UNIVERSITY OF CALIFORNIA
Santa Barbara

Monoids and Categories of Noetherian Modules

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy
in
Mathematics
by
Gary John Brookfield

Committee in Charge:
Professor K. R. Goodearl, Chairperson
Professor B. Huisgen-Zimmermann
Professor J. Zelmanowitz

June 1997

The Dissertation of Gary John Brookfield is approved

Committee Chairperson

June 1997

VITA

August 3, 1954 — Born — Southampton, UK

1978 — B.Sc. Physics, University of British Columbia

1981 — M.Sc. Physics, University of British Columbia

PUBLICATIONS

Direct Sum Cancellation of Noetherian Modules,
accepted by the Journal of Algebra, March 1997

FIELDS OF STUDY

Major Field. Mathematics

Specialties: Ring Theory, Noetherian Modules

Professor K. R. Goodearl

Abstract

Monoids and Categories of Noetherian Modules

by

Gary John Brookfield

In this dissertation we will investigate the structure of module categories by considering how modules can be constructed by extensions from other modules. The natural way to do this is to define for a module category $\mathbf{S}$, a commutative monoid $(M(\mathbf{S}), +)$ and a map $\Lambda\colon \mathbf{S} \to M(\mathbf{S})$ so that the relevant information about $\mathbf{S}$ is transferred to $M(\mathbf{S})$ via Λ. Specifically, we construct $M(\mathbf{S})$ so that if $A, B, C \in \mathbf{S}$ and there is a short exact sequence $0 \to A \to B \to C \to 0$, then $\Lambda(B) = \Lambda(A) + \Lambda(C)$ in $M(\mathbf{S})$. We also require that $M(\mathbf{S})$ be the largest monoid with this property, meaning that $M(\mathbf{S})$ has the universal property for such maps.

Let R be a ring, R-**Mod** the category of left R-modules and R-**Noeth** the category of Noetherian left R-modules. Then we will prove

- $M(R\text{-}\mathbf{Mod})$ is a conical commutative refinement monoid.
- $M(R\text{-}\mathbf{Noeth})$ is semi-Artinian and strongly separative, meaning that
$$2a = a + b \implies a = b$$
for all $a, b \in M(R\text{-}\mathbf{Noeth})$.
- The Krull dimension of a Noetherian module can be determined from its image in $M(R\text{-}\mathbf{Noeth})$.
- If R is left fully bounded Noetherian, in particular, if R is commutative Noetherian, then $M(R\text{-}\mathbf{Noeth})$ is Artinian, primely generated and weakly cancellative, and has $\leq$-multiplicative cancellation.

The strong separativity of $M(R\text{-}\mathbf{Noeth})$ leads to one of the most interesting results of this dissertation: If A, B, C are left R-modules such that $A \oplus C \cong B \oplus C$ with C Noetherian, then A and B have isomorphic submodule series.

Contents

1 Introduction

It is the hope of all module theorists that modules can be understood and classified by considering how they can be built from a family of simpler building blocks. The prototype for this idea is the classification of Noetherian $\mathbb{Z}$-modules, that is, finitely generated Abelian groups: Every such group is a finite direct sum of the groups $\mathbb{Z}$ and $\mathbb{Z}_{p^n}$ for $p, n \in \mathbb{N}$ and p prime. Thus the building blocks for finitely generated Abelian groups are the groups $\mathbb{Z}$ and $\mathbb{Z}_{p^n}$, and the glue that holds them together is the direct sum operation. Moreover, two finitely generated Abelian groups are isomorphic if and only if they are built from the same building blocks. This is the best possible situation and reflects a complete understanding of Noetherian $\mathbb{Z}$-modules.

This happy situation, when it occurs, has many simple consequences. Some of these which will be important in discussing other module categories, we list here:

- *Cancellation:* If A, B and C are Noetherian $\mathbb{Z}$-modules such that $A \oplus C \cong B \oplus C$, then $A \cong B$.
- *Multiplicative Cancellation:* If A and B are Noetherian $\mathbb{Z}$-modules such that $A^n \cong B^n$ for some $n \in \mathbb{N}$, then $A \cong B$.
- *Refinement:* If A_1, A_2, B_1 and B_2 are Noetherian $\mathbb{Z}$-modules such that $A_1 \oplus A_2 \cong B_1 \oplus B_2$, then there are Noetherian $\mathbb{Z}$-modules C_{11}, C_{12}, C_{21} and C_{22} such that

$$\begin{aligned} A_1 &\cong C_{11} \oplus C_{12} & A_2 &\cong C_{21} \oplus C_{22} \\ B_1 &\cong C_{11} \oplus C_{21} & B_2 &\cong C_{12} \oplus C_{22}. \end{aligned}$$

If we write $A \lesssim B$ when A is isomorphic to a direct summand of B, then we have

- *Antisymmetry:* If A and B are Noetherian $\mathbb{Z}$-modules such that $A \lesssim B \lesssim A$, then $A \cong B$.
- $\lesssim$*-Cancellation:* If A, B and C are Noetherian $\mathbb{Z}$-modules such that $A \oplus C \lesssim B \oplus C$, then $A \lesssim B$.
- $\lesssim$*-Multiplicative Cancellation:* If A and B are Noetherian $\mathbb{Z}$-modules such that $A^n \lesssim B^n$ for some $n \in \mathbb{N}$, then $A \lesssim B$.
- *Decomposition:* If A, B_1 and B_2 are Noetherian $\mathbb{Z}$-modules such that $A \lesssim B_1 \oplus B_2$, then there are Noetherian $\mathbb{Z}$-modules C_1, C_2 such that $C_1 \lesssim B_1$, $C_2 \lesssim B_2$ and $A \cong C_1 \oplus C_2$.
- *Descending Chain Condition:* If $A_1 \gtrsim A_2 \gtrsim A_3 \gtrsim \ldots$ is a decreasing sequence of Noetherian $\mathbb{Z}$-modules then there is an $N \in \mathbb{N}$ such that $A_n \cong A_N$ for all $n \geq N$.

All of the above properties derive from the classification theorem for $\mathbb{Z}$-**Noeth**. (The notation used in this dissertation is explained at the end of this section.) The reason for their importance is that, for other rings and in other circumstances, they occur even without a complete classification theorem, and hence can be considered consolation prizes when a complete classification is not possible. Indeed, almost no category of modules can be classified the simple way that $\mathbb{Z}$-**Noeth** is. Even though the full classification may not be possible, there is still some hope that some of these other weaker properties may occur.

Before considering more general module categories, we note the special properties that the building blocks of $\mathbb{Z}$-**Noeth** have: A nonzero Noetherian $\mathbb{Z}$-module P is one of the building blocks described above if it has the following property:

- *Primeness:* If A and B are Noetherian $\mathbb{Z}$-modules such that $P \lesssim A \oplus B$, then either $P \lesssim A$ or $P \lesssim B$.

It will be very useful to add an element of abstraction to the classification of Noetherian $\mathbb{Z}$-modules which will enable us to discuss the successes and failures of similar classification schemes for other module categories:

Definition 1.1. *Let R be a unital ring and $\mathbf{S} \subseteq R$-**Mod** a class of left R-modules which is closed under finite direct sums and contains the zero module. For any module $A \in \mathbf{S}$, let $\{\cong A\}$ be its isomorphism class. Let $V(\mathbf{S})$ be class of all isomorphism classes of $\mathbf{S}$. We will write $+$ for the operation on $V(\mathbf{S})$ induced by the direct sum, that is,*

$$\{\cong A\} + \{\cong B\} = \{\cong A \oplus B\}$$

for all $A, B \in \mathbf{S}$.

It is easy to see that $V(\mathbf{S})$ is a well defined commutative monoid (though perhaps not a set; see 5.1). The identity element of $V(\mathbf{S})$ is the image of the zero module, $\{\cong 0\} = \{0\}$, which we will write as $0 \in V(\mathbf{S})$.

The classification of Noetherian $\mathbb{Z}$-modules discussed above is then a statement about $V(\mathbb{Z}\text{-}\mathbf{Noeth})$:

- $V(\mathbb{Z}\text{-}\mathbf{Noeth})$ is a free commutative monoid with basis $\{\cong \mathbb{Z}\}$ and $\{\cong \mathbb{Z}_{p^n}\}$ for $p, n \in \mathbb{N}$ and p prime (5.13).

