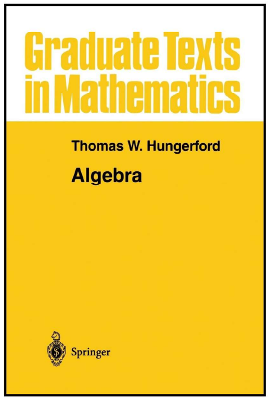
Modern Algebra

Chapter III. Rings

III.5. Rings of Polynomials and Formal Power Series—Proofs of Theorems



Theorem III.5.5

Theorem III.5.5. Let R and S be commutative rings with identity and $\varphi : R \to S$ is a homomorphism of rings such that $\varphi(1_R) = 1_S$. If $s_1, s_2, \ldots, s_n \in S$ then there is a unique homomorphism of rings $\overline{\varphi} : R[x_1, x_2, \ldots, x_n] \to S$ such that $\overline{\varphi}|_R = \varphi$ and $\overline{\varphi}(x_i) = s_i$ for $i = 1, 2, \ldots, n$. This property (that is, the mapping properties of φ and $\overline{\varphi}$; Hungerford calls this "a universal mapping property") completely determines the polynomial ring $R[x_1, x_2, \ldots, x_n]$ up to isomorphism.

Proof. If $f \in R[x_1, x_2, \ldots, x_n]$ then by Theorem III.5.4(v) $f = \sum_{i=0}^m a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}$ for some $a_i \in R$ and $k_{ij} \in \mathbb{N}$ (we omit x_j^0 terms). As described above, $\overline{\varphi}(f) = \varphi(f(s_1, s_2, \ldots, s_n)) = \sum_{i=0}^m \varphi(a_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}}$ is well-defined and $\overline{\varphi}_R = \varphi$ and $\overline{\varphi}(x_i) = s_i$. Now we show that $\overline{\varphi}$ is a ring homomorphism. Let $f = \sum_{i=0}^m a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}$ and $g = \sum_{i=0}^m b_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}$ (we include the x_i with 0 exponent here).

Theorem III.5.5 (continued 1)

Proof(continued). Then

$$\begin{aligned}
\overline{\varphi}(f+g) &= \overline{\varphi}\left(\sum_{i=0}^m a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}} + \sum_{i=0}^m b_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right) \\
&= \overline{\varphi}\left(\sum_{i=0}^m (a_i + b_i) x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right) \text{ by the definition} \\
&\qquad \text{of } + \text{ in } R[x_1, x_2, \ldots, x_n] \\
&= \varphi\left(\sum_{i=0}^m (a_i + b_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}}\right) \text{ by the definition of } \overline{\varphi} \\
&= \sum_{i=0}^m \varphi(a_i + b_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}} \text{ by the definition of } \varphi
\end{aligned}$$

Theorem III.5.5 (continued 2)

Proof(continued). Then

$$\begin{aligned}
&= \sum_{i=0}^m \varphi(a_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}} + \sum_{i=0}^m \varphi(b_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}} \\
&\qquad \text{since } \varphi \text{ is a homomorphism} \\
&= \varphi\left(\sum_{i=0}^m a_i s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}}\right) + \varphi\left(\sum_{i=0}^m b_i s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}}\right) \\
&\qquad \text{by the definition of } \varphi \\
&= \overline{\varphi}\left(\sum_{i=0}^m a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right) + \overline{\varphi}\left(\sum_{i=0}^m b_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right) \\
&\qquad \text{by the definition of } \overline{\varphi} \\
&= \overline{\varphi}(f) + \overline{\varphi}(g).
\end{aligned}$$

Theorem III.5.5 (continued 3)

Proof(continued). Next, "we find" that

$$\overline{\varphi}(fg) = \overline{\varphi}\left(\left(\sum_{i=0}^{m} a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right)\left(\sum_{i=0}^{m} b_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right)\right)$$
$$= \cdots = \overline{\varphi}(f)\overline{\varphi}(g)$$

by the Binomial Theorem (Theorem III.1.6), the rules of exponents as given in Theorem III.5.4(iii,iv) and the fact that φ is a homomorphism. So $\overline{\varphi}$ is a ring homomorphism. Suppose that $\psi : R[x_1, x_2, \ldots, x_n] \to S$ is a homomorphism such that $\psi|_R = \varphi$ and $\psi(x_i) = s_i$ for all i. Then

$$\begin{aligned}\psi(f) &= \psi\left(\sum_{i=0}^{m} a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}\right)\\ &= \sum_{i=0}^{m} \psi(a_i)\psi(x_1^{k_{i1}})\psi(x_2^{k_{i2}}) \cdots \psi(x_n^{k_{in}})\\ &\qquad \text{since } \psi \text{ is a ring homomorphism}\end{aligned}$$

