SPLITTING OF SHORT EXACT SEQUENCES FOR MODULES

KEITH CONRAD

1. Introduction

Let R be a commutative ring. A sequence of R-modules and R-linear maps

$$N \xrightarrow{f} M \xrightarrow{g} P$$

is called *exact* at M if $\operatorname{im} f = \ker g$. For example, to say $0 \longrightarrow M \xrightarrow{h} P$ is exact at M means h is injective, and to say $N \xrightarrow{h} M \longrightarrow 0$ is exact at M means h is surjective. The linear maps coming out of 0 or going to 0 are unique, so there is no need to label them.

A *short exact sequence* of R-modules is a sequence of R-modules and R-linear maps

$$0 \longrightarrow N \xrightarrow{f} M \xrightarrow{g} P \longrightarrow 0 \tag{1.1}$$

which is exact at N, M, and P. That means f is injective, g is surjective, and $\operatorname{im} f = \ker g$.

Example 1.1. For an R-module M and submodule N, there is a short exact sequence

$$0 \longrightarrow N \longrightarrow M \longrightarrow M/N \longrightarrow 0,$$

where the map $N \to M$ is the inclusion and the map $M \to M/N$ is reduction modulo N.

Example 1.2. For R-modules N and P, the direct sum $N \oplus P$ fits into the short exact sequence

$$0 \longrightarrow N \longrightarrow N \oplus P \longrightarrow P \longrightarrow 0,$$

where the map $N \to N \oplus P$ is the embedding $n \mapsto (n, 0)$ and the map $N \oplus P \to P$ is the projection $(n, p) \mapsto p$.

Example 1.3. Let I and J be ideals in R such that $I + J = R$. Then there is a short exact sequence

$$0 \longrightarrow I \cap J \longrightarrow I \oplus J \xrightarrow{+} R \longrightarrow 0,$$

where the map $I \oplus J \to R$ is addition, whose kernel is $\{(x, -x) : x \in I \cap J\}$, and the map $I \cap J \to I \oplus J$ is $x \mapsto (x, -x)$. This is *not* the short exact sequence $0 \longrightarrow I \longrightarrow I \oplus J \longrightarrow J \longrightarrow 0$ as in Example 1.2, even though the middle modules in both are $I \oplus J$.

Any short exact sequence that looks like the short exact sequence of a direct sum in Example 1.2 is called a *split* short exact sequence. More precisely, a short exact sequence $0 \longrightarrow N \xrightarrow{f} M \xrightarrow{g} P \longrightarrow 0$ is called split when there is an R-module isomorphism $\theta\colon M \to N \oplus P$ such that the diagram

$$\begin{tikzcd}
0 \arrow[r] & N \arrow[r, "f"] \arrow[d, "\mathrm{id}"'] & M \arrow[r, "g"] \arrow[d, "\theta"'] & P \arrow[r] \arrow[d, "\mathrm{id}"'] & 0 \\
0 \arrow[r] & N \arrow[r] & N \oplus P \arrow[r] & P \arrow[r] & 0
\end{tikzcd} \tag{1.2}$$

commutes. The point is not simply that M is isomorphic to $N \oplus P$, but *how* the isomorphism works. It allows us to regard f as the embedding $N \to N \oplus P$ and g as the projection $N \oplus P \to P$. (Notice also that the outer vertical maps in (1.2) are both the identities.)

In Section 2 we will give two ways to characterize when a short exact sequence of R-modules splits. Section 3 will discuss a few consequences.

2. When a Short Exact Sequence Splits

Theorem 2.1. *Let* $0 \longrightarrow N \xrightarrow{f} M \xrightarrow{g} P \longrightarrow 0$ *be a short exact sequence of R-modules. The following are equivalent:*

1. *There is an R-linear map $f' \colon M \to N$ such that $f'(f(n)) = n$ for all $n \in N$.*
2. *There is an R-linear map $g' \colon P \to M$ such that $g(g'(p)) = p$ for all $p \in P$.*
3. *The short exact sequence splits: there is an isomorphism $\theta \colon M \to N \oplus P$ such that the diagram (1.2) commutes.*

If we replace R-modules with groups and R-linear maps with group homomorphisms, conditions (1) and (2) are not equivalent: for a short exact sequence $1 \longrightarrow H \xrightarrow{f} G \xrightarrow{g} H \longrightarrow 1$, (1) corresponds to G being a direct product of H and K while (2) corresponds to G being a semidirect product of H and K. The reason (1) and (2) are no longer equivalent for groups is related to noncommutativity. For an exact sequence of abelian groups, (1) and (2) are equivalent (the special case $R = \mathbf{Z}$, since abelian groups are $\mathbf{Z}$-modules).

Proof. We will first show (1) and (3) are equivalent, and then (2) and (3) are equivalent.

(1) $\Rightarrow$ (3): Define $\theta \colon M \to N \oplus P$ by

$$\theta(m) = (f'(m), g(m)).$$

Since f' and g are R-linear, θ is R-linear.