Corresponding to the properties of $\mathbb{Z}$-**Noeth** described above are the following monoid properties. These we define for an arbitrary commutative monoid M:

P1: *Cancellation:* If $a, b, c \in M$ such that $a + c = b + c$, then $a = b$.

P2: *Multiplicative Cancellation:* If $a, b \in M$ such that $na = nb$ for some $n \in \mathbb{N}$, then $a = b$.

P3: *Refinement:* If $a_1, a_2, b_1, b_2 \in M$ such that $a_1 + a_2 = b_1 + b_2$, then there are $c_{11}, c_{12}, c_{21}, c_{22} \in M$ such that

$$\begin{aligned} a_1 &= c_{11} + c_{12} & \qquad a_2 &= c_{21} + c_{22} \\ b_1 &= c_{11} + c_{21} & \qquad b_2 &= c_{12} + c_{22}. \end{aligned}$$

For $A, B \in \mathbb{Z}$-**Noeth** we have

$$(A \lesssim B) \iff (\exists C \text{ such that } A \oplus C \cong B) \iff (\exists C \text{ such that } \{\cong A\} + \{\cong C\} = \{\cong B\}).$$

Accordingly, in an arbitrary commutative monoid M, we define a relation $\leq$ by

$$a \leq b \iff \exists c \in M \text{ such that } a + c = b$$

for $a, b \in M$ (6.1). Thus, $A \lesssim B$ in $\mathbb{Z}$-**Noeth** if and only if $\{\cong A\} \leq \{\cong B\}$ in $V(\mathbb{Z}\text{-}\mathbf{Noeth})$. Corresponding to the remaining properties of $\mathbb{Z}$-**Noeth** we then get the following monoid properties:

P4: *Antisymmetry:* If $a, b \in M$ such that $a \leq b \leq a$, then $a = b$.

P5: $\leq$*-Cancellation:* If $a, b, c \in M$ such that $a + c \leq b + c$, then $a \leq b$.

P6: $\leq$*-Multiplicative Cancellation:* If $a, b \in M$ such that $na \leq nb$ for some $n \in \mathbb{N}$, then $a \leq b$.

P7: *Decomposition:* If $a, b_1, b_2 \in M$ such that $a \leq b_1 + b_2$, then there are $c_1, c_2 \in M$ such that $c_1 \leq b_1$, $c_2 \leq b_2$ and $a = c_1 + c_2$.

P8: *Descending Chain Condition:* If $a_1 \geq a_2 \geq a_3 \geq \ldots$ is a decreasing sequence in M, then there is an $N \in \mathbb{N}$ such that $a_n \geq a_N$ for all $n \geq N$. Note that we require $a_n \geq a_N$ rather than $a_n = a_N$. See Section 2 for details.

The basis elements of $V(\mathbb{Z}$-**Noeth**$)$ can be characterized within $M = V(\mathbb{Z}$-**Noeth**$)$ as those elements $0 \neq q \in M$ with the following property:

- *Primeness:* If $a, b \in M$ such that $q \leq a + b$, then either $q \leq a$ or $q \leq b$.

Elements of a commutative monoid which satisfy this condition are called **prime** elements.

Any free commutative monoid, in particular $V(\mathbb{Z}$-**Noeth**$)$, has all of the properties P1-P8. In arbitrary commutative monoids these properties are not independent of each other. For example, if a commutative monoid has refinement then it has decomposition. One of the main purposes of this dissertation is to prove other interdependencies. For example, we will show that if a commutative monoid has refinement and the descending chain condition then it also has $\leq$-multiplicative cancellation (13.1). These same conditions do not suffice to make the monoid cancellative.

With this new level of abstraction we can discuss structure theorems for module categories of the following form: Let $\mathbf{S}$ be a class of modules and $\Lambda\colon \mathbf{S} \to M$ a surjective map where M is a commutative monoid with monoid operation $+$ induced from some module composition operation in $\mathbf{S}$. Then a structure theorem for $\mathbf{S}$ is simply some statement, with P1-P8 serving as prototypes, about the structure of M. The significance and utility of such a theorem depends on these criteria:

C1: How big is $\mathbf{S}$? The bigger, the better.

C2: How small are the equivalence classes $\Lambda^{-1}(a)$ for $a \in M$? The ideal situation is that these equivalence classes are isomorphism classes of modules, so that knowing $\Lambda(A) = \Lambda(B)$ in M implies $A \cong B$ as modules. Such a map provides the most detailed information on $\mathbf{S}$.

C3: Which of the properties P1-P8 does the monoid M have? The more it has, the closer the theorem is to the ideal represented by the classification of $\mathbb{Z}$-**Noeth**.

C4: Is there a small class of prime elements of M such that every element of M is a finite sum of these elements? If so, are elements of M uniquely expressible as sums of these elements? Is M isomorphic to the free commutative monoid generated by these elements?

As is usual, any structure theorem involves making a compromise of these conflicting goals. It is the main theme of this dissertation that gains can be made in items C1, C3 and C4 of this list by sacrificing the ideal situation in C2. Before explaining this, we will show by example that even for a commutative Noetherian ring R and $\mathbf{S} = R$-**Noeth**, maintaining the ideal situation in C2, requires compromises in C3. Specifically, we will show that $V(R$-**Noeth**$)$ does not, in general, have refinement, decomposition, cancellation, multiplicative cancellation, or $\leq$-multiplicative cancellation.

Example 1.2. *Let $S = \mathbb{R}[X,Y]/(X^2+Y^2-1)$, the coordinate ring of the unit circle. We will write x, y for the images of X, Y in S. Let $\Phi\colon S \to S$ be the ring automorphism such that $\Phi(x) = -x$, $\Phi(y) = -y$ and $\Phi(r) = r$ for all $r \in \mathbb{R}$. Let $R = \{r \in S \mid \Phi(r) = r\}$. One readily confirms that R is the subring of S consisting of all polynomials such that the total degree of each term is even. Considered as functions on the unit circle, R contains all polynomials $f(x,y)$ such that $f(x,y) = f(-x,-y)$.*

Let $A = \{a \in S \mid \Phi(a) = -a\}$. Considered as functions on the unit circle, A contains all polynomials $f(x,y)$ such that $f(x,y) = -f(-x,-y)$. A is not a subring of S, but it is an R-submodule, since if $r \in R$ and $a \in A$, then $\Phi(ra) = \Phi(r)\Phi(a) = -ra$, that is, $ra \in A$.

We will prove that $A \not\cong R$ as R-modules by showing that A is not cyclic: Suppose $f \in A$. Then it is a topological fact that there must be some point (x_0, y_0) on the unit circle such that $f(x_0, y_0) = 0$. If $x_0 \neq 0$, then the polynomial $x \in A$ is nonzero at (x_0, y_0), so is not in Rf. Similarly, if $y_0 \neq 0$, then the polynomial $y \in A$ is nonzero at (x_0, y_0), so is not in Rf. Since, of course, $x_0 = y_0 = 0$ is not possible, we have $Rf \neq A$. Thus A is not cyclic and can not be isomorphic to R.

Next we will show that $A \oplus A \cong R \oplus R$ as R-modules: Define $\Psi\colon S \oplus S \to S \oplus S$ by

$$\Psi(s_1, s_2) = (s_1 x + s_2 y, s_1 y - s_2 x)$$

for all $(s_1, s_2) \in S \oplus S$. It is easy to check that Ψ is an R-module homomorphism, that $\Psi \circ \Psi = \mathrm{id}_S$, and that $\Psi(R \oplus R) = A \oplus A$ and $\Psi(A \oplus A) = R \oplus R$. Thus $A \oplus A \cong R \oplus R$.

Let $r = \{\cong R\}$ and $a = \{\cong A\}$ in the monoid $V(R\text{-}\mathbf{Noeth})$. Then we have $2a = 2r$ but $a \neq r$. Thus $V(R\text{-}\mathbf{Noeth})$ does not have multiplicative cancellation.