Theorem III.5.5 (continued 4)

Proof(continued).

$$\begin{aligned}\psi(f) &= \sum_{i=0}^{m} \psi(a_i)(\psi(x_1))^{k_{i1}}(\psi(x_2))^{k_{i2}} \cdots (\psi(x_n))^{k_{in}}\\ &\qquad \text{since } \psi \text{ is a ring homomorphism}\\ &= \sum_{i=0}^{m} \varphi(a_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}} \text{ by hypotheses on the } \psi \text{ values}\\ &= \varphi(f(s_1, s_2, \ldots, s_n)) \text{ by definition of } \varphi\\ &= \overline{\varphi}(f) \text{ by definition of } \overline{\varphi}.\end{aligned}$$

Whence $\psi = \overline{\varphi}$ and $\overline{\varphi}$ is unique.

Theorem III.5.5 (continued 5)

Proof(continued). Finally, in order to show that $R[x_1, x_2, \ldots, x_n]$ is completely determined by the property $\overline{\varphi}|_R = \varphi$ and $\psi(x_i) = s_i$, define category $\mathcal{C}$ whose objects are all $(n+2)$-tuples $(\psi, K, s_1, s_2, \ldots, s_n)$ where K is a commutative ring with identity, $s_i \in K$, and $\psi : R \to K$ is a homomorphism with $\psi(1_R) = 1_K$. A morphism in $\mathcal{C}$ from $(\psi, J, s_1, s_2, \ldots, s_n)$ to $(\theta, T, t_1, t_2, \ldots, t_n)$ is a homomorphism of rings $\zeta : K \to T$ such that $\zeta(1_K) = 1_T$, $\zeta\psi = \theta$, and $\zeta(s_i) = t_i$. Since these morphisms are functions then the definition of "category" (Definition I.7.1) is satisfied (compositions, associativity, identity). Recall that a morphism is an equivalence if it has a left and right inverse. So a morphism is one to one if and only if it has a left inverse by Theorem 0.3.1(i); a morphism is onto if and only it it has a right inverse by Theorem 0.3.1(ii). Hence, a morphism is an equivalence if and only if it is one to one and onto; that is, if and only if it is a ring isomorphism. Let $\iota : R \to R[x_1, x_2, \ldots, x_n]$ be the inclusion map which maps each $r \in R$ to the "constant polynomial" $r \in R[x_1, x_2, \ldots, x_n]$.

Theorem III.5.5 (continued 6)

Proof(continued). Consider $(\iota, R[x_1, x_2, \ldots, x_n], x_1, x_2, \ldots, x_n)$ in $\mathcal{C}$. For any $(\psi, K, s_1, s_2, \ldots, s_n) \in \mathcal{C}$ we know by the first paragraph of the proof, since $\psi : R \to K$ is a ring homomorphism (φ of the first paragraph) then there is a unique $\overline{\psi} : R[x_1, x_2, \ldots, x_n] \to K$ a ring homomorphism with $\overline{\psi}|_R = \psi$ and $\overline{\psi}(x_i) = s_i$. Notice that $\overline{\psi}(1_{R[x_1, x_2, \ldots, x_n]}) = \psi(1_R) = 1_K$ and $\overline{\psi}\iota = \psi$ (since $\overline{\psi}\iota$ is literally $\overline{\psi}$ restricted to R). So $\overline{\psi}$ is a morphism from $(\iota, R[x_1, x_2, \ldots, x_n], x_1, x_2, \ldots, x_n)$ to $(\psi, K, s_1, s_2, \ldots, s_n)$ and $\overline{\psi}$ is a unique such morphism. So $(\iota, R[x_1, x_2, \ldots, x_n], x_1, x_2, \ldots, x_n)$ is a universal object in $\mathcal{C}$ (by definition, since the morphism $\overline{\psi}$ exists for any object in $\mathcal{C}$ and is unique). By Theorem I.7.10, any two universal objects in $\mathcal{C}$ are equivalent (and equivalence here corresponds to a ring isomorphism, as explained above). "This property" (that is, the mapping properties of φ and $\overline{\varphi}$) therefore determine $R[x_1, x_2, \ldots, x_n]$ up to isomorphism. □

Corollary III.5.6

Corollary III.5.6. If $\varphi : R \to S$ is a homomorphism of commutative rings and $s_1, s_2, \ldots, s_n \in S$, then the map $R[x_1, x_2, \ldots, x_n] \to S$, where $f = \sum_{i=0}^{m} a_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}}$ is mapped to $\overline{\varphi}(f) = \varphi(f(s_1, s_2, \ldots, s_n)) = \sum_{i=0}^{m} \varphi(a_i) s_1^{k_{i1}} s_2^{k_{i2}} \cdots s_n^{k_{in}}$, is a homomorphism of rings.

