Skew Polynomial Rings

John A. Beachy

NIU
October 31, 2018
Example 1: \( \mathbb{C}[z] \)

If \( R \) is any ring (not necessarily commutative), we can define the polynomial ring \( R[x] \), in which \( x \) commutes with elements of \( R \), and each element of \( R[x] \) has the form \( \sum_{i=0}^{n} a_i x^i \), for \( a_i \in R \).

Our goal is to study general “polynomial” rings in which the indeterminate need not commute with elements from the coefficient ring \( R \).

Consider the ring \( \mathbb{C}[z] \) of polynomials over the field \( \mathbb{C} \) of complex numbers. For \( c \in \mathbb{C} \) define \( zc = \bar{c}z \), where \( \bar{c} \) is the complex conjugate of \( c \). There are axioms to check, but this does define a noncommutative ring structure on \( \mathbb{C}[z] \). Example: for \( i \) we have \((iz)^2 = (iz)(iz) = i(zi)z = i(-iz)z = -i^2z^2 = z^2 \), whereas in the ordinary polynomial ring we would have \((iz)^2 = -z^2 \). Note that complex conjugation defines an automorphism of \( \mathbb{C} \).
Example 2: \( \mathbb{C}[z; D] \)

Consider the homogeneous linear differential equation

\[
a_n(z) \frac{d^n f}{dz^n} + \cdots + a_1(z) \frac{df}{dz} + a_0(z) f = 0,
\]

where the solution \( f(z) \) is a polynomial with complex coefficients, and the terms \( a_i(z) \) also belong to \( \mathbb{C}[z] \). The equation can be written in compact form as \( L(f) = 0 \), where \( L \) is the differential operator

\[
L = a_n(z) D^n + \cdots + a_1(z) D + a_0(z),
\]

with \( D = d/dz \). The differential operator can be thought of as a polynomial in the two indeterminates \( z \) and \( D \), but in this case the indeterminates do not commute, since

\[
Dz[f(z)] = D(zf(z)) = f(z) + zD(f(z)) = (zD + 1)[f(z)].
\]
The composite of two differential operators can be written as

\[ a_n(z)D^n + \cdots + a_1(z)D + a_0(z). \]

The resulting ring, with elements written in this standard form, is denoted by \( \mathbb{C}[z][D] \) or \( \mathbb{C}[z; D] \), and is called the **ring of differential operators**.

In taking the product of \( a_n(z)D^n + \cdots + a_1(z)D + a_0(z) \) and \( b_m(z)D^m + \cdots + b_1(z)D + b_0(z) \) we have

\[ D \cdot b_m(z)D^m = b_mD^{m+1} + a'_n(z)D^m, \]

so the leading term of the product is \( a_n(z)b_m(z)D^{n+m} \). Thus the product of two nonzero elements is nonzero, so \( \mathbb{C}[z; D] \) is a noncommutative domain.
In $\mathbb{C}[x; D]$, for a polynomial $a(x) \in \mathbb{C}[x]$, we have the general calculation

\[ Da(z)[f(z)] = D(a(z)f(z)) = D(a(z))f(z) + a(z)D(f(z)) = (a(z)D + D(a(z)))[f(z)]. \]

It turns out to be convenient to write the identity we get as

\[ Da(z) = a(z)D + \frac{d}{dz}a(z), \]

separating out the action of $D$ on $a(z)$. 
In the noncommutative polynomial ring we hope to construct, we would like to be able to express each polynomial uniquely in the form \( f(x) = \sum a_i x^i \), for some \( a_i \in R \). We would also like multiplication to respect degrees in the usual way, so that we will have \( \deg(f(x)g(x)) \leq \deg(f(x)) + \deg(g(x)) \). Furthermore, \( x^n a \) should have degree \( n \), for any \( a \in R \). In particular, \( xa \) should have degree \( 1 \), and so we should have

\[
xa \in Rx + R.
\]

