4.3. Lagrange Interpolation

Throughout this section we shall assume F is a fixed field and that $t_0, t_1, \ldots, t_n$ are $n + 1$ *distinct* elements of F. Let V be the subspace of $F[x]$ consisting of all polynomials of degree less than or equal to n (together with the 0-polynomial), and let L_i be the function from V into F defined for f in V by

$$L_i(f) = f(t_i), \qquad 0 \leq i \leq n.$$

By part (i) of Theorem 2, each L_i is a linear functional on V, and one of the things we intend to show is that the set consisting of $L_0, L_1, \ldots, L_n$ is a basis for V^*, the dual space of V.

Of course in order that this be so, it is sufficient (cf. Theorem 15 of Chapter 3) that $\{L_0, L_1, \ldots, L_n\}$ be the dual of a basis $\{P_0, P_1, \ldots, P_n\}$ of V. There is at most one such basis, and if it exists it is characterized by

$$L_j(P_i) = P_i(t_j) = \delta_{ij}. \tag{4-11}$$

The polynomials

$$\begin{aligned} P_i &= \frac{(x - t_0) \cdots (x - t_{i-1})(x - t_{i+1}) \cdots (x - t_n)}{(t_i - t_0) \cdots (t_i - t_{i-1})(t_i - t_{i+1}) \cdots (t_i - t_n)} \\ &= \prod_{j \neq i} \left(\frac{x - t_j}{t_i - t_j}\right) \end{aligned} \tag{4-12}$$

are of degree n, hence belong to V, and by Theorem 2, they satisfy (4-11).

If $f = \sum\limits_i c_iP_i$, then for each j

$$f(t_j) = \sum_i c_iP_i(t_j) = c_j. \tag{4-13}$$

Since the 0-polynomial has the property that $0(t) = 0$ for each t in F, it follows from (4-13) that the polynomials $P_0, P_1, \ldots, P_n$ are linearly independent. The polynomials $1, x, \ldots, x^n$ form a basis of V and hence the dimension of V is $(n + 1)$. So, the independent set $\{P_0, P_1, \ldots, P_n\}$ must also be a basis for V. Thus for each f in V

$$f = \sum_{i=0}^{n} f(t_i)P_i. \tag{4-14}$$

The expression (4-14) is called **Lagrange's interpolation formula.** Setting $f = x^j$ in (4-14) we obtain

$$x^j = \sum_{i=0}^{n} (t_i)^jP_i.$$

Now from Theorem 7 of Chapter 2 it follows that the matrix

$$\begin{bmatrix} 1 & t_0 & t_0^2 & \cdots & t_0^n \\ 1 & t_1 & t_1^2 & \cdots & t_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^n \end{bmatrix} \tag{4-15}$$

is invertible. The matrix in (4-15) is called a **Vandermonde matrix;** it is an interesting exercise to show directly that such a matrix is invertible, when $t_0, t_1, \ldots, t_n$ are $n + 1$ distinct elements of F.

If f is any polynomial over F we shall, in our present discussion, denote by $f^\sim$ the polynomial function from F into F taking each t in F into $f(t)$. By definition (cf. Example 4, Chapter 2) every polynomial function arises in this way; however, it may happen that $f^\sim = g^\sim$ for two polynomials f and g such that $f \neq g$. Fortunately, as we shall see, this unpleasant situation only occurs in the case where F is a field having only a finite number of distinct elements. In order to describe in a precise way the relation between polynomials and polynomial functions, we need to define the product of two polynomial functions. If f, g are polynomials over F, the product of $f^\sim$ and $g^\sim$ is the function $f^\sim g^\sim$ from F into F given by

$$(f^\sim g^\sim)(t) = f^\sim(t)g^\sim(t), \qquad t \text{ in } F. \tag{4-16}$$

By part (ii) of Theorem 2, $(fg)(t) = f(t)g(t)$, and hence

$$(fg)^\sim(t) = f^\sim(t)g^\sim(t)$$

for each t in F. Thus $f^\sim g^\sim = (fg)^\sim$, and is a polynomial function. At this point it is a straightforward matter, which we leave to the reader, to verify that the vector space of polynomial functions over F becomes a linear algebra with identity over F if multiplication is defined by (4-16).

