamma = 0$ and $\sigma = 1$. Thus $\oplus$ and $\odot$ would be the ordinary operations and (4) would be

$$a + b = g^{-1}(g(a) + g(b)) ,$$
$$a \cdot b = g^{-1}(g(a) \cdot g(b)) .$$

It follows that g is an automorphism of $(F, \oplus, \odot)$. But there exist exactly n automorphisms of a field with p^n elements [4, § 38], Since there are $(p^n - 2)!$ permutations g of F leaving 0 and 1 invariant, and since $(p^n - 2)! > n$ if $p^n \geqq 5$, it follows that Theorem 1 fails for finite fields of order $q = p^n \geqq 5$.

On the other hand suppose that F is a finite field of order $q = p^n = 2, 3$, or 4. Suppose further that there are polynomials P and Q for which the operations $a \oplus b = P(a, b)$ and $a \odot b = Q(a, b)$ yield a field $(F, \oplus, \odot)$. Using the mapping $f(a) = (u - z)a + z$, we apply f^{-1} as in equations (2). Thus we move from $(F, \oplus, \odot)$ to $(F, \oplus', \odot')$

having 0 and 1 as additive and multiplicative identities. Now simple examination of the addition and multiplication tables for finite fields with 2, 3 or 4 elements shows that the operations $\oplus'$ and $\odot'$ must be the ordinary operations of addition and multiplication. Thus we can get equations (3) and the rest of the proof follows as in Theorem 1. We have proved the following result.

THEOREM 2. *Theorem* 1 *holds for only those finite fields with* 2, 3 *or* 4 *elements.*

4. The complex case: a simplification. The definition (II) allows considerably more latitude for the operations $\oplus$ and $\odot$ than exists in the one-dimensional case, and the problem appears to be correspondingly more difficult. To simplify things we show first that there is no great loss in generality in presuming that the additive ard multiplicative identities of the field $(R \times R, \oplus, \odot)$ are (0, 0) and (1, 0). For let the zero and unity of the field be denoted by (p, q) and (r, s). We define

$$(5) \qquad [a, b] = (ar - ap - bs + bq + p, as - aq + br - bp + q) ,$$

and note that

$$[0, 0] = (p, q), [1, 0] = (r, s) .$$

The right member of (5) is simply

$$(a, b)(r - p, s - q) + (p, q) ,$$

where the multiplication and addition are as in the field of complex numbers. Since $(p, q) \neq (r, s)$ we see that $(r - p, s - q) \neq (0, 0)$ and so (5) is a one-to-one mapping of $R \times R$ onto $R \times R$. If we extend the multiplications $\oplus$ and $\odot$ to the pairs $[a, b]$ by the use of (5) we see that

$$[0, 0] \oplus [a, b] = [1, 0] \odot [a, b] = [a, b] .$$

Furthermore, $[a, b] = (x, y)$ implies not only that x and y are polynomials in a and b by (5), but also that a and b are polynomials in x and y. Hence any system of pairs (a, b) with $\oplus$ and $\odot$ defined by (II) can be transformed into an isomorphic system of pairs $[a, b]$ with $\oplus$ and $\odot$ defined by (5) and (II). Thus all fields of the required sort can be generated in a simple way as in § 2 from those having (0, 0) and (1, 0) as zero and unit.

5. The complex case with linearity.

THEOREM 3. *Let the operations* $\oplus$ *and* $\odot$ *be defined as in* (II), *and assume that each of* P_1, P_2, Q_1, Q_2 *is linear in each argument*

separately. Then $(R \times R, \oplus, \odot)$ *is a field with* $(0, 0)$ *and* $(1, 0)$ *as zero and unity if and only if*

$$(a, b) \oplus (c, d) = (a + c, b + d) \quad \textit{and}$$
$$(a, b) \odot (c, d) = (ac + \gamma bd, ad + bc + \delta bd)$$

for some $\gamma \in R$ *and* $\delta \in R$ *with* $\delta^2 + 4\gamma < 0$. *When these conditions are satisfied,* $(R \times R, \oplus, \odot)$ *is isomorphic to the field of complex numbers, that is,* $(R \times R, \oplus, \odot) \cong (C, +, \cdot)$.

Proof. First we assume that $(R \times R, \oplus, \odot)$ is a field. By the commutative property $P_1(a, b, c, d)$ is symmetric in a and c and also in b and d; likewise for P_2, Q_1 and Q_2. Thus we can write

$$\begin{aligned} P_1(a, b, c, d) = {} & \alpha_0 + \alpha_1(a + c) + \alpha_2(b + d) + \alpha_{12}(ab + cd) \\ & + \alpha_{13}ac + \alpha_{24}bd + \alpha_{14}(ad + bc) + \alpha_{123}(abc + acd) \\ & + \alpha_{124}(abd + bcd) + \alpha_{1234}abcd \,. \end{aligned}$$