Since $X^2 + Y^2 - 1$ is irreducible, S is a domain, and since R is a subring, R is also a domain. In particular, R has no nontrivial direct summands, so that $A \not\lesssim R$ and $a \not\leq r$. It is then easy to see that the inequality $a \leq r + r$ can not be decomposed, which means $V(R\text{-}\mathbf{Noeth})$ does not have decomposition or refinement. Also, we have $2a \leq 2r$ but $a \not\leq r$, so that $V(R\text{-}\mathbf{Noeth})$ does not have $\leq$-multiplicative cancellation.

We also want to show that cancellation can fail in $V(R\text{-}\mathbf{Noeth})$, for R a commutative Noetherian ring. One standard example of this, due to Kaplansky and Swan [29] is the following:

Example 1.3. *Let $R = \mathbb{R}[X,Y,Z]/(X^2+Y^2+Z^2-1)$, the coordinate ring of the unit sphere. We will write x, y, z for the images of X, Y, Z in R. Let $\eta\colon R \oplus R \oplus R \to R$ be the R-module homomorphism defined by $\eta(a,b,c) = ax + by + cz$. Since $\eta(x,y,z) = 1$, this homomorphism is surjective. Let $P = \ker \eta$, then we get the short exact sequence*

$$0 \to P \to R \oplus R \oplus R \xrightarrow{\eta} R \to 0.$$

Since R is projective, this sequence splits to give $(R \oplus R) \oplus R \cong P \oplus R$. In [29, Theorem 3] *and* [23, 11.2.3] *a topological argument, similar to the argument used in Example 1.2, is used to show that $P \not\cong R \oplus R$.*

Let $r = \{\cong R\}$ and $p = \{\cong P\}$ in the monoid $V(R\text{-}\mathbf{Noeth})$. Then we have $2r + r = p + r$ but $2r \neq p$. Thus $V(R\text{-}\mathbf{Noeth})$ is not cancellative.

For many more examples of this type see [21].

In particular, from these examples we see that for R a commutative Noetherian ring, $V(R\text{-}\mathbf{Noeth})$ can not be a free commutative monoid, and so there is little hope of understanding the structure of this monoid in terms of sums of building blocks as was done for

$V(\mathbb{Z}$-**Noeth**). Of course, these difficulties are not made any easier if we want to study noncommutative rings or non-Noetherian modules.

We should mention here a few other situations where, in spite of the above examples, one can say something about $V(\mathbf{S})$. The properties of $V(\mathbb{Z}$-**Noeth**) that we have been discussing occur for any PID, commutative or noncommutative. So if R is a PID, then $V(R$-**Noeth**) is a free commutative monoid with basis $\{\cong R\}$ and all elements $\{\cong A\}$ where A is an indecomposable Noetherian R-module. In this case $V(R$-**Noeth**) has all of the properties P1-P8.

If R is any ring and $\mathbf{S} = R$-**len**, the class of R-modules of finite length, then the Krull-Remak-Schmidt-Azumaya Theorem is equivalent to saying that $V(\mathbf{S})$ is a free commutative monoid with basis all elements $\{\cong A\}$ where A is an indecomposable finite length module. In more generality, we get the same result if $\mathbf{S}$ is the set of all R-modules which are finite direct sums of modules with local endomorphism rings.

It seems that, except for some simple cases, $V(R$-**Noeth**) is too complicated to get a handle on – the monoid $V(R$-**Noeth**) retains too much of the complexity of R-**Noeth** itself. One approach to this problem, the one taken in this dissertation, is to define a commutative monoid $M(\mathbf{S})$ for each class of modules $\mathbf{S}$ and a map $\Lambda\colon \mathbf{S} \to M(\mathbf{S})$ in which we abandon the ideal hoped for in C2, in exchange for gains in C1, C3 and C4. We do this by changing the glue that we use to hold modules together. Instead of the direct sum operation, we use extensions. Specifically, we will consider a module B to be built from the modules A and C if there is a short exact sequence

$$0 \to A \to B \to C \to 0$$

in R-**Mod**. Of course, in this circumstance, the module B is not determined, even up to isomorphism, by the modules A and C, and this is exactly where we make our compromise in C2.

The details of the construction of the monoid $M(\mathbf{S})$ are in Section 16 so we will just outline the main features here: Let $\mathbf{S}$ be a Serre subcategory (16.1) of R-**Mod**. The monoid $M(\mathbf{S})$ and the map $\Lambda\colon \mathbf{S} \to M(\mathbf{S})$ are constructed so that the monoid operation is induced from extensions in the same way that the operation in $V(\mathbf{S})$ is induced from direct sums. Thus we require that if $A, B, C \in \mathbf{S}$ and there is a short exact sequence $0 \to A \to B \to C \to 0$, then $\Lambda(B) = \Lambda(A) + \Lambda(C)$. Such maps are said to respect short exact sequences. We also require that the monoid $M(\mathbf{S})$ be universal for maps which respect short exact sequences from $\mathbf{S}$ to arbitrary monoids. Another way of saying this is that we require Λ and $M(\mathbf{S})$ to compromise C2 as little as possible given our choice of glue.

It turns out that it is possible to construct uniquely (up to monoid isomorphism) such a monoid. The image of a module A in $M(\mathbf{S})$ will be written as $[A]$. For any Serre subcategory $\mathbf{S} \subseteq R$-**Mod**, the monoid $M(\mathbf{S})$ is contained in $M(R$-**Mod**) so we do not need to distinguish $[A] \in M(\mathbf{S})$ from $[A] \in M(R$-**Mod**). The image of the zero module $[0]$ is the identity element of $M(\mathbf{S})$.

What do we lose by studying $M(\mathbf{S})$ rather than $V(\mathbf{S})$? Certainly $M(\mathbf{S})$ contains less information about $\mathbf{S}$ than $V(\mathbf{S})$. In regard to C2, we can be quite precise about when two modules map to the same element of $M(\mathbf{S})$: If $A, B \in \mathbf{S}$, then $[A] = [B]$ if and only if the modules have isomorphic submodules series. That is, there are submodule series

$0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ and $0 = B_0 \leq B_1 \leq \cdots \leq B_n = B$ and a permutation of the indices σ, such that $A_i/A_{i-1} \cong B_{\sigma(i)}/B_{\sigma(i)-1}$ for $i = 1, 2, \ldots, n$.

For a simple example, consider again $\mathbb{Z}$-**Noeth** and the modules $\mathbb{Z}$ and $\mathbb{Z} \oplus \mathbb{Z}_2$. These are, of course, non-isomorphic modules since they map to distinct elements of $V(\mathbb{Z}\text{-}\mathbf{Noeth})$. On the other hand, we have the standard short exact sequences

$$0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}_2 \to 0$$

and

$$0 \to \mathbb{Z} \to \mathbb{Z} \oplus \mathbb{Z}_2 \to \mathbb{Z}_2 \to 0$$

in $\mathbb{Z}$-**Noeth**, so in $M(\mathbb{Z}\text{-}\mathbf{Noeth})$ we get $[\mathbb{Z}] = [\mathbb{Z}] + [\mathbb{Z}_2] = [\mathbb{Z} \oplus \mathbb{Z}_2]$. In our analogy, each of the modules $\mathbb{Z}$ and $\mathbb{Z} \oplus \mathbb{Z}_2$ is built from $\mathbb{Z}$ and $\mathbb{Z}_2$, but they are glued together in different ways. The monoid $M(\mathbb{Z}\text{-}\mathbf{Noeth})$ records only that they are made from the same parts, and ignores the way they are constructed from these parts. This means that no property of $M(\mathbb{Z}\text{-}\mathbf{Noeth})$ can distinguish these two modules. For example, the fact that $\mathbb{Z}$ is a uniform module and $\mathbb{Z} \oplus \mathbb{Z}_2$ is not, can not be seen in the structure of $M(\mathbb{Z}\text{-}\mathbf{Noeth})$. On the plus side, both of these modules have Krull dimension 1, and this is a property that can be seen in the monoid, since the Krull dimension function on $\mathbb{Z}$-**Noeth** respects short exact sequences. See 17.4. This loss of information about $\mathbf{S}$ in going from $V(\mathbf{S})$ to $M(\mathbf{S})$ should not be considered a disadvantage if, in compensation, we have better information about $M(\mathbf{S})$ than about $V(\mathbf{S})$.