Definition. *Let* F *be a field and let* $\mathfrak{A}$ *and* $\mathfrak{A}^\sim$ *be linear algebras over* F. *The algebras* $\mathfrak{A}$ *and* $\mathfrak{A}^\sim$ *are said to be* **isomorphic** *if there is a one-to-one mapping* $\alpha \to \alpha^\sim$ *of* $\mathfrak{A}$ *onto* $\mathfrak{A}^\sim$ *such that*

(a) $$(c\alpha + d\beta)^\sim = c\alpha^\sim + d\beta^\sim$$

(b) $$(\alpha\beta)^\sim = \alpha^\sim\beta^\sim$$

for all α, β *in* $\mathfrak{A}$ *and all scalars* c, d *in* F. *The mapping* $\alpha \to \alpha^\sim$ *is called an* **isomorphism** *of* $\mathfrak{A}$ *onto* $\mathfrak{A}^\sim$. *An isomorphism of* $\mathfrak{A}$ *onto* $\mathfrak{A}^\sim$ *is thus a vector-space isomorphism of* $\mathfrak{A}$ *onto* $\mathfrak{A}^\sim$ *which has the additional property* (b) *of 'preserving' products.*

EXAMPLE 4. Let V be an n-dimensional vector space over the field F. By Theorem 13 of Chapter 3 and subsequent remarks, each ordered basis $\mathfrak{B}$ of V determines an isomorphism $T \to [T]_{\mathfrak{B}}$ of the algebra of linear operators on V onto the algebra of $n \times n$ matrices over F. Suppose now that U is a fixed linear operator on V and that we are given a polynomial

$$f = \sum_{i=0}^{n} c_i x^i$$

with coefficients c_i in F. Then

$$f(U) = \sum_{i=0}^{n} c_i U^i$$

and since $T \to [T]_{\mathcal{B}}$ is a linear mapping

$$[f(U)]_{\mathcal{B}} = \sum_{i=0}^{n} c_i[U^i]_{\mathcal{B}}.$$

Now from the additional fact that

$$[T_1T_2]_{\mathcal{B}} = [T_1]_{\mathcal{B}}[T_2]_{\mathcal{B}}$$

for all T_1, T_2 in $L(V, V)$ it follows that

$$[U^i]_{\mathcal{B}} = ([U]_{\mathcal{B}})^i, \qquad 2 \leq i \leq n.$$

As this relation is also valid for $i = 0, 1$ we obtain the result that

$$[f(U)]_{\mathcal{B}} = f([U]_{\mathcal{B}}). \tag{4-17}$$

In words, if U is a linear operator on V, the matrix of a polynomial in U, in a given basis, is the same polynomial in the matrix of U.

Theorem 3. *If* F *is a field containing an infinite number of distinct elements, the mapping* f $\to$ f$^{\sim}$ *is an isomorphism of the algebra of polynomials over* F *onto the algebra of polynomial functions over* F.

Proof. By definition, the mapping is onto, and if f, g belong to $F[x]$ it is evident that

$$(cf + dg)^{\sim} = df^{\sim} + dg^{\sim}$$

for all scalars c and d. Since we have already shown that $(fg)^{\sim} = f^{\sim}g^{\sim}$, we need only show that the mapping is one-to-one. To do this it suffices by linearity to show that $f^{\sim} = 0$ implies $f = 0$. Suppose then that f is a polynomial of degree n or less such that $f' = 0$. Let $t_0, t_1, \ldots, t_n$ be any $n + 1$ distinct elements of F. Since $f^{\sim} = 0$, $f(t_i) = 0$ for $i = 0, 1, \ldots, n$, and it is an immediate consequence of (4-14) that $f = 0$. ∎

From the results of the next section we shall obtain an altogether different proof of this theorem.

Exercises

1. Use the Lagrange interpolation formula to find a polynomial f with real coefficients such that f has degree ≤ 3 and $f(-1) = -6$, $f(0) = 2$, $f(1) = -2$, $f(2) = 6$.