We represent P_2, Q_1 and Q_2 by similar expressions with the α's replaced by β's, γ's and δ's respectively. From the relation $(a, b) \oplus (0, 0) = (a, b)$ we deduce

$$P_1(a, b, 0, 0) = a \,, \qquad P_2(a, b, 0, 0) = b \,,$$

from which it follows that

$$\alpha_1 = \beta_2 = 1 \quad \text{and} \quad \alpha_0 = \beta_0 = \alpha_2 = \beta_1 = \alpha_{12} = \beta_{12} = 0 \,.$$

Now define (h, k) by the relation $(1, 0) \oplus (1, 0) = (h, k)$. Then the distributive property implies that

$$(a, b) \odot (h, k) = (a, b) \oplus (a, b)$$

and so we obtain

$$\begin{aligned} P_1(a, b, a, b) &= Q_1(a, b, h, k) \\ &= 2a + \alpha_{13}a^2 + \alpha_{24}b^2 + 2\alpha_{14}ab + 2\alpha_{123}a^2b \\ &\quad + 2\alpha_{124}ab^2 + \alpha_{1234}a^2b^2 \,. \end{aligned}$$

But $Q_1(a, b, h, k)$ is linear in a and b, and hence

$$\alpha_{13} = \alpha_{24} = \alpha_{14} = \alpha_{123} = \alpha_{124} = \alpha_{1234} = 0 \,.$$

The relation $P_2(a, b, a, b) = Q_2(a, b, h, k)$ yields an analogous result for the β's, and so we get

$$(a, b) \oplus (c, d) = (a + c, b + d) \,.$$

Next, from the relation $(a, b) \odot (0, 0) = (0, 0)$ we see that

$$Q_1(a, b, 0, 0) = Q_2(a, b, 0, 0) = 0 ,$$

and so

$$\gamma_0 = \gamma_1 = \gamma_2 = \gamma_{12} = \delta_0 = \delta_1 = \delta_2 = \delta_{12} = 0 .$$

From $Q_1(a, b, 1, 0) = a$ and $Q_2(a, b, 1, 0) = b$ we obtain

$$\gamma_{13} = \delta_{14} = 1 , \qquad \delta_{13} = \gamma_{14} = \gamma_{123} = \delta_{123} = 0 .$$

Thus we have

$$\begin{aligned} Q_1(a, b, c, d) &= ac + \gamma_{24}bd + \gamma_{124}(bcd + abd) + \gamma_{1234}abcd , \\ Q_2(a, b, c, d) &= ad + bc + \delta_{24}bd + \delta_{124}(bcd + abd) + \delta_{1234}abcd . \end{aligned}$$

Also the equations

$$\begin{aligned} (a, b) \odot (1, 1) &= (a, b) \odot (1, 0) \oplus (a, b) \odot (0, 1) \\ &= (a, b) \oplus (a, b) \odot (0, 1) \end{aligned}$$

imply that

$$Q_1(a, b, 1, 1) = a + Q_1(a, b, 0, 1), Q_2(a, b, 1, 1) = b + Q_2(a, b, 0, 1) .$$

This yields

$$\gamma_{124} = \gamma_{1234} = \delta_{124} = \delta_{1234} = 0 ,$$

and so we have, removing subscripts,

$$(a, b) \odot (c, d) = (ac + \gamma bd, ad + bc + \delta bd) .$$

Finally, if $(a, b) \neq (0, 0)$, there must exist real numbers x and y such that $(a, b) \odot (x, y) = (1, 0)$. This gives a pair of linear equations with determinant $a^2 + \delta ab - \gamma b^2$. This must not vanish except for $a = 0$ and $b = 0$, and so we conclude that

$$\delta^2 + 4\gamma < 0 .$$

Conversely, to prove that the operations $\oplus$ and $\odot$ in the statement of the theorem do give a field isomorphic to the field of complex numbers, define α and β by

$$\alpha = \frac{\delta}{2} , \qquad \beta = \frac{\sqrt{-4\gamma - \delta^2}}{2} .$$

Since $\beta \neq 0$ the mapping

$$\phi: (a, b) \to (a + \alpha b, \beta b)$$

is one-to-one from C onto itself. As in Theorem 1 we point out that by a not difficult calculation

$$(a, b) \oplus (c, d) = \phi^{-1}(\phi(a, b) + \phi(c, d))$$

and

$$(a, b) \odot (c, d) = \phi^{-1}(\phi(a, b) \cdot \phi(c, d)) .$$

Thus the mapping ϕ is an isomorphism from $(R \times R, \oplus, \odot)$ to $(C, +, \cdot)$.

As a variation on Theorem 3 we prove the following; see [2, p. 251] for a related result.

THEOREM 4. *In Theorem* 3 *replace the hypothesis that* P_1, P_2, Q_1 *and* Q_2 *are linear by the assumption*

$$(a, b) \odot (c, 0) = (ac, bc) \tag{6}$$

for all a, b, c *in* R. *Then the conclusion of Theorem* 3 *holds.*

Proof. If first we assume the definitions of $\oplus$ and $\odot$ as in the equations of Theorem 3, then we have a field, and we note that (6) follows. Conversely, suppose that $(R \times R, \oplus, \odot)$ is a field with the usual zero and unity and such that (6) holds. Then we note that

$$\begin{aligned}(aP_1(x, y, z, w), aP_2(x, y, z, w)) &= (a, 0) \odot (P_1(x, y, z, w), P_2(x, y, z, w)) \\ &= (a, 0) \odot ((x, y) \oplus (z, w)) \\ &= (ax, ay) \oplus (az, aw) \\ &= (P_1(ax, ay, az, aw), P_2(ax, ay, az, aw)) .\end{aligned}$$

Thus P_1 and P_2 are homogeneous and linear.

