Semidirect Products: \( x \mapsto ax + b \) as a First Example

SHREERAM S. ABHYANKAR
Purdue University
West Lafayette, IN 47907
ram@cs.purdue.edu

CHRIS CHRISTENSEN
Northern Kentucky University
Highland Heights, KY
christensen@nku.edu

After precalculus, mathematics students often leave behind the familiar family of transformations, \( x \mapsto ax + b \). We will show that this family, when examined in the right light, leads us to some interesting and important ideas in group theory.

By building on this accessible example, it is possible to introduce the semidirect product (a topic usually first seen at the graduate level) in an undergraduate abstract algebra course. After being introduced to the semidirect product, students are able to better understand the structures of some of the groups of small order that are typically discussed in a first course in abstract algebra.

**Constructions** Early in any study of abstract algebra, we learn how to construct new groups by taking the (external) direct product of two groups. For this construction, we need not put any restrictions on the operations of the groups we are combining because the group operation for the direct product is defined componentwise using the operations of the factors.

Here is an example that we will later use for comparison. Let \( \mathbb{R}^+ \) denote the additive group of the real numbers \( \mathbb{R} \), and let \( \mathbb{R}^\times \) denote the multiplicative group of the nonzero elements of \( \mathbb{R} \). Consider the external direct product \( \mathbb{R}^+ \times \mathbb{R}^\times \). The group operation on the subgroup \( \{(b, 1)\} \) looks like addition

\[(d, 1)(b, 1) = (d + b, 1),\]

and the group operation on the subgroup \( \{(0, a)\} \) looks like multiplication

\[(0, c)(0, a) = (0, ca),\]

but these operations merge into one operation for \( \mathbb{R}^+ \times \mathbb{R}^\times \)

\[(d, c)(b, a) = (d + b, ca).\]

Notice that \( \mathbb{R}^+ \times \mathbb{R}^\times \) is abelian because both \( \mathbb{R}^+ \) and \( \mathbb{R}^\times \) are abelian and the operation is defined componentwise. It is easy to show that the external direct product \( G = G_1 \times G_2 \) has the following properties: If \( H = \{(h, 1_{G_2})\} \) and \( K = \{(1_{G_1}, k)\} \), then \( H \approx G_1 \) and \( K \approx G_2 \) (where we denote isomorphism by \( \approx \)). Furthermore, \( HK = G; H \cap K = \{(1_{G_1}, 1_{G_2})\}, \) the identity in \( G; \) and \( H \) and \( K \) are normal in \( G \).

More interesting than constructing arbitrary external direct products is discovering which groups are internal direct products of two (or more) of their subgroups. Usually we encounter this early—at order four—when we notice that the Klein four-group \( V_4 \approx \mathbb{Z}_2 \times \mathbb{Z}_2 \). Later we learn in the fundamental theorem of finite abelian groups that every finite abelian group is the direct product of cyclic groups.
Another, more subtle way to construct groups is by the semidirect product. This interesting structure is not usually examined in undergraduate abstract algebra texts, although some of the groups we encounter most often in a typical undergraduate abstract algebra course are semidirect products—\( D_{2n} \), the dihedral group of order \( 2n \) (which consists of the symmetries of a regular \( n \)-gon), and \( S_n \), the symmetric group of degree \( n \) (which consists of the permutations of \( \{1, 2, \ldots, n\} \)). There is one notable exception; Goodman [8] has a section discussing the semidirect product. Hungerford [11] and Saracino [14] relegate the semidirect product to exercises: Hungerford defines the internal semidirect product and has an exercise to show that each of \( S_3, D_4 \), and \( S_4 \) is a semidirect product of two of its subgroups, and Saracino defines the external semidirect product and uses it in an exercise to show that if \( p \) and \( q \) are primes and \( p \) divides \( q - 1 \), then there exists a nonabelian group of order \( pq \).

It is certainly not necessary to wait until graduate school to encounter the semidirect product. We can take the first step toward it by carefully exploring \( x \mapsto ax + b \).