The above example shows another important feature of $M(\mathbb{Z}\text{-}\mathbf{Noeth})$, that it is not cancellative: We have $[\mathbb{Z}] = [\mathbb{Z}] + [\mathbb{Z}_2]$, but $[\mathbb{Z}_2] \neq [0] = 0$. The reason that cancellation fails here is that $[\mathbb{Z}]$ is too big compared with $[\mathbb{Z}_2]$. One of the main results of this dissertation (17.4) is that we do get cancellation in $M(R\text{-}\mathbf{Noeth})$ when the size of the canceled element is controlled relative to the remaining elements. Specifically, for any ring R, if $a, b, c \in M(R\text{-}\mathbf{Noeth})$ with $a + c = b + c$ and $c \leq a$, then $a = b$. This cancellation property is called strong separativity and is equivalent to

$$2a = a + b \implies a = b$$

for all elements a, b of a monoid. See 8.12.

We show that strong separativity occurs in Example 1.3:

Example 1.3 (continued)**.** *We have* $R \oplus R \oplus R \cong P \oplus R$ *with, of course,* R *and* P *Noetherian. In* $M(R\text{-}\mathbf{Noeth})$ *we get* $3[R] = [R] + [P]$ *and hence* $2(2[R]) = (2[R]) + [P]$. *Since* $M(R\text{-}\mathbf{Noeth})$ *is strongly separative, this implies* $2[R] = [P]$ *and* $[R \oplus R] = [P]$. *We will confirm that* $R \oplus R$ *and* P *have isomorphic submodule series even though* $R \oplus R \not\cong P\ldots$

It is easily checked that the homomorphism $\tau\colon R \oplus R \oplus R \to P$ *given by* $\tau(a, b, c) = (a, b, c) - \eta(a, b, c)(x, y, z)$ *is the projection from* $R \oplus R \oplus R$ *onto* P. *Thus* P *is generated by* $\tau(1, 0, 0)$, $\tau(0, 1, 0)$ *and* $\tau(0, 0, 1)$.

Note that $\tau(x, y, z) = 0$ *and, more generally,* $\tau(a, b, c) = 0$ *if and only if* (a, b, c) *is a multiple of* (x, y, z). *With this fact, a simple calculation shows that* τ *is monic when restricted to* $R \oplus R \oplus 0$, *so that the submodule* $Q = \tau(R \oplus R \oplus 0) = R\tau(1, 0, 0) + R\tau(0, 1, 0)$ *is isomorphic to* $R \oplus R$.

To investigate the quotient module P/Q *we define the homomorphism* $\gamma\colon R \to P/Q$ *by* $\gamma(c) = \tau(0, 0, c) + Q$. *This homomorphism is surjective by construction and a calculation shows that* $\ker\gamma = Rz$. *Thus* $P/Q \cong R/Rz$.

Since R is a domain, we also have $Rz \cong R$. Thus $0 \leq Q \leq P$ and $0 \leq R \oplus Rz \leq R \oplus R$ are isomorphic submodule series for P and $R \oplus R$ with factors isomorphic to $R \oplus R$ and R/Rz.

In contrast, the same example shows that if R is commutative Noetherian, $V(R\text{-}\mathbf{Noeth})$ may not be strongly separative: We have in Example 1.3 that $2r + r = p + r$, and hence $2(2r) = (2r) + p$, but $2r \neq p$.

The strong separativity of $M(R\text{-}\mathbf{Noeth})$ leads to one of the most interesting results of this dissertation (17.10): If A, B, C are modules such that $A \oplus C \cong B \oplus C$ with $C \in R\text{-}\mathbf{Noeth}$, then $[A] = [B]$ in $M(R\text{-}\mathbf{Mod})$.

The strong separativity of $M(R\text{-}\mathbf{Noeth})$ is a consequence of two other important properties of $M(R\text{-}\mathbf{Mod})$. The first of these is that for any ring R, the monoid $M(R\text{-}\mathbf{Mod})$ has refinement (16.10). This is itself a result of the Schreier refinement theorem for submodule series in modules. The second of these properties is that $M(R\text{-}\mathbf{Noeth})$ is a semi-Artinian monoid (14.1). This is a weak type of descending chain condition for monoids which will be discussed in Sections 14 and 13.

The situation for semi-Artinian refinement monoids seen in $M(R\text{-}\mathbf{Noeth})$ is an example of another of the main themes of this dissertation, that commutative refinement monoids which have descending chain conditions also have cancellation properties. As well as the semi-Artinian monoids already mentioned, we will define Artinian monoids which satisfy the descending chain condition of P9 (2.13). In Section 19, we will show that if R is a fully bounded Noetherian (FBN) ring, in particular, if R is commutative Noetherian, then $M(R\text{-}\mathbf{Noeth})$ is an Artinian monoid.

Artinian refinement monoids have the strongest descending chain condition and so have many cancellation properties. These include $\leq$-multiplicative cancellation, as well as the weak cancellation and midseparativity properties that are defined in 9.1. They also have the property that every element is a sum of prime elements. This is getting quite close to the ideal situation in C4, though there are some subtleties about the unique representation of elements by sums of primes.

In Section 19 we will show another interesting aspect of $M(R\text{-}\mathbf{Noeth})$ when R is a Dedekind domain, that the ideal class group of the ring is embedded in the monoid in a natural way. Thus one can consider the monoid $M(R\text{-}\mathbf{Noeth})$ as a generalization of the ideal class group in the sense that it contains similar information. Though the ideal class group is defined only for Dedekind domains, the monoid is defined for any ring.

Apart from this introduction, the dissertation falls naturally into three parts:

Part B. *Ordered Classes.* As mentioned above, the order structure of a commutative refinement monoid is crucial in understanding its cancellation properties. In this part of the dissertation we collect all the definitions and theorems about ordered classes which we will need to study monoids in Part C. In addition, we will investigate the Krull length function which will enable us to get a handle on the order structure of Noetherian and Artinian modules. This in turn will allow us to derive the order properties of the corresponding monoids $M(R\text{-}\mathbf{Noeth})$ and $M(R\text{-}\mathbf{Art})$ in Part D.

Part C. *Commutative Monoids.* The bulk of this dissertation is devoted to the investigation of the relation between the order and cancellation properties of commutative refinement monoids. Because of the absence of a suitable source in the semigroup literature, we will have to start at a very basic level with the discussion of various isomorphism theorems, free monoids, direct products and direct sums of monoids. The preorder on monoids and the refinement and decomposition properties are introduced in Sections 6 and 7. In Sections 8 and 9, we define the main cancellation properties: separativity, strong separativity, weak cancellation and midseparativity. Primely generated refinement monoids have a lot of cancellation properties. These we discuss in Section 11.

Finally, in the last four sections of Part C, we consider cancellation in commutative monoids with chain conditions. Artinian refinement monoids have the strongest chain condition and hence the most cancellation properties. Semi-Artinian refinement monoids have a weaker chain condition and a correspondingly smaller number of cancellation properties.

Part D. *Modules.* In the final part of the dissertation, we apply our understanding of order and cancellation in monoids to monoids constructed from categories of modules. The construction itself and its basic properties are discussed in Section 16. One of the main results of this dissertation is that $M(R\text{-}\mathbf{Noeth})$ is a semi-Artinian refinement monoid with no proper regular elements, and so is strongly separative. This theorem and its consequences are in Section 17.

In the final section we discuss the case where R is an FBN or commutative Noetherian ring. We will show that, in this circumstance, $M(R\text{-}\mathbf{Noeth})$ is an Artinian refinement monoid and has as a consequence strong cancellation properties. In fact, considerable information about $M(R\text{-}\mathbf{Noeth})$ can be obtained because we know what the prime elements look like. This is used in the last theorem of the dissertation which shows that for a Dedekind domain R, the ideal class group is embedded in $M(R\text{-}\mathbf{Noeth})$.

No graduate student works in a vacuum. The study of modules, in particular, has a long history filled with folklore, open conjectures, failed conjectures, personalities, hopes and ideas. Complementary to this history is current information: Who is doing what? Who wants to know what? What's hot? What's not? What might make a good thesis topic? All this behind-the-scenes information, though not written down, is as vital to success as the theorems that one claims to be learning.