Thus

\[
xa = \tau(a)x + \delta(a),
\]

for some elements \( \tau(a) \) and \( \delta(a) \) in \( R \).
The distributive law $x(a + b) = xa + xb$ must be satisfied, so

$$
\tau(a + b)x + \delta(a + b) = \tau(a)x + \delta(a) + \tau(b)x + \delta(b)
$$

for all $a, b \in R$, and so the functions $\tau$ and $\delta$ must be additive. To preserve the associative law we need $x(ab) = (xa)b$, and so the following expressions must be equal for all $a, b \in R$.

$$
\tau(ab)x + \delta(ab) = x(ab) = (xa)b = (\tau(a)x + \delta(a))b = \tau(a)\tau(b)x + \tau(a)\delta(b) + \delta(a)b
$$

It follows that $\tau(ab)$ must equal $\tau(a)\tau(b)$, and so $\tau$ must be an endomorphism of $R$. Furthermore, $\delta(ab)$ must equal $\tau(a)\delta(b) + \delta(a)b$, and so $\delta$ must be a $\tau$-derivation of $R$. The pair $(\tau, \delta)$ is called a (left) skew derivation, or $\delta$ is called a $\tau$-derivation.
Remember that in our generalized “polynomial” ring, for the indeterminate $x$ and an element $a \in R$, we want to have

$$xa = \tau(a)x + \delta(a),$$

where $\tau$ is an endomorphism of $R$ and $\delta$ is a derivation on $R$. In the ring $\mathbb{C}[z; D]$ we had

$$Da(z) = a(z)D + \frac{d}{dz}a(z),$$

so in this case $D$ is our indeterminate, and since $R = \mathbb{C}[z]$, we have a derivation $\delta = \frac{d}{dz} : \mathbb{C}[z] \to \mathbb{C}[z]$, and the endomorphism $\tau : \mathbb{C}[z] \to \mathbb{C}[z]$ is just the identity mapping.
The definition

Definition

Let $S$ be a ring containing $R$ as a subring. We say that $S$ is a **skew polynomial ring** over $R$ or that $S$ is an **Ore extension** of $R$ if there exists an element $x \in S$ for which $S$ is a free left $R$-module with basis $1, x, x^2, \ldots$ such that $xR \subseteq Rx + R$.

In this case, there exist an endomorphism $\tau$ of $R$ and a $\tau$-derivation $\delta$ on $R$ such that $xa = \tau(a)x + \delta(a)$, for all $a \in R$. To summarize this data we write $S = R[x; \tau, \delta]$. 
Let $S = R[x; \tau, \delta]$ be a skew polynomial ring. Each nonzero element of $S$ can be expressed uniquely in the form $a_nx^n + \ldots + a_1x + a_0$. As usual, the integer $n$ is called the degree of the element, and $a_n$ is called the leading coefficient.

In multiplying two elements $\sum_{i=0}^n a_ix^i$ and $\sum_{i=0}^m b_ix^i$, the candidate for the leading coefficients is $a_nx^n \cdot b_mx^m = a_n\tau^n(b_m)x^{n+m} + \text{terms of lower degree}$, since $x \cdot b_mx^m = (\tau(b_m)x + \delta(b_m))x^m = \tau(b_m)x^{m+1} + \delta(b_m)x^m$.

If $R$ is a domain, and $\tau$ is injective, then $a_n\tau^n(b_m)$ is nonzero when $a_n$ and $b_m$ are nonzero, so the degree of a product is the sum of the degrees, In this case $S$ is a domain.
A somewhat curious fact

The conditions making the pair \((\tau, \delta)\) a skew derivation can also be expressed in the following way. If \(\tau, \delta\) are functions from \(R\) to \(R\), consider the function \(\phi : R \rightarrow M_2(R)\) defined for all \(r \in R\) by

\[
\phi(r) = \begin{bmatrix}
\tau(r) & \delta(r) \\
0 & r
\end{bmatrix}.
\]