**Internal semidirect products** Consider the vector space \( \mathbb{R}^1 \) and the transformations

\[
x \mapsto ax + b.
\]

We will require \( a \neq 0 \), without which this mapping is not a transformation. Also, notice that if \( b \neq 0 \), then this is not a linear transformation of \( \mathbb{R}^1 \) (because 0 is not fixed). Let \( G \) be the set of transformations \( \{x \mapsto ax + b\} \). We will show that this is a group, where the operation is composition of transformations. First, check that we have closure under composition: if we take two maps, say, \( T_{a,b} : x \mapsto ax + b \) and \( T_{c,d} : x \mapsto cx + d \) and compose them, we get \( T_{c,d} \circ T_{a,b} : x \mapsto c(ax + (cb + d)), \) which is in \( G \), since \( ca \neq 0 \). Associativity is automatic because composition of arbitrary maps is associative. Taking \( a = 1 \) and \( b = 0 \) gives the identity transformation. Finally, we need to check for inverses, and we notice that a condition is required. If \( T_{a,b} : x \mapsto ax + b \), then \( T_{a,b}^{-1} : x \mapsto x/a - b/a \). Since \( a \neq 0 \), having \( a \) in the denominator causes no problems.

We call the mappings \( x \mapsto ax + b \) affine transformations and the group \( G \) of mappings \( x \mapsto ax + b \) with \( a \neq 0 \) the **affine general linear group** of \( \mathbb{R}^1 \).

Notice that an affine transformation \( T_{a,b} : x \mapsto ax + b \) does two distinct things to \( x \). It scales \( x \) by \( a \neq 0 \), and it translates by \( b \). Clearly, \( \mathbb{R}^+ \) is isomorphic to the subgroup \( N = \{x \mapsto ax + b : a = 1\} = \{x \mapsto x + b\} \), which are the translations; and \( \mathbb{R}^\times \) is isomorphic to the subgroup \( A = \{x \mapsto ax + b : a \neq 0, b = 0\} = \{x \mapsto ax : a \neq 0\} \), which are the scalings.

The group of affine linear transformations of \( \mathbb{R}^1 \) is a product of the two subgroups \( N = \{x \mapsto x + b\} \) and \( A = \{x \mapsto ax : a \neq 0\} \). But, it is not the direct product (we essentially constructed that earlier as \( \mathbb{R}^+ \times \mathbb{R}^\times \)); it is a semidirect product.

To explore this product, let us examine how the two subgroups \( N \) and \( A \) sit inside \( G \). First, notice that every affine transformation \( T_{a,b} : x \mapsto ax + b : a \neq 0 \) is made up by composing an element \( T_{1,0} \) from \( N = \{x \mapsto x + b\} \) with an element \( T_{a,0} \) from \( A = \{x \mapsto ax : a \neq 0\} \). Next, we notice that \( N \cap A = \{x \mapsto x\} = \{T_{1,0}\} \), the identity transformation. Finally, notice that \( N \) is normal in \( G \), denoted \( N \triangleleft G \) (because if \( T_{1,d} \in N \) and \( T_{a,b} \in G \), then \( T_{a,b} \circ T_{1,d} \circ T_{a,b}^{-1} : x \mapsto x + ad \in N \)). These three properties expressing how \( N \) and \( A \) sit inside \( G \) characterize the structure called an internal semidirect product.

**Definition.** A group \( G \) is the **internal semidirect product** of \( N \) by \( A \), which we denote by \( G = N \rtimes A \), when \( G \) contains subgroups \( N \) and \( A \) such that \( NA = G, \ N \cap A = \{1\}, \) and \( N \triangleleft G \).
Notice that the internal semidirect product is a generalization of the internal direct product, where we require that both subgroups be normal in G. The symmetry of N and A in the definition of the direct product \( N \times A \) is reflected in the use of the symmetric symbol \( \times \). The symbol for the semidirect product \( \rtimes \) is asymmetric to remind us that the two subgroups do not sit symmetrically in the semidirect product; the symbol points to the left to remind us of which subgroup is normal.

The normality of \( N \) enables the manipulation
\[
(n_1a_1)(n_2a_2) = (n_1a_1)(n_1(a_1^{-1}a_1)a_2) = (n_1(a_1n_2a_1^{-1}))(a_1a_2) \in NA,
\]
which verifies the closure of the operation in \( G = NA \).

Let us consider some familiar groups that are semidirect products.