Turning to the operation $\odot$ we note that

$$\begin{aligned}(aQ_1(x, y, z, w), aQ_2(x, y, z, w)) &= (a, 0) \odot ((x, y) \odot (z, w)) \\ &= (ax, ay) \odot (z, w) \\ &= (Q_1(ax, ay, z, w), Q_2(ax, ay, z, w)) .\end{aligned}$$

Applying the commutative property we get

$$\begin{aligned}(a^2Q_1(x, y, z, w), a^2Q_2(x, y, z, w)) &= (Q_1(ax, ay, az, aw), Q_2(ax, ay, az, aw))\end{aligned}$$

and hence Q_1 and Q_2 are homogeneous of degree 2. Now the relations

$$Q_1(a, b, 0, 0) = Q_2(a, b, 0, 0) = 0$$

show that $Q_1(a, b, c, d)$ and $Q_2(a, b, c, d)$ have no a^2 or b^2 terms. From the commutative property it follows that Q_1 and Q_2 have no c^2 or d^2 terms. Thus Q_1 and Q_2 are linear in each argument separately, as also are P_1 and P_2, and so we can apply Theorem 3 to complete the proof.

6. Linearity not necessary. Here we show that $(R \times R, \oplus, \odot)$ with operations defined by (II) may be a field with the usual zero and unity even though P_1, P_2, Q_1 and Q_2 are not linear in the separate arguments. For let T be any polynomial in one variable with real coefficients and set $S(x) = x(x-1)T(x)$. Define the mapping ϕ by

$$\phi\colon (a, b) \to (a + S(b), b) .$$

Then ϕ is a one-to-one mapping of C onto itself which leaves $(0, 0)$ and $(1, 0)$ invariant. Thus if we define

$$(a, b) \oplus (c, d) = \phi^{-1}(\phi(a, b) + \phi(c, d)),$$
$$(a, b) \odot (c, d) = \phi^{-1}(\phi(a, b), \phi(c, d)),$$

we get $(R \times R, \oplus, \odot)$ isomorphic to $(C, +, \cdot)$, the two field representations having common zero and unity. It is clear that the polynomials P_1, P_2, Q_1 and Q_2 may be given arbitrarily high degrees by the proper choice of T.

7. A general theorem. A question left unanswered in the preceding three sections is whether any field satisfying (II) must be isomorphic to the complex numbers. That the answer is yes is a special case of the following result.

THEOREM 5. *Let f and g be continuous mappings from $R^n \times R^n$ into R^n, and suppose that the binary operations $\oplus$ and $\odot$ defined on R^n by*

$$x \oplus y = f(x, y), x \odot y = g(x, y)$$

make $(R^n, \oplus, \odot)$ a field. Then $n = 1$ or 2 and the field is the real field or the field of complex numbers accordingly.

Proof. Let $\ominus x$ and x^* denote the inverses of x under $\oplus$ and $\odot$ respectively. We will show that the maps

$$x \to \ominus x \quad \text{and} \quad x \to x^*$$

are continuous and thus $(R^n, \oplus, \odot)$ is a topological field. Then the known result that any locally compact connected topological field satisfying the first axiom of countability is either the real or the complex numbers will yield the theorem; cf. [3, p. 173].

Consider the map $T: R^{2n} \to R^{2n}$ defined by $T: (x, y) \to (x, x \oplus y)$, where x and y belong to R^n. It is easily seen that T is continuous, one-to-one and onto. It is claimed that T is a homeomorphism. For suppose that A is an open subset of R^{2n} and $a \in A$. Let K be a compact neighborhood of a contained in A. Then T is a homeomorphism of K onto $T[K]$ and so by Brouwer's theorem [1, p. 100] on the invariance of domains the interior of K maps onto an open set. Thus $T(a)$ is an interior point of $T[A]$; we see that T takes open sets onto open sets.

Now T^{-1} is the mapping $(x, s) \to (x, s \ominus x)$, and so, letting s be the additive identity of $(R^n, \oplus, \odot)$, we see that the map $x \to \ominus x$ is continuous. The verification that $x \to x^*$ is a continuous map runs along the same lines. Thus with the usual topology $(R^n, \oplus, \odot)$ is either the reals or the complexes. Since R^m homeomorphic to R^n implies $m = n$, the theorem follows.

REFERENCES

1. S. Lefschetz, *Topology*, 2nd edition, Chelsea, New York, 1959.
2. L. J. Paige and J. D. Swift, *Elements of Linear Algebra*, Ginn, Boston, 1961.
3. L. Pontrjagin, *Topological Groups*, Princeton, 1939.
4. B. L. van der Waerden, *Modern Algebra*, vol. 1, Unger, 1949.

UNIVERSITY OF OREGON