2. Let $\alpha, \beta, \gamma, \delta$ be real numbers. We ask when it is possible to find a polynomial f over R, of *degree not more than* 2, such that $f(-1) = \alpha$, $f(1) = \beta$, $f(3) = \gamma$ and $f(0) = \delta$. Prove that this is possible if and only if

$$3\alpha + 6\beta - \gamma - 8\delta = 0.$$

3. Let F be the field of real numbers,

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$p = (x - 2)(x - 3)(x - 1).$$

(a) Show that $p(A) = 0$.

(b) Let P_1, P_2, P_3 be the Lagrange polynomials for $t_1 = 2$, $t_2 = 3$, $t_3 = 1$. Compute $E_i = P_i(A)$, $i = 1, 2, 3$.

(c) Show that $E_1 + E_2 + E_3 = I$, $E_iE_j = 0$ if $i \neq j$, $E_i^2 = E_i$.

(d) Show that $A = 2E_1 + 3E_2 + E_3$.

4. Let $p = (x - 2)(x - 3)(x - 1)$ and let T be any linear operator on R^4 such that $p(T) = 0$. Let P_1, P_2, P_3 be the Lagrange polynomials of Exercise 3, and let $E_i = P_i(T)$, $i = 1, 2, 3$. Prove that

$$E_1 + E_2 + E_3 = I, \qquad E_iE_j = 0 \quad \text{if} \quad i \neq j,$$

$$E_i^2 = E_i, \quad \text{and} \quad T = 2E_1 + 3E_2 + E_3.$$

5. Let n be a positive integer and F a field. Suppose A is an $n \times n$ matrix over F and P is an invertible $n \times n$ matrix over F. If f is any polynomial over F, prove that

$$f(P^{-1}AP) = P^{-1}f(A)P.$$

6. Let F be a field. We have considered certain special linear functionals on $F[x]$ obtained via 'evaluation at t':

$$L(f) = f(t).$$

Such functionals are not only linear but also have the property that $L(fg) = L(f)L(g)$. Prove that if L is any linear functional on $F[x]$ such that

$$L(fg) = L(f)L(g)$$

for all f and g, then either $L = 0$ or there is a t in F such that $L(f) = f(t)$ for all f.

4.4. Polynomial Ideals

In this section we are concerned with results which depend primarily on the multiplicative structure of the algebra of polynomials over a field.

Lemma. *Suppose* f *and* d *are non-zero polynomials over a field* F *such that deg* d $\leq$ *deg* f. *Then there exists a polynomial* g *in* F[x] *such that either*

$$f - dg = 0 \quad \textit{or} \quad \deg (f - dg) < \deg f.$$

Proof. Suppose

$$f = a_mx^m + \sum_{i=0}^{m-1} a_ix^i, \qquad a_m \neq 0$$

and that

$$d = b_nx^n + \sum_{i=0}^{n-1} b_ix^i, \qquad b_n \neq 0.$$

Then $m \geq n$, and

$$f - \left(\frac{a_m}{b_n}\right)x^{m-n}d = 0 \quad \text{or} \quad \deg\left[f - \left(\frac{a_m}{b_n}\right)x^{m-n}d\right] < \deg f.$$

Thus we may take $g = \left(\frac{a_m}{b_n}\right) x^{m-n}$. ∎

Using this lemma we can show that the familiar process of 'long division' of polynomials with real or complex coefficients is possible over any field.

Theorem 4. *If* f, d *are polynomials over a field* F *and* d *is different from* 0 *then there exist polynomials* q, r *in* F[x] *such that*

(i) $f = dq + r$.
(ii) *either* $r = 0$ *or* $\deg r < \deg d$.