Since

\[
\phi(r)\phi(s) = \begin{bmatrix}
\tau(r) & \delta(r) \\
0 & r
\end{bmatrix} \begin{bmatrix}
\tau(s) & \delta(s) \\
0 & s
\end{bmatrix} =
\begin{bmatrix}
\tau(r)\tau(s) & \tau(r)\delta(s) + \delta(r)s \\
0 & rs
\end{bmatrix},
\]

to have a ring homomorphism we need \(\delta(rs) = \tau(r)\delta(s) + \delta(r)s\).
Then \(\phi\) is a ring homomorphism iff the pair \((\tau, \delta)\) is a skew derivation.
Definition

Let $R$ be a ring, and let $\tau$ be an endomorphism of $R$. The skew polynomial ring $R[x; \tau]$ is defined to be the set of all left polynomials of the form $a_0 + a_1x + \ldots + a_nx^n$ with coefficients $a_0, \ldots, a_n$ in $R$. Addition is defined as usual, and multiplication is defined by using the relation $xa = \tau(a)x$, for all $a \in R$.

Theorem

Let $R$ be a ring, and let $\tau$ be an endomorphism of $R$. The set $R[x; \tau]$ of skew polynomials over $R$ is a ring.
Proof: It is clear that \( R[x; \tau] \) is a group under addition. The associative law holds for multiplication of monomials, as shown by the following computations:

\[
(ax^i \cdot bx^j) \cdot cx^k = (a\tau^i(b)x^{i+j}) \cdot cx^k = (a\tau^i(b))(x^{i+j} \cdot cx^k) \\
= (a\tau^i(b))(\tau^{i+j}(c)x^{i+j+k}) \\
= (a\tau^i(b))(\tau^{i+j}(c))x^{i+j+k}
\]

\[
ax^i \cdot (bx^j \cdot cx^k) = ax^i \cdot (b\tau^j(c)x^{j+k}) = a(x^i \cdot b\tau^j(c))x^{j+k} \\
= a(\tau^i(b\tau^j(c))x^i)x^{j+k} = a(\tau^i(b)\tau^i(\tau^j(c)))x^{i+j+k} \\
= a(\tau^i(b)\tau^{i+j}(c))x^{i+j+k} \\
= (a\tau^i(b))\tau^{i+j}(c)x^{i+j+k}
\]
We extend the definition of multiplication to all polynomials in \( R[ x; \tau ] \) by repeatedly using the distributive laws. Since \( \tau \) is a ring homomorphism, we have \( \tau(1) = 1 \), and so the constant polynomial 1 serves as a multiplicative identity element.

\[ \square \]

**Theorem**

Let \( K \) be a division ring, and let \( \tau \) be a nontrivial endomorphism of \( K \). Then the skew polynomial ring \( K[ x; \tau ] \) is a noncommutative domain in which every left ideal is a principal left ideal.

**Proof:** Let \( K[ x; \tau ] = S \), and let \( I \) be a nonzero left ideal of \( S \). Among the nonzero elements of \( I \) we can choose one of minimal degree \( m \), say \( p(x) = a_0 + \ldots + a_m x^m \).
Since $K$ is a division ring, we can assume without loss of generality that $a_m = 1$. (If not, consider the polynomial $a_m^{-1}p(x)$, which is also in $I$ and still has degree $m$.)

We claim that $I$ is the left ideal generated by $p(y)$, so that $I = S \cdot p(y)$. The proof is by induction on the degree of the nonzero elements of $I$. Let $f(y) = b_0 + \ldots + b_n y^n$ belong to $I$, with $\deg(f(y)) = n$. Since $n \geq m$, consider the polynomial $g(y) = f(y) - b_n y^{n-m} p(y)$. Because the leading coefficient of $p(y)$ is 1, the endomorphism $\tau$ has no effect on the product $y^{n-m} \cdot y^m$, and so the degree of $g(y)$ is strictly less than the degree of $f(y)$. If $n = m$, then we conclude from the choice of $p(y)$ that $g(y) = 0$, and so $f(y) \in S \cdot p(y)$. Now assume that the induction hypothesis holds for all elements of $I$ with degree $\leq k$, and that $n = k + 1$. Then it follows that $g(y)$ belongs to $I$, and so $f(y) = g(y) + b_n y^{n-m} p(y)$ belongs to $I$. □