The smallest order nonabelian group that we meet has order six—\( D_6 \) the symmetries of an equilateral triangle. Now, there are two groups of order six—\( D_6 \) and \( Z_6 \) (\( Z_6 \) denotes the cyclic group of order 6.) Each has a (normal) subgroup of order three, but \( Z_6 \) has only one subgroup of order two (hence, it is normal) while \( D_6 \) has three subgroups of order two (none of which is normal). \( Z_6 \cong Z_2 \times Z_3 \), but \( D_6 \approx N \rtimes A \) where \( N \) is the subgroup of order three (the rotations) and \( A \) is one of the subgroups of order two (one of the subgroups generated by a reflection). In general, \( D_{2n} \), the symmetries of a regular \( n \)-gon, is the semidirect product of its rotation subgroup by one of the subgroups generated by a reflection.

Recall that \( S_3 \approx D_6 \). The alternating group \( A_3 \), which consists of the even permutations corresponds to the subgroup of rotations of \( D_6 \). It is easy to see in general that \( S_n \) is the internal semidirect product of \( A_n \) by a subgroup of order two generated by one of the transpositions or two-cycles. (Gallian [7, p. 183] gives a construction of \( D_{12} \) as an internal direct product.)

Another semidirect product that students may have seen is the group of Euclidean isometries of the plane or of space. Each of these transformation groups is the semidirect product of the subgroup of translations by the subgroup of isometries that fix the origin. An interesting feature of the spatial case is that the second factor, namely the isometries of the sphere, is a noncommutative group.

Finally, using the internal semidirect product, we can generalize the definition of the affine general linear group of \( \mathbb{R}^1 \) to \( \mathbb{R}^n \) (or any finite dimensional vector space \( V \)—we do this near the end of this paper). Let \( A \) be the group of all nonsingular (that is, invertible) linear transformations of \( \mathbb{R}^n \); this group is called the general linear group of \( \mathbb{R}^n \) and is denoted by \( GL(\mathbb{R}^n) \). Consider transformations \( x \mapsto a(x) + b \) where \( a \in A \) and \( b \) is an \( n \)-dimensional vector of \( \mathbb{R}^n \). Notice that for \( \mathbb{R}^1 \), \( A \) is isomorphic to \( \{ x \mapsto ax : a \neq 0 \} \). We define the affine general linear group of \( \mathbb{R}^n \) by \( AGL(\mathbb{R}^n) = \{ x \mapsto a(x) + b : a \in GL(\mathbb{R}^n), b \in \mathbb{R}^n \} \). Notice that the affine general linear group of \( \mathbb{R}^n \) is the internal semidirect product of \( N \) by \( A \) where \( N \) denotes the normal subgroup of transformations \( \{ x \mapsto x + b \} \) of \( \mathbb{R}^n \).

**External semidirect products** It is easy to construct external direct products because the operations of the factors do not really merge. The external semidirect product is another matter. We need a clue to suggest to us when we can merge the operations of two unrelated groups into one operation in the semidirect product. Again, we carefully examine \( x \mapsto ax + b \) for a clue.

Recall that when we examined the affine general linear group of \( \mathbb{R}^1 \), we saw that \( N \approx \mathbb{R}^+ \) and \( A \approx \mathbb{R}^x \) do not play equal roles in the internal semidirect product \( G = N \rtimes A \); \( N \) is a normal subgroup. Consider conjugation of \( N \) by elements of \( A \). Let \( T_{1,d} \in N \) and \( T_{a,0} \in A \). Then \( T_{a,0} \circ T_{1,d} \circ T_{a,0}^{-1} : x \mapsto x + ad \in N \). Notice the \( ad \), but rather than thinking of \( a \) and \( d \) as elements of \( \mathbb{R} \) and \( ad \) as a product in \( \mathbb{R} \), think of...
Let $X$ and $A$ be groups, and let $\theta$ be a given action of $A$ on $X$; that is, a homomorphism $\theta : A \to \text{Aut}(X)$ (where $\text{Aut}(X)$ denotes the group of automorphisms of $X$). Then, for $c \in A, \theta(c) : X \to X$, and if $b \in X$, we denote its image under this automorphism by $\theta(c)(b)$. The external semidirect product $X \rtimes_\theta A$ of the group $X$ and the group $A$ relative to $\theta$ is, as a set, simply $X \times A$. We make this into a group by defining $(d, c)(b, a) = (d\theta(c)(b), ca)$. (When $\theta$ is clear, we will write $X \rtimes A$.)