Thus, in the long development of this dissertation, I have enormously benefitted from the help of my supervisor, Ken Goodearl. He has kept me from trying to prove obviously false theorems, pointed out the probably true ones, and directed me to the appropriate references for those already known. He provided examples and counterexamples, asked the right questions, and knew how to find the answers. Not to be forgotten are the hours he spent checking the proofs, reducing the number of typos and making suggestions for improvements to this dissertation. Perhaps most important, he made the process of getting a Ph.D. a pleasure.

A Note on Notation

We collect here most of the notation that will be used in this dissertation complete with references to the precise definitions whenever possible.

$\mathbb{N}$	Natural numbers, $\{1, 2, 3, \ldots\}$
$\mathbb{Z}$	Integers
$\mathbb{Z}^+$	Non-negative integers, $\{0, 1, 2, 3, \ldots\}$
$(\mathbb{Z}^+)^\infty$	$\mathbb{Z}^+ \cup \{\infty\}$
$\mathbb{Z}_n$	Cyclic group of order $n \in \{2, 3, 4, \ldots\}$
$\mathbb{R}$	Real numbers
$\mathbb{R}^+$	Non-negative real numbers
$\mathbb{R}^{++}$	Strictly positive real numbers
$\mathbb{C}$	Complex numbers
$\{0, \infty\}$	The monoid such that $\infty + \infty = \infty$
Ord	The class of ordinal numbers
$\mathbf{Ord}^*$	$\mathbf{Ord} \cup \{-1\}$
Card	The class of cardinal numbers
Krull	The Krull monoid. See 3.21.

Throughout this dissertation R will be a unital ring. The following are Serre subcategories of R-modules:

R-**Mod**	The category of left R-modules.
R-**Noeth**	The full subcategory of R-**Mod** consisting of all Noetherian modules.
R-**Art**	The full subcategory of R-**Mod** consisting of all Artinian modules.
R-**len**	The full subcategory of R-**Mod** consisting of all modules of finite length.

Though we are using here the nomenclature and notation of category theory, we will only be interested in full subcategories of R-**Mod**. So we will think of categories as subclasses of the objects of R-**Mod**, and modules as elements, rather than objects, of these categories.

Let A and B be modules.

$A \cong B$	A is isomorphic to B.
$\{\cong A\}$	Isomorphism class containing the module A
$A \leq B$	A is a submodule of B
$A \lesssim B$	A is isomorphic to a direct summand of B
$A \sim B$	Modules A and B have isomorphic submodule series. See 16.2.
$[A]$	The $\sim$-equivalence class containing the module A.

Let $\mathcal{L}$ be a preordered class, $\mathcal{M} \subseteq \mathcal{L}$ and $x, y \in \mathcal{L}$.

$\mathcal{L}^\circ$	The dual of $\mathcal{L}$. See 2.1.
$\{\leq x\}$	$\{y \in \mathcal{L} \mid y \leq x\}$
$\{\geq x\}$	$\{y \in \mathcal{L} \mid y \geq x\}$
$[x, y]$	$\{z \in \mathcal{L} \mid x \leq z \leq y\}$
$x \equiv y$	$x \leq y \leq x$
$\{\equiv x\}$	$\{y \in \mathcal{L} \mid y \equiv x\}$
$\overline{\mathcal{L}}$	The universal poclass associated with $\mathcal{L}$. See 2.2.
$\downarrow\mathcal{M}$	The lower class generated by $\mathcal{M}$.
$\Downarrow\mathcal{L}$	The class of lower classes of $\mathcal{L}$ ordered by inclusion. See 2.4.
$\operatorname{Arad}\mathcal{L}$	The Artinian radical of $\mathcal{L}$. See 2.16.
$x \vee y$	The supremum (join) of x and y.
$x \wedge y$	The infimum (meet) of x and y.

Let M and M' be commutative monoids, N a submonoid of M, $\alpha \in \mathbf{Ord}$ and $a, b \in M$.

$M \cong M'$	M is isomorphic to M'. See 5.3.
$I \leq M$	I is an order ideal of M. See 6.12.
$a \leq b$	There is some $c \in M$ such that $a + c = b$
$\{\leq a\}$	$\{b \in M \mid b \leq a\}$
$a \equiv b$	$a \leq b \leq a$. See 6.3.
$\{\equiv a\}$	$\{b \in M \mid b \leq a \leq b\}$, the $\equiv$-equivalence class containing $a \in M$. See 6.3.
$\overline{M}$	The universal partially ordered monoid constructed from M. See 6.3.
$a \ll b$	$a + b \leq b$. See 6.5.
$\{\ll a\}$	$\{b \in M \mid b \ll a\}$
$a \prec b$	There is some $n \in \mathbb{N}$ such that $a \leq nb$. See 6.16.
$\{\prec a\}$	$\{b \in M \mid b \prec a\}$
$a \asymp b$	$a \prec b \prec a$. See 6.18.
$\{\asymp a\}$	$\{b \in M \mid b \prec a \prec b\}$, the Archimedean component containing a.
$\widetilde{M}$	The universal semilattice monoid constructed from M. See 6.18.
$a \sim_r b$	$a + r = b + r$. See 10.2.
$[a]_r$	The $\sim_r$-congruence class containing $a \in M$.
$a \sim_N b$	See 5.5
$[a]_N$	The $\sim_N$-congruence class containing $a \in M$.
$\operatorname{Arad} M$	The Artinian radical of M. See 12.3.
$\operatorname{Arad}_\alpha M$	See 14.3.
$\operatorname{srad} M$	The semi-Artinian radical of M. See 14.3.
$\operatorname{soc} M$	The socle of M. See 14.21.
$\operatorname{soc}_\alpha M$	See 14.25.
$\operatorname{Lrad} M$	The Loewy radical of M. See 14.25.

2 Artinian Ordered Classes

In this section we present the notation, definitions, and propositions we will need to discuss the order structure of monoids, and hence, the order structure of module categories. The monoids that we will study may not be sets - typically a monoid will be constructed from a module category, and the fact that the objects of a category do not, in general, form a set is reflected in the size of the monoids we will have to deal with. As a consequence, we discuss in this section ordered classes rather than ordered sets.

The prototypical ordered class is the class of ordinals, $(\mathbf{Ord}, \leq)$. **Ord** has the property that for any ordinal α, the elements of **Ord** which are less than α form a set. This is typical of the ordered classes which will arise from module categories. See, in particular, 16.13.

Definition 2.1.

- *A* **partially ordered class**, *$(\mathcal{L}, \leq)$, is a nonempty class $\mathcal{L}$ and a relation $\leq$ on $\mathcal{L}$ which is*
  1. *reflexive:* $(\forall x \in \mathcal{L})\ (x \leq x)$
  2. *transitive:* $(\forall x, y, z \in \mathcal{L})\ (x \leq y \text{ and } y \leq z \implies x \leq z)$
  3. *antisymmetric:* $(\forall x, y \in \mathcal{L})\ (x \leq y \text{ and } y \leq x \implies x = y)$.

  We will abbreviate "partially ordered class" as "poclass", and "partially ordered set" as "poset". The greatest and least elements of any poclass $\mathcal{L}$, if they exist will be labeled $\top$ and $\perp$ respectively. A **bounded poclass** *is a poclass which has both a maximum and a minimum element.*
- *A* **preordered class**, *$(\mathcal{L}, \leq)$, is a nonempty class $\mathcal{L}$ and a relation $\leq$ on $\mathcal{L}$ which is reflexive and transitive, that is, it satisfies 1 and 2 above. Every preordered class $\mathcal{L}$ has a* **dual** *preordered class, $\mathcal{L}^\circ$, with the same elements as $\mathcal{L}$ but reverse order.*
- *A function $\psi\colon \mathcal{K} \to \mathcal{L}$ between preordered classes is* **increasing** *if*

$$(\forall x_1, x_2 \in \mathcal{K})\ (x_1 \leq x_2 \implies \psi(x_1) \leq \psi(x_2)).$$