Proof. This is just the first paragraph of the proof of Theorem III.5.5 (without the uniqueness part; we may not have rings with identity here, but the presence of an identity is not used in this part of the proof of Theorem III.5.5). □

Corollary III.5.7

Corollary III.5.7. Let R be a commutative ring with identity and n a positive integer. For each k (with $1 \le k < n$) there are isomorphic rings

$$R[x_1, x_2, \ldots, x_k][x_{k+1}, x_{k+2}, \ldots, x_n] \cong R[x_1, x_2, \ldots, x_n]$$

$$\cong R[x_{k+1}, x_{k+2}, \ldots, x_n][x_1, x_2, \ldots, x_k].$$

Proof. Let S be a commutative ring with identity and $\varphi : R \to S$ a ring homomorphism. Let $s_1, s_2, \ldots, s_n \in S$. By Theorem III.5.5 there exists a ring homomorphism $\overline{\varphi} : R[x_1, x_2, \ldots, x_k] \to S$ such that $\overline{\varphi}|_R = \varphi$ and $\varphi(x_i) = s_i$. Applying Theorem III.5.5 to ring $R[x_1, x_2, \ldots, x_k]$ and homomorphism $\overline{\varphi} : R[x_1, x_2, \ldots, x_k] \to S$, there is a homomorphism $\overline{\overline{\varphi}} : (R[x_1, x_2, \ldots, x_k])[x_{k+1}, x_{k+2}, \ldots, x_n] \to S$ such that $\overline{\overline{\varphi}}|_{R[x_1, x_2, \ldots, x_k]} = \overline{\varphi}$ and $\overline{\overline{\varphi}}(x_i) = s_i$. Suppose that $\psi : R[x_1, x_2, \ldots, x_k][x_{k+1}, x_{k+2}, \ldots, x_n] \to S$ is a homomorphism such that $\psi|_R = \varphi$ and $\psi(x_i) = s_i$. Then the uniqueness argument of Theorem III.5.5 (paragraph 1 of the proof) holds to show that $\psi|_{R[x_1, x_2, \ldots, x_n]} = \overline{\overline{\varphi}}$.

Corollary III.5.7 (continued)

Corollary III.5.7. Let R be a commutative ring with identity and n a positive integer. For each k (with $1 \le k < n$) there are isomorphic rings

$$R[x_1, x_2, \ldots, x_k][x_{k+1}, x_{k+2}, \ldots, x_n] \cong R[x_1, x_2, \ldots, x_n]$$

$$\cong R[x_{k+1}, x_{k+2}, \ldots, x_n][x_1, x_2, \ldots, x_k].$$

Proof (continued). Consequently, $R[x_1, x_2, \ldots, x_k][x_{k+1}, x_{k+2}, \ldots, x_n]$ has the desired "universal mapping property" (i.e., the mapping properties of φ and $\overline{\overline{\varphi}}$), so by Theorem III.5.5, $R[x_1, x_2, \ldots, x_k][x_{k+1}, x_{k+2}, \ldots, x_n] \cong R[x_1, x_2, \ldots, x_n]$. The other isomorphism is similar. □

Proposition III.5.9

Proposition III.5.9. Let R be a ring with identity and $f = \sum_{i=0}^{\infty} a_i x^i \in R[[x]]$.

(i) f is a unit in $R[[x]]$ if and only if its constant term a_0 is a unit in R.

(ii) If a_0 is irreducible in R, then f is irreducible in $R[[x]]$.

Proof. (i) Suppose f is a unit. Then there exists $g = \sum_{i=0}^{\infty} b_i x^i \in R[[x]]$ such that $fg = gf = 1_R \in R[[x]]$. Then $a_0 b_0 = b_0 a_0 = 1_R$, and so a_0 is a unit in R. Conversely, suppose a_0 is a unit in R. With $g = \sum_{i=0}^{\infty} b_i x^i \in R[[x]]$ where $fg = 1_R$ we have the following equations satisfied:

$$\begin{aligned} a_0 b_0 &= 1_R \\ a_0 b_1 + a_1 b_0 &= 0 \\ &\vdots \\ a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 &= 0 \\ &\vdots \end{aligned}$$

Proposition III.5.9 (continued 1)

Proposition III.5.9. Let R be a ring with identity and $f = \sum_{i=0}^{\infty} a_i x^i \in R[[x]]$.