The action $\theta$ becomes conjugation in the semidirect product, that is, $\theta(a)\theta(b) = (1, a)(b, 1)(1, a)^{-1}$. Let $\overline{X} = \{(b, 1) : b \in X\}$, and let $\overline{A} = \{(1, a) : a \in A\}$. Then $X \rtimes_\theta A$ is easily seen to be the internal semidirect product of $X$ by $A$.

If $\theta$ is trivial, that is, if the image of $\theta = 1$, then $X \rtimes_\theta A$ is simply the external direct product $X \times A$.

We now construct external semidirect products that are isomorphic to the two affine general linear groups we mentioned above.

Because each $c \in \mathbb{R}^\times$ acts on $\mathbb{R}^+$ as a group automorphism $c(b) = cb$, we can construct $\mathbb{R}^+ \rtimes \mathbb{R}^\times$. The elements of our group are the elements of the Cartesian product $\mathbb{R}^+ \times \mathbb{R}^\times$, and the group operation is $(d, c)(b, a) = (cb + d, ca)$. This group is isomorphic to the affine general linear group of $\mathbb{R}^1$. Notice that although $\mathbb{R}^+ \rtimes \mathbb{R}^\times$ are abelian, $\mathbb{R}^+ \rtimes \mathbb{R}^\times$ is not abelian (unlike $\mathbb{R}^+ \times \mathbb{R}^\times$).

Because $\text{GL}(\mathbb{R}^n)$ acts on $\mathbb{R}^n$, we can construct $\mathbb{R}^n \rtimes \text{GL}(\mathbb{R}^n)$. The elements of our group are the elements of the Cartesian product $\mathbb{R}^n \times \text{GL}(\mathbb{R}^n) = \{(b, a) : b \in \mathbb{R}^n, a \in \text{GL}(\mathbb{R}^n)\}$, and the group operation is $(d, c)(b, a) = (c(b) + d, ca)$. This group is isomorphic to $\text{AGL}(\mathbb{R}^n)$.

**Holomorphs and Cayley’s theorem**  The transformation $x \mapsto ax + b$ is also a key to understanding Cayley’s theorem. But before looking for this representation, we will use the semidirect product to introduce another idea from group theory—the idea of a holomorph. Consider an arbitrary group $X$, which as usual is written multiplicatively. Let $A$ be a subgroup of $\text{Aut}(X)$. Notice that $\text{Aut}(X)$ is a subgroup of the symmetric group on $X$, denoted $\text{Sym}(X)$. Recall that the symmetric group on $X$ is the group of all bijections of $X$; that is, the set of permutations of $X$. For any $a \in A$ and $b \in X$, let $\Phi_{a,b}$ be the bijection of $X$ under which the value of any $x \in X$ is obtained by multiplying $a(x)$ from the left by $b$, that is, $\Phi_{a,b}(x) = ba(x)$; note that if $X$ is an additive abelian group then this reduces to the now very familiar $\Phi_{a,b}(x) = a(x) + b$.

We define the (relative) holomorph $H(X, A)$ of $(X, A)$ to be the subgroup of $\text{Sym}(X)$ consisting of all $\Phi_{a,b}$ with $a$ varying in $A$ and $b$ varying in $X$.

Consider the relative holomorph $H(X, 1)$. This is the image in $\text{Sym}(X)$ of the mapping $b \mapsto \Phi_{1,b}, b \in X$. Notice that $\Phi_{1,b}$ is simply the mapping of $X$ into $\text{Sym}(X)$ given by left multiplication by $b$. This reminds us of Cayley’s theorem.

In 1854, Cayley [4] stated his theorem that every group of order $n$ is isomorphic to a subgroup of $S_n$: “A set of symbols, 1, $\alpha$, $\beta$, $\ldots$ all of them different, and such that the product of any two of them into itself, belongs to the set, is said to be a group. It follows that if the entire group is multiplied by any one of the symbols, either as further or nearer factor, the effect is simply to reproduce the group . . . .”
$H(X, 1)$, obtained from the proof of Cayley’s theorem, is called the left regular representation of $X$ in $\text{Sym}(X)$. In the case of the additive group $\mathbb{R}^+$, the set of translations $\{\Phi_{1,b} = \{x \mapsto x + b\}$ is the regular representation of $\mathbb{R}^+$.