The polynomials q, r *satisfying* (i) *and* (ii) *are unique.*

Proof. If f is 0 or $\deg f < \deg d$ we may take $q = 0$ and $r = f$. In case $f \neq 0$ and $\deg f \geq \deg d$, the preceding lemma shows we may choose a polynomial g such that $f - dg = 0$ or $\deg (f - dg) < \deg f$. If $f - dg \neq 0$ and $\deg (f - dg) \geq \deg d$ we choose a polynomial h such that $(f - dg) - dh = 0$ or

$$\deg [f - d(g + h)] < \deg (f - dg).$$

Continuing this process as long as necessary, we ultimately obtain polynomials q, r such that $r = 0$ or $\deg r < \deg d$, and $f = dq + r$. Now suppose we also have $f = dq_1 + r_1$ where $r_1 = 0$ or $\deg r_1 < \deg d$. Then $dq + r = dq_1 + r_1$, and $d(q - q_1) = r_1 - r$. If $q - q_1 \neq 0$ then $d(q - q_1) \neq 0$ and

$$\deg d + \deg (q - q_1) = \deg (r_1 - r).$$

But as the degree of $r_1 - r$ is less than the degree of d, this is impossible and $q - q_1 = 0$. Hence also $r_1 - r = 0$. ∎

Definition. *Let* d *be a non-zero polynomial over the field* F. *If* f *is in* F[x], *the preceding theorem shows there is at most one polynomial* q *in* F[x] *such that* $f = dq$. *If such a* q *exists we say that* d **divides** f, *that* f *is* **divisible** *by* d, *that* f *is a* **multiple** *of* d, *and call* q *the* **quotient** *of* f *and* d. *We also write* $q = f/d$.

Corollary 1. *Let* f *be a polynomial over the field* F, *and let* c *be an element of* F. *Then* f *is divisible by* $x - c$ *if and only if* $f(c) = 0$.

Proof. By the theorem, $f = (x - c)q + r$ where r is a scalar polynomial. By Theorem 2,

$$f(c) = 0q(c) + r(c) = r(c).$$

Hence $r = 0$ if and only if $f(c) = 0$. ∎

Definition. *Let* F *be a field. An element* c *in* F *is said to be a* **root** *or a* **zero** *of a given polynomial* f *over* F *if* f(c) = 0.

Corollary 2. *A polynomial* f *of degree* n *over a field* F *has at most* n *roots in* F.

Proof. The result is obviously true for polynomials of degree 0 and degree 1. We assume it to be true for polynomials of degree $n - 1$. If a is a root of f, $f = (x - a)q$ where q has degree $n - 1$. Since $f(b) = 0$ if and only if $a = b$ or $q(b) = 0$, it follows by our inductive assumption that f has at most n roots. ∎

The reader should observe that the main step in the proof of Theorem 3 follows immediately from this corollary.

The formal derivatives of a polynomial are useful in discussing multiple roots. The **derivative** of the polynomial

$$f = c_0 + c_1x + \cdots + c_nx^n$$

is the polynomial

$$f' = c_1 + 2c_2x + \cdots + nc_nx^{n-1}.$$

We also use the notation $Df = f'$. Differentiation is linear, that is, D is a linear operator on $F[x]$. We have the higher order formal derivatives $f'' = D^2f$, $f^{(3)} = D^3f$, and so on.

Theorem 5 (Taylor's Formula). *Let* F *be a field of characteristic zero,* c *an element of* F, *and* n *a positive integer. If* f *is a polynomial over* f *with* deg f ≤ n, *then*

$$\mathrm{f} = \sum_{\mathrm{k}=0}^{\mathrm{n}} \frac{(\mathrm{D^kf})}{\mathrm{k}!}(\mathrm{c})(\mathrm{x} - \mathrm{c})^{\mathrm{k}}.$$

Proof. Taylor's formula is a consequence of the binomial theorem and the linearity of the operators $D, D^2, \ldots, D^n$. The binomial theorem is easily proved by induction and asserts that

$$(a + b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k}\, b^k$$

where

$$\binom{m}{k} = \frac{m!}{k!(m-k)!} = \frac{m(m-1)\cdots(m-k+1)}{1 \cdot 2 \cdots k}$$

is the familiar binomial coefficient giving the number of combinations of m objects taken k at a time. By the binomial theorem

$$\begin{aligned} x^m &= [c + (x - c)]^m \\ &= \sum_{k=0}^{m} \binom{m}{k} c^{m-k}(x - c)^k \\ &= c^m + mc^{m-1}(x - c) + \cdots + (x - c)^m \end{aligned}$$