(i) f is a unit in $R[[x]]$ if and only if its constant term a_0 is a unit in R.

Proof (continued). (i) Conversely, if the system of equations is satisfied by $(b_0, b_1, \ldots)$ then $g = \sum_{i=0}^{\infty} b_i x^i \in R[[x]]$ satisfies $fg = 1_R$ in $R[[x]]$. Now we show there is a solution and hence g is a right inverse of f. Since a_0 is a unit there is a solution to the first equation, namely $b_0 = a_0^{-1}$. Then we can solve the second equation to get $b_1 = a_0^{-1}(-a_1 b_0) = -a_0^{-1}(a_1 a_0^{-1})$. Inductively, we can find each $b_n = a_0(-a_1 b_{n-1} - a_2 b_{n-2} - \cdots - a_n b_0)$ (in terms of $a_0^{-1}, a_1, a_2, \ldots, a_n$ and $b_0, b_1, \ldots, b_{n-1}$). We can then (inductively) express each b_n in terms of the a_i's above. Therefore, there exists $g \in R[[x]]$ such that $fg = 1_R$. Similarly, there exists $h \in R[[x]]$ such that $hf = 1_R$. But then $h = h1_R = h(fg) = (hf)g = 1_R g = g$. So (i) follows.

Proposition III.5.9 (continued 2)

Proposition III.5.9. Let R be a ring with identity and $f = \sum_{i=0}^{\infty} a_i x^i \in R[[x]]$.

(ii) If a_0 is irreducible in R, then f is irreducible in $R[[x]]$.

Proof. (ii) Recall that f a nonzero nonunit in a ring is irreducible if $f = gh$ implies that either g or h is a unit. With $f = \sum_{i=0}^{\infty} a_i x^i$, $g = \sum_{i=0}^{\infty} b_i x^i$, $h = \sum_{i=0}^{\infty} c_i x^i$, $f = gh$ implies $a_0 = b_0 c_0$. If a_0 is irreducible then either b_0 or c_0 is a unit. So by (i), either g or h is a unit. Therefore, f is irreducible. □

Corollary III.5.10

Corollary III.5.10. If R is a division ring, then the units in $R[[x]]$ are precisely those power series with nonzero constant terms. The principal ideal (x) consists precisely of the nonunits in $R[[x]]$ and is the unique maximal ideal of $R[[x]]$. Thus if R is a field, $R[[x]]$ is a local ring.

Proof. First, if R is a division ring then each nonzero element of R is a unit. So by Proposition III.5.9(i), a formal power series is a unit if and only if the constant term is nonzero.
Now $x = (0, 1_R, 0, \ldots)$ commutes with every element of $R[[x]]$, so x is in the center of $R[[x]]$ and $(x) = \{xf \mid f \in R[[x]]\}$ (by Theorem III.2.5(iii)). Consequently, every nonzero element xf of (x) has zero constant term, whence by Proposition III.5.9(i), xf is a nonunit. Conversely, for every nonunit $f \in R[[x]]$, by Theorem III.5.9(i), we have $f = \sum_{i=0}^{\infty} a_i x^i$ with $a_0 = 0$. Let $g = \sum_{i=0}^{\infty} b_i x^i$ where $b_i = a_{i+1}$. Then $xg = f$ whence $f \in (x)$. So (x) consists precisely of the nonunits in $R[[x]]$.

Corollary III.5.10 (continued)

Corollary III.5.10. If R is a division ring, then the units in $R[[x]]$ are precisely those power series with nonzero constant terms. The principal ideal (x) consists precisely of the nonunits in $R[[x]]$ and is the unique maximal ideal of $R[[x]]$. Thus if R is a field, $R[[x]]$ is a local ring.

Proof (continued). Finally, since $1_R \notin (x)$ by the first claim of this result then $(x) \neq R[[x]]$. Furthermore, every ideal I of $R[[x]]$ with $I \neq R[[x]]$ must contain no units (see "Remark" on page 123 or the "Note" on page 2 of the class notes for Section II.2). So I consists only of nonunits. Since (x) is the set of all nonunits by the previous paragraph, then $I \subset (x)$. Thus every ideal of $R[[x]]$ (except $R[[x]]$ itself) is contained in (x) and so (x) is the only maximal ideal of $F[[x]]$. □