  Note that if ψ is increasing as a function from $\mathcal{K}$ to $\mathcal{L}$, it is also increasing as a function from $\mathcal{K}^\circ$ to $\mathcal{L}^\circ$. A function $\psi\colon \mathcal{K} \to \mathcal{L}$ between preordered classes is **decreasing** *if it is increasing as a function between $\mathcal{K}$and $\mathcal{L}^\circ$.*
- *A function $\psi\colon \mathcal{K} \to \mathcal{L}$ between preordered classes is an* **isomorphism** *if it is increasing and there exists an increasing function $\phi\colon \mathcal{L} \to \mathcal{K}$ such that $\phi \circ \psi$ and $\psi \circ \phi$ are the identity maps on $\mathcal{K}$ and $\mathcal{L}$ respectively. If such a function exists then $\mathcal{K}$ and $\mathcal{L}$ are* **isomorphic**. *Note that a bijective increasing function may not be an isomorphism. See 2.9 for an example.*

  If $\mathcal{K}$ and $\mathcal{L}^\circ$ are isomorphic, then we say $\mathcal{K}$ and $\mathcal{L}$ are **anti-isomorphic**.
- *If a function $\psi\colon \mathcal{K} \to \mathcal{L}$ between preordered classes restricts to an isomorphism $\psi\colon \mathcal{K} \to \psi(\mathcal{K})$, we will say that ψ is an* **embedding**, *and that $\mathcal{K}$ is embedded in $\mathcal{L}$. It is easy to check that ψ is an embedding if and only if*

$$(\forall x_1, x_2 \in \mathcal{K})\ (x_1 \leq x_2 \iff \psi(x_1) \leq \psi(x_2)).$$

Of course, any subclass of a preordered class $\mathcal{L}$ is embedded in $\mathcal{L}$ by the inclusion map.

It is convenient to use the the following notation: Let $\mathcal{L}$ be a preordered class, and $x \in \mathcal{L}$. We will write

$$\{\leq x\} = \{z \in \mathcal{L} \mid z \leq x\}$$
$$\{\geq x\} = \{z \in \mathcal{L} \mid z \geq x\}.$$

This form of notation we will extend to other relations as appropriate, for example, $\{< x\}$, $\{\equiv x\}$, $\{\prec x\}$. Also, if $x \leq y$ in $\mathcal{L}$, we will write

$$[x, y] = \{z \in \mathcal{L} \mid x \leq z \leq y\}.$$

Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two preordered classes. We will write $\mathcal{L}_1 \times \mathcal{L}_2$ for the Cartesian product of $\mathcal{L}_1$ and $\mathcal{L}_2$ with order given by

$$(x_1, x_2) \leq (y_1, y_2) \iff (x_1 \leq y_1 \text{ and } x_2 \leq y_2)$$

for all $(x_1, x_2), (y_1, y_2) \in \mathcal{L}_1 \times \mathcal{L}_2$. Notice that $\{\leq (x_1, x_2)\} = \{\leq x_1\} \times \{\leq x_2\}$ for $(x_1, x_2) \in \mathcal{L}_1 \times \mathcal{L}_2$.

There are projection maps $\pi_1\colon \mathcal{L}_1 \times \mathcal{L}_2 \to \mathcal{L}_1$ and $\pi_2\colon \mathcal{L}_1 \times \mathcal{L}_2 \to \mathcal{L}_2$ defined by $\pi_1(x_1, x_2) = x_1$ and $\pi_2(x_1, x_2) = x_2$. Also each of $\mathcal{L}_1$ and $\mathcal{L}_2$ can be embedded in $\mathcal{L}_1 \times \mathcal{L}_2$: If $x_1 \in \mathcal{L}_1$ then the map from $\mathcal{L}_2$ to $\mathcal{L}_1 \times \mathcal{L}_2$ given by $x_2 \mapsto (x_1, x_2)$ is an embedding. We have defined preordered classes to be nonempty so that such embeddings always exist.

It is easy to show that $\mathcal{L}_1 \times \mathcal{L}_2 \cong \mathcal{L}_2 \times \mathcal{L}_1$ and $(\mathcal{L}_1 \times \mathcal{L}_2) \times \mathcal{L}_3 \cong \mathcal{L}_1 \times (\mathcal{L}_2 \times \mathcal{L}_3)$.

If $\psi_1\colon \mathcal{K}_1 \to \mathcal{L}_1$ and $\psi_2\colon \mathcal{K}_2 \to \mathcal{L}_2$ are maps between preordered classes, then we will write $\psi_1 \times \psi_2$ for the map from $\mathcal{K}_1 \times \mathcal{K}_2$ to $\mathcal{L}_1 \times \mathcal{L}_2$ given by

$$(\psi_1 \times \psi_2)(x_1, x_2) = (\psi_1(x_1), \psi_2(x_2)).$$

If ψ_1 and ψ_2 are increasing, then so is $\psi_1 \times \psi_2$.

Any preordered class $\mathcal{L}$ has a poclass, to be called $\overline{\mathcal{L}}$, associated with it:

Definition 2.2. *Let $\mathcal{L}$ be a preordered class. Define a relation $\equiv$ on $\mathcal{L}$ by*

$$x \equiv y \iff x \leq y \text{ and } y \leq x.$$

It is easily shown that $\equiv$ is an equivalence relation. We will write $\{\equiv x\}$ for the equivalence class containing $x \in \mathcal{L}$, and $\overline{\mathcal{L}} = \mathcal{L}/\equiv$ for the class of equivalence classes. $\overline{\mathcal{L}}$ is a poclass when given the order

$$\{\equiv x\} \leq \{\equiv y\} \iff x \leq y.$$

Thus the map $\{\equiv\ \}\colon \mathcal{L} \to \overline{\mathcal{L}}$ taking x to $\{\equiv x\}$ is increasing. Of course, if $\mathcal{L}$ happened to be partially ordered to start with, then this map is an isomorphism of $\mathcal{L}$ and $\overline{\mathcal{L}}$.

Proposition 2.3. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be an increasing function between preordered classes. If $\mathcal{L}$ is a poclass, then there exists a unique increasing function $\bar{\psi}$ making the following diagram commute.*

$$\begin{tikzcd} \mathcal{K} \arrow[r, "\psi"] \arrow[d, "\{\equiv\ \}"'] & \mathcal{L} \\ \overline{\mathcal{K}} \arrow[ur, "\bar{\psi}"'] & \end{tikzcd}$$

Proof. Straight forward. □

If $\mathcal{M}$ is a nonempty subclass of a preordered class $\mathcal{L}$, then $\overline{\mathcal{M}}$ is defined as above, independently of its embedding in $\mathcal{L}$. But since $\overline{\mathcal{M}}$ is easily seen to be isomorphic to $\{\{\equiv x\} \in \overline{\mathcal{L}} \mid x \in \mathcal{M}\} \subseteq \overline{\mathcal{L}}$, we will make the following identification:

$$\overline{\mathcal{M}} = \{\{\equiv x\} \in \overline{\mathcal{L}} \mid x \in \mathcal{M}\}.$$

One can readily check that for preordered classes $\mathcal{L}_1$ and $\mathcal{L}_2$, we have $\overline{\mathcal{L}_1 \times \mathcal{L}_2} = \overline{\mathcal{L}}_1 \times \overline{\mathcal{L}}_2$.

Definition 2.4. *A subclass $\mathcal{J} \subseteq \mathcal{L}$ of a preordered class is a* **lower class** *of $\mathcal{L}$ if*

$$(\forall x \in \mathcal{J})\ (\{\leq x\} \subseteq \mathcal{J}).$$

The class of lower classes of $\mathcal{L}$ we will denote by $\Downarrow\mathcal{L}$, and we will order $\Downarrow\mathcal{L}$ by inclusion. Some authors use "hereditary subclass" for "lower class".

A subclass $\mathcal{J} \subseteq \mathcal{L}$ of a preordered class is an **upper class** *of $\mathcal{L}$ if*

$$(\forall x \in \mathcal{J})\ (\{\geq x\} \subseteq \mathcal{J}).$$

The class of upper classes of $\mathcal{L}$ we will denote by $\Uparrow\mathcal{L}$, and we will order $\Uparrow\mathcal{L}$ by inclusion.

Thus $\emptyset$ and $\mathcal{L}$ are upper classes and lower classes of $\mathcal{L}$. Also $\{\leq x\}$ is a lower class and $\{\geq x\}$ is an upper class of $\mathcal{L}$ for any $x \in \mathcal{L}$. Note that $\Downarrow\mathcal{L}$ and $\Uparrow\mathcal{L}$ are poclasses even if $\mathcal{L}$ is only preordered.