Notice that the left regular representation of $X$, $H(X, 1)$, is a normal subgroup of $H(X, A)$. (Usually the left regular representation is not a normal subgroup of Sym($X$), but it is a normal subgroup of $H(X, A)$ which is a subgroup of Sym($X$).) Moreover, for any subgroup $A$ of Aut($X$), the relative holomorph $H(X, A)$ is the internal semidirect product of the left regular representation $H(X, 1)$ by $A$.

We define the (full) holomorph $H(X)$ of $X$ by putting $H(X) = H(X, \text{Aut}(X))$. The holomorph of $X$ essentially combines $X$ and its group of automorphisms Aut($X$) to form a larger group.

**Affine general linear group** By using the idea of holomorph, we can define the affine general linear group of an arbitrary finite dimensional vector space $V$. Recall that the group of all nonsingular linear transformations of a finite dimensional vector space $V$ over a field $k$ is called the general linear group of $V$ and is denoted by GL($V$). We define the affine general linear group $\text{AGL}(V)$ of $V$ by putting $\text{AGL}(V) = H(V, \text{GL}(V))$. In this way, we define $\text{AGL}(V)$ as an internal semidirect product. Also, because GL($V$) is a subgroup of Aut($V$), we may think of $\text{AGL}(V)$ as the external semidirect product $V \rtimes \text{GL}(V)$. Of course, the internal semidirect product and the external semidirect product are isomorphic.

**A Theorem of Burnside** Abhyankar [1] begins with $x \mapsto ax + b$ and proves a theorem of Burnside that a 2-transitive permutation group has a unique minimal normal subgroup and that subgroup is either elementary abelian or nonabelian simple. This theorem, which first appeared in section 134 of the 1897 edition of Burnside’s book [3] and later in section 154 of the second edition (see also theorem 4.1B and theorem 7.2E of Dixon and Mortimer [5]), suffices to settle a portion of Hilbert’s Thirteenth Problem and is a cornerstone of the Classification of Doubly Transitive groups (see section 7.7 of Dixon and Mortimer [5]). The Classification of Doubly Transitive Groups was the first amazing application of the Classification Theorem of Finite Simple Groups, which Gorenstein called the Enormous Theorem.

A first step to all of these results can be taken by carefully exploring the familiar transformation $x \mapsto ax + b$.

**More** Of course, more information about the ideas that were mentioned above can be found in standard graduate texts on abstract algebra (Dummit and Foote [6] and Hungerford [11]) or in advanced texts on the theory of groups [2, 9, 10, 13]. A long section containing many examples and exercises about semidirect products can be found in Weinstein [15]. Rotman [13] discusses the affine general linear group and holomorphs. Holomorphs also were discussed by Burnside [3].

**Acknowledgment.** Abhyankar’s work was partly supported by NSF Grant DMS 97-32592 and NSA grant MDA 904-97-1-0010.

**REFERENCES**

There Are Three Methods for Solving Problems

**Method 0**
Get the answer from the teacher.

**Method 1**
Oh no! A problem! Oh no!
I can’t do it! I can’t!
Oh no! Oh no!

How will I solve this problem?
How will I get the right answer? Oh no!

There must be a method.
I’ll find the right one!

How will I find the right method?
What if I find the wrong one!
Oh no!

I might find the right method,
maybe, with luck.

But what if I make a mistake,
even just one?
Oh no! Oh no!

I’ll try and I’ll try.
I’ll think and I’ll thunk.

Here’s an answer.
It might be the wrong one.
Oh no! Oh no!

Teacher, teacher.
Is my answer correct?

**Method 2**
That’s an interesting puzzle.
Please show me that piece, it looks nice.

That piece over there looks nice too.
I wonder if these two pieces connect.

I’ll put this one here and that one nearby.
Let’s see where this one goes next.

I could try it right there,
Perhaps it goes here.

This spot is just right.
Now my puzzle is solved.

And I have my answer, too.
That’s nice.

My solution is correct.
And my answer’s just right!

Now I have three methods, I do.
Which one should I use just now?
I wonder, I do.

Could there be any more methods?
I wonder that too.

—CAROL LE GUENNEC
SOLANO COMMUNITY COLLEGE
SUISUN, CA 94585