To see that the diagram (1.2) commutes, going around the top and right of the first square has the effect $n \mapsto f(n) \mapsto \theta(f(n)) = (f'(f(n)), g(f(n))) = (n, 0)$ and going around the left and bottom has the effect $n \mapsto n \mapsto (n, 0)$. Going both ways around the second square sends $m \in M$ to $g(m) \in P$.

To see θ is injective, suppose $\theta(m) = (0, 0)$, so $f'(m) = 0$ and $g(m) = 0$. From exactness at M, the condition $g(m) = 0$ implies $m = f(n)$ for some $n \in N$. Then $0 = f'(m) = f'(f(n)) = n$, so $m = f(n) = f(0) = 0$.

To show θ is surjective, let $(n, p) \in N \oplus P$. Since g is onto, $p = g(m)$ for some $m \in M$, so $p = g(m) = g(m + f(x))$ for any $x \in N$. To have $\theta(m + f(x)) = (n, p)$, we seek an $x \in N$ such that

$$n = f'(m + f(x)) = f'(m) + f'(f(x)) = f'(m) + x.$$

So *define* $x := n - f'(m)$. Then $m + f(x) = m + f(n) - f(f'(m))$ and

$$\begin{aligned} \theta(m + f(x)) &= (f'(m + f(x)), g(m + f(x))) \\ &= (n, g(m)) \\ &= (n, p). \end{aligned}$$

Thus θ is an isomorphism of R-modules.

(3) $\Rightarrow$ (1): Suppose there is an R-module isomorphism $\theta \colon M \to N \oplus P$ making (1.2) commute. From commutativity of the second square in (1.2), $\theta(m) = (*, g(m))$. Let the first coordinate of $\theta(m)$ be $f'(m)$: $\theta(m) = (f'(m), g(m))$. Then $f' \colon M \to N$. Since θ is

R-linear, f' is R-linear. By commutativity in the first square of (1.2), $\theta(f(n)) = (n, 0)$ for $n \in N$, so $(f'(f(n)), g(f(n))) = (n, 0)$, so $f'(f(n)) = n$ for all $n \in N$.

(2) $\Rightarrow$ (3): To get an isomorphism $M \to N \oplus P$, it is easier to go the other way. Let $h\colon N \oplus P \to M$ by

$$h(n, p) = f(n) + g'(p).$$

This is R-linear since f and g' are R-linear.

To show h is injective, if $h(n, p) = 0$ then $f(n) + g'(p) = 0$. Applying g to both sides, $g(f(n)) + g(g'(p)) = 0$, which simplifies to $p = 0$. Then $0 = f(n) + g'(0) = f(n)$, so $n = 0$ since f is injective.

To show h is surjective, pick $m \in M$. We want to find $n \in N$ and $p \in P$ such that

$$f(n) + g'(p) = m.$$

Applying g to both sides, we get

$$g(f(n)) + g(g'(p)) = g(m) \Rightarrow p = g(m).$$

So we define $p := g(m)$ and then ask if there is $n \in N$ such that $f(n) = m - g'(g(m))$. Since $\operatorname{im} f = \ker g$, whether or not there is such an n is equivalent to checking $m - g'(g(m)) \in \ker g$:

$$\begin{aligned} g(m - g'(g(m))) &= g(m) - g(g'(g(m)) \\ &= g(m) - g(m) \\ &= 0. \end{aligned}$$

Thus $h\colon N \oplus P \to M$ is an isomorphism of R-modules. Let $\theta = h^{-1}$ be the inverse isomorphism.

To show the diagram (1.2) commutes, it is equivalent to show the "flipped" diagram

$$\begin{tikzcd} 0 \arrow[r] & N \arrow[r, "f"] & M \arrow[r, "g"] & P \arrow[r] & 0 \\ 0 \arrow[r] & N \arrow[r] \arrow[u, "\mathrm{id}"'] & N \oplus P \arrow[r] \arrow[u, "h"'] & P \arrow[r] \arrow[u, "\mathrm{id}"'] & 0 \end{tikzcd}$$

commutes ($h = \theta^{-1}$)). For $n \in N$, going around the first square along the left and top has the effect $n \mapsto n \mapsto f(n)$, and going around the other way has the effect $n \mapsto (n, 0) \mapsto h(n, 0) = f(n) + g'(0) = f(n)$. In the second square, for $(n, p) \in N \oplus P$ going around the left and top has the effect $(n, p) \mapsto g(h(n, p)) = g(f(n)) + g(g'(p)) = 0 + p = p$, while going around the other way has the effect $(n, p) \mapsto p \mapsto p$.