It is easily checked that arbitrary unions and intersections of lower classes are also lower classes, (thus $\Downarrow\mathcal{L}$ is a complete bounded lattice). In particular, if $\mathcal{M} \subseteq \mathcal{L}$ is an arbitrary subclass, then $\mathcal{M}$ generates a lower class $\downarrow\mathcal{M}$ defined by

$$\downarrow\mathcal{M} = \bigcap\{\mathcal{J} \in \Downarrow\mathcal{L} \mid \mathcal{M} \subseteq \mathcal{J}\},$$

or equivalently,

$$\downarrow\mathcal{M} = \bigcup\{\{\leq x\} \mid x \in \mathcal{M}\}.$$

If $\mathcal{J}$ is a lower class in $\mathcal{L}$, then it is a union of $\equiv$-equivalence classes and $\overline{\mathcal{J}}$ is a lower class in $\overline{\mathcal{L}}$. In fact, it is easy to see that the function $\{\equiv\ \}$ induces an isomorphism $\{\equiv\ \}\colon\ \Downarrow\mathcal{L} \to \Downarrow\overline{\mathcal{L}}$.

Of course, everything we have said in the previous paragraph applies to upper classes as well. In particular, we will write $\uparrow\mathcal{M}$ for the upper class generated by $\mathcal{M} \subseteq \mathcal{L}$. If $\mathcal{J}$ is a lower class of $\mathcal{L}$ then the complement of $\mathcal{J}$, $\mathcal{L} \setminus \mathcal{J}$, is an upper class, and this is easily seen to provide an anti-isomorphism between $\Downarrow\mathcal{L}$ and $\Uparrow\mathcal{L}$.

We introduce next certain classes of functions (exact, exact° and strictly increasing) which, we will find later (2.17), behave well with respect to Artinian and Noetherian preordered classes.

First we note that increasing functions can be characterized by how they act on upper and lower classes:

Proposition 2.5. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a function between preordered classes. Then the following are equivalent:*

1. *ψ is an increasing function.*

2. $\{\leq \psi(x)\} \supseteq \psi(\{\leq x\})$ *for all* $x \in \mathcal{K}$.
3. ψ^{-1} *maps lower classes of* $\mathcal{L}$ *to lower classes of* $\mathcal{K}$.
4. $\{\geq \psi(x)\} \supseteq \psi(\{\geq x\})$ *for all* $x \in \mathcal{K}$.
5. ψ^{-1} *maps upper classes of* $\mathcal{L}$ *to upper classes of* $\mathcal{K}$.

Proof.

$1 \Rightarrow 2$ If $y \in \psi(\{\leq x\})$, then there is $x' \leq x$ such that $y = \psi(x')$. Since ψ is increasing, $\psi(x') \leq \psi(x)$, that is, $y \in \{\leq \psi(x)\}$.

$2 \Rightarrow 3$ Let $\mathcal{M}$ be a lower class of $\mathcal{L}$. If $x \in \psi^{-1}(\mathcal{M})$ then we have $\psi(x) \in \mathcal{M}$ and also $\psi(\{\leq x\}) \subseteq \{\leq \psi(x)\} \subseteq \mathcal{M}$. Thus $\{\leq x\} \subseteq \psi^{-1}(\mathcal{M})$.

$3 \Rightarrow 1$ If $x \in \mathcal{K}$ then $\{\leq \psi(x)\}$ is a lower class in $\mathcal{L}$ so by hypothesis, $\psi^{-1}(\{\leq \psi(x)\})$ is a lower class in $\mathcal{L}$. Since x is in $\psi^{-1}(\{\leq \psi(x)\})$, so is any $x' \leq x$, that is $\psi(x') \leq \psi(x)$.

By duality, 1 implies 4 implies 5 implies 1. □

We will be interested in functions which have properties "inverse" to those of increasing functions:

Proposition 2.6. *Let* $\psi\colon \mathcal{K} \to \mathcal{L}$ *be a function between preordered classes. Then the following are equivalent:*

1. $\{\leq \psi(x)\} \subseteq \psi(\{\leq x\})$ *for all* $x \in \mathcal{K}$.
2. ψ *maps lower classes of* $\mathcal{K}$ *to lower classes of* $\mathcal{L}$.

Also the following are equivalent:

3. $\{\geq \psi(x)\} \subseteq \psi(\{\geq x\})$ *for all* $x \in \mathcal{K}$.
4. ψ *maps upper classes of* $\mathcal{K}$ *to upper classes of* $\mathcal{L}$.

Proof.

$1 \Rightarrow 2$ Let $\mathcal{M}$ be a lower class of $\mathcal{K}$ and $y \in \psi(\mathcal{M})$. Then there is some $x \in \mathcal{M}$ such that $y = \psi(x)$, so $\{\leq y\} = \{\leq \psi(x)\} \subseteq \psi(\{\leq x\}) \subseteq \psi(\mathcal{M})$.

$2 \Rightarrow 1$ If $x \in \mathcal{K}$ then $\{\leq x\}$ is a lower class in $\mathcal{K}$ and hence, $\psi(\{\leq x\})$ is a lower class in $\mathcal{L}$. Since $\psi(x)$ is in $\psi(\{\leq x\})$, we have $\{\leq \psi(x)\} \subseteq \psi(\{\leq x\})$.

Of course, the equivalence of 3 and 4 follows by duality. □

Items 1 and 2 are <u>not</u> equivalent to 3 and 4, so we are led to define two types of functions:

Definition 2.7. *A function* $\psi\colon \mathcal{K} \to \mathcal{L}$ *between poclasses is* **exact** *if it satisfies either 1 or 2 of the conditions of the previous proposition. It is* **exact°(dual-exact)** *if it satisfies either of conditions 3 or 4.*

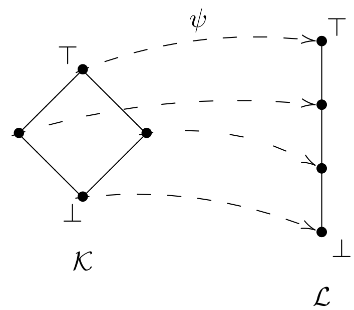
Some easy examples:

- The map $\{\equiv\ \}\colon \mathcal{L} \to \overline{\mathcal{L}}$ is increasing, exact, exact° and surjective.
- If $\mathcal{M} \subseteq \mathcal{L}$, then the embedding of $\mathcal{M}$ into $\mathcal{L}$ is always increasing, but it is exact if and only if $\mathcal{M}$ is a lower class, and exact° if and only if $\mathcal{M}$ is an upper class. Thus the restriction of an exact function to a lower class is an exact function.
- The projections π_1 and π_2 of $\mathcal{L}_1 \times \mathcal{L}_2$ onto $\mathcal{L}_1$ and $\mathcal{L}_2$, are increasing, exact and exact°.

Proposition 2.8. *Let* $\psi\colon \mathcal{K} \to \mathcal{L}$ *be a bijection. Then* ψ *is exact if and only if* ψ *is exact° if and only if* ψ^{-1} *is increasing.*

Proof. Easy consequence of 2.5 and 2.6. □

Example 2.9. *A function $\psi\colon \mathcal{K} \to \mathcal{L}$ which illustrates exactness, and the relations among the previous three propositions is:*



The function ψ is increasing but not exact or exact$^\circ$. The function ψ^{-1} is exact and exact$^\circ$ but not increasing. Note also that ψ is an increasing bijection but not an isomorphism.