(3) $\Rightarrow$ (2): Let $g'\colon P \to M$ by $g'(p) = \theta^{-1}(0, p)$. Since $p \mapsto (0, p)$ and θ^{-1} are R-linear, g' is R-linear. For $p \in P$, the commutativity of the diagram

$$\begin{tikzcd} M \arrow[r, "g"] \arrow[d, "\theta"'] & P \arrow[d, "\mathrm{id}"] \\ N \oplus P \arrow[r] & P \end{tikzcd}$$

implies commutativity of the diagram

$$\begin{tikzcd} M \arrow[r, "g"] & P \\ N \oplus P \arrow[r] \arrow[u, "\theta^{-1}"] & P \arrow[u, "\mathrm{id}"'] \end{tikzcd}$$

so $g(g'(p)) = g(\theta^{-1}(0,p)) = p$. $\square$

3. Consequences

Let's take another look at the short exact sequence in Example 1.3:

$$0 \longrightarrow I \cap J \longrightarrow I \oplus J \xrightarrow{+} R \longrightarrow 0, \tag{3.1}$$

where I and J are ideals with $I + J = R$ and the map from $I \cap J$ to $I \oplus J$ is $x \mapsto (x, -x)$. It turns out this splits: $I \oplus J$ is isomorphic to $(I \cap J) \oplus R$ in a manner compatible with the maps in the short exact sequence. That is, the diagram

$$\begin{tikzcd} 0 \arrow[r] & I \cap J \arrow[r] \arrow[d, "\mathrm{id}"'] & I \oplus J \arrow[r, "+"] \arrow[d, "\theta"'] & R \arrow[r] \arrow[d, "\mathrm{id}"'] & 0 \\ 0 \arrow[r] & I \cap J \arrow[r] & (I \cap J) \oplus R \arrow[r] & R \arrow[r] & 0 \end{tikzcd}$$

commutes for some isomorphism θ. The bottom row is the usual short exact sequence for a direct sum of R-modules. To show the sequence (3.1) splits, we use the equivalence of (2) and (3) in Theorem 2.1. From $I + J = R$ we have $x_0 + y_0 = 1$ for some $x_0 \in I$ and $y_0 \in J$. Let $g' \colon R \to I \oplus J$ by $g'(r) = (rx_0, ry_0)$. Then $rx_0 + ry_0 = r$, so g' is a right inverse to the addition map $I \oplus J \to R$ and that shows (3.1) splits.

Although $I \oplus J \cong (I \cap J) \oplus R$ as R-modules, it need not be the case that either I or J is isomorphic to $I \cap J$ or R.

Example 3.1. Let $R = \mathbf{Z}[\sqrt{-5}]$, $I = (3, 1 + \sqrt{-5})$, and $J = (3, 1 - \sqrt{-5})$. Then $I + J$ contains 3 and $1+\sqrt{-5}+1-\sqrt{-5} = 2$, so it contains 1 and thus $I+J = R$. From $I+J = R$, $I \cap J = IJ$ and $IJ = 3R \cong R$. Therefore

$$I \oplus J \cong R \oplus R$$

as R-modules. The ideals I and J are not isomorphic to R as R-modules since they are nonprincipal ideals: $I^2 = (2 - \sqrt{-5})$, $J^2 = (2 + \sqrt{-5})$, and neither $\pm(2 + \sqrt{-5})$ nor $\pm(2 - \sqrt{-5})$ are squares in $\mathbf{Z}[\sqrt{-5}]$.

Using Theorem 2.1, we can describe when a submodule $N \subset M$ is a direct summand.

Theorem 3.2. *For a submodule $N \subset M$, the following conditions are equivalent:*

1. *N is a direct summand: $M = N \oplus P$ for some submodule $P \subset M$.*
2. *There is an R-linear map $f' \colon M \to N$ such that $f'(n) = n$ for all $n \in N$.*

Proof. (1) $\Rightarrow$ (2): Let $f' \colon M \to N$ by $f'(n + p) = n$. This is well-defined from the meaning of a direct sum decomposition, and it is R-linear. Obviously $f'(n) = n$ for $n \in N$.

(2) $\Rightarrow$ (1): There is a standard short exact sequence

$$0 \longrightarrow N \longrightarrow M \longrightarrow M/N \longrightarrow 0.$$

Since f' is a left inverse to the inclusion map $N \to M$ in this short exact sequence, the equivalence of (1) and (3) in Theorem 2.1 implies there is a commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & N \arrow[r] \arrow[d, "\mathrm{id}"'] & M \arrow[r] \arrow[d, "\theta"'] & M/N \arrow[r] \arrow[d, "\mathrm{id}"'] & 0 \\ 0 \arrow[r] & N \arrow[r] & N \oplus (M/N) \arrow[r] & M/N \arrow[r] & 0 \end{tikzcd}$$

where $\theta(m) = (f'(m), \overline{m})$ is an R-module isomorphism. For $n \in N$, $\theta(n) = (f'(n), \overline{n}) = (n, 0)$, so using θ^{-1} shows M has a direct sum decomposition with N as the first summand. $\square$

Theorem 3.3. *For an injective R-linear map $N \xrightarrow{f} M$, the following conditions are equivalent:*

1. *$f(N)$ is a direct summand of M.*
2. *There is an R-linear map $f' \colon M \to N$ such that $f'(f(n)) = n$ for all $n \in N$.*

The proof is similar to that of Theorem 3.2.