Definition 2.10. *A function $\psi\colon \mathcal{K} \to \mathcal{L}$ between preordered classes is* **strictly increasing** *if it is increasing and*

$$(\forall x_1, x_2 \in \mathcal{K})\ (x_1 \leq x_2 \text{ and } \psi(x_1) \geq \psi(x_2) \implies x_1 \geq x_2).$$

If $\mathcal{K}$ and $\mathcal{L}$ are poclasses, then this definition coincides with the usual one, namely, $x_1 < x_2$ implies $\psi(x_1) < \psi(x_2)$. The reason for the peculiar form of this definition is that we want a map $\psi\colon \mathcal{K} \to \mathcal{L}$ between preordered classes to be strictly increasing if and only if the induced map between $\overline{\mathcal{K}}$ and $\overline{\mathcal{L}}$ is strictly increasing in the usual sense. Exact and exact$^\circ$ maps also have this property:

Proposition 2.11. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be an increasing function between preordered classes. Then there is a unique increasing function $\bar{\psi}\colon \overline{\mathcal{K}} \to \overline{\mathcal{L}}$ which makes the following diagram commute:*

$$\begin{tikzcd}
\mathcal{K} \arrow[r, "\psi"] \arrow[d, "\{\equiv\ \}"'] & \mathcal{L} \arrow[d, "\{\equiv\ \}"] \\
\overline{\mathcal{K}} \arrow[r, "\bar{\psi}"] & \overline{\mathcal{L}}
\end{tikzcd}$$

Further, $\bar{\psi}$ is exact (exact$^\circ$, strictly increasing) if and only if ψ is exact (exact$^\circ$, strictly increasing).

Proof. Using 2.3, we get a unique increasing function from $\overline{\mathcal{K}}$ to $\overline{\mathcal{L}}$ making the diagram commute. The claim about exact and exact$^\circ$ functions follows from the fact that the functions $\{\equiv\ \}$ induce isomorphisms between $\Downarrow\mathcal{L}$ and $\Downarrow\overline{\mathcal{L}}$, and between $\Downarrow\mathcal{K}$ and $\Downarrow\overline{\mathcal{K}}$.

It is a simple calculation to establish the property of strictly increasing functions. □

It is also an easy calculation to check the following:

- The composition of increasing (strictly increasing, exact, exact$^\circ$) functions is again increasing (strictly increasing, exact, exact$^\circ$).

- If $\psi_1\colon \mathcal{K}_1 \to \mathcal{L}_1$ and $\psi_2\colon \mathcal{K}_2 \to \mathcal{L}_2$ are maps between preordered classes, and ψ_1 and ψ_2 are increasing (strictly increasing, exact, exact°), then so is $\psi_1 \times \psi_2$.

Lemma 2.12. *Consider the following commutative diagram with $\mathcal{K}$, $\mathcal{L}$ and $\mathcal{M}$ preordered classes:*

$$\begin{tikzcd} \mathcal{K} \arrow[r, "\sigma"] \arrow[dr, "\psi"'] & \mathcal{L} \arrow[d, "\tau"] \\ & \mathcal{M} \end{tikzcd}$$

1. *If σ is surjective and increasing, and ψ is exact (exact°), then τ is exact (exact°).*
2. *If τ is injective and increasing, and ψ is exact (exact°), then σ is exact (exact°).*

Proof. We prove 1 in the exact case only. The other cases are similar:
Let $\mathcal{J}$ be a lower class in $\mathcal{L}$. Since σ is surjective, we have $\sigma(\sigma^{-1}(\mathcal{J})) = \mathcal{J}$, and so

$$\psi(\sigma^{-1}(\mathcal{J})) = \tau(\sigma(\sigma^{-1}(\mathcal{J}))) = \tau(\mathcal{J}).$$

The function ψ is exact and σ is increasing, so $\psi(\sigma^{-1}(\mathcal{J})) = \tau(\mathcal{J})$ is a lower class in $\mathcal{M}$. Since τ maps lower classes of $\mathcal{L}$ to lower classes of $\mathcal{M}$, it is an exact function. $\square$

The following definitions are standard for poclasses:

Definition 2.13.

- *A poclass $\mathcal{L}$ is* **Artinian** *if either of the following equivalent properties is true:*
  1. *Every nonempty subclass of $\mathcal{L}$ has a minimal element.*
  2. *For every decreasing sequence $x_1 \geq x_2 \geq x_3 \geq \dots$ in $\mathcal{L}$, there is an $N \in \mathbb{N}$ such that $x_n = x_N$ for all $n \geq N$.*
- *A poclass $\mathcal{L}$ is* **Noetherian** *if either of the following equivalent properties is true:*
  3. *Every nonempty subclass of $\mathcal{L}$ has a maximal element.*
  4. *For every increasing sequence $x_1 \leq x_2 \leq x_3 \leq \dots$ in $\mathcal{L}$, there is an $N \in \mathbb{N}$ such that $x_n = x_N$ for all $n \geq N$.*

The equivalence of 1 and 2 (or 3 and 4) is an easy and standard result. Of course, any subclass of an Artinian (Noetherian) poclass is also Artinian (Noetherian).

We extend these definitions to apply to preordered classes in a rather simple minded but nonetheless useful way:

Definition 2.14. *A preordered class $\mathcal{L}$ is* **Artinian (Noetherian)** *if and only if $\overline{\mathcal{L}}$ is Artinian (Noetherian).*

It is easily checked that if $\mathcal{M}$ is a subclass of a preordered class $\mathcal{L}$, then $\mathcal{M}$ is Artinian (Noetherian) if and only if $\overline{\mathcal{M}}$ is an Artinian (Noetherian) subclass of $\overline{\mathcal{L}}$.

It is a standard result that for poclasses $\mathcal{L}_1$ and $\mathcal{L}_2$, we have that $\mathcal{L}_2 \times \mathcal{L}_2$ is Artinian (Noetherian) if and only if $\mathcal{L}_1$ and $\mathcal{L}_2$ are Artinian (Noetherian). This result extends easily to preordered classes using the equation $\overline{\mathcal{L}_2 \times \mathcal{L}_2} = \overline{\mathcal{L}}_1 \times \overline{\mathcal{L}}_2$, which is valid for preordered classes.

It is convenient to define minimal (maximal) elements for preordered classes in such a way that the definitions of Artinian and Noetherian preordered classes take the same form as in 2.13:

Definition 2.15. *If $\mathcal{M}$ is a subclass of a preordered class $\mathcal{L}$, then an element $x \in \mathcal{M}$ is* **minimal** *in $\mathcal{M}$ if*

$$(\forall y \in \mathcal{M})\ (y \leq x \implies y \geq x).$$

Maximal elements are defined dually.

In a poclass, of course, minimal and maximal have their usual meanings. We have made these definitions so that $x \in \mathcal{M}$ is minimal in $\mathcal{M}$ if and only if $\{\equiv x\}$ is minimal in $\overline{\mathcal{M}}$.

With this definition, a preordered class $\mathcal{L}$ is Artinian if and only if any of the following are true:

1. Every nonempty subclass of $\mathcal{L}$ has a minimal element.
2. For every decreasing sequence $x_1 \geq x_2 \geq x_3 \geq \ldots$ there is an $N \in \mathbb{N}$ such that $x_n \geq x_N$ for all $n \geq N$.

Of course, a dual statement is true for Noetherian preordered classes.

We have noted already that an arbitrary union of lower classes is a lower class. It is also easy to see that an arbitrary union of Artinian lower classes is Artinian: Any decreasing sequence in the union, $x_1 \geq x_2 \geq x_3 \geq \ldots$, is contained in the same Artinian lower class that contains x_1. Thus we define

Definition 2.16. *The* **Artinian radical** *of a preordered class $\mathcal{L}$, written* $\operatorname{Arad}\mathcal{L}$*, is the union of all Artinian lower classes. As we have seen, it is also the largest Artinian lower class.*

Clearly, a lower class is Artinian if and only if it is contained in the Artinian radical.

Now we can we present the main theorem of this section:

Theorem 2.17. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a function between preordered classes and $\emptyset \neq \mathcal{M} \subseteq \mathcal{L}$.*

1. *If ψ is strictly increasing and $\mathcal{M}$ is Artinian, then $\psi^{-1}(\mathcal{M})$ is Artinian.*
2. *If ψ is strictly increasing and $\mathcal{M}$ is Noetherian, then $\psi^{-1}(\mathcal{M})$ is Noetherian.*
3. *If ψ is exact and increasing, and $\mathcal{K}$ is Artinian, then $\psi(\mathcal{K})$ is Artinian.*
4. *If ψ is exact$^\circ$ and increasing, and $\mathcal{K}$ is Noetherian, then $\psi(\mathcal{K})$ is Noetherian.*

If $\mathcal{K}$ is a poclass, then the hypothesis that ψ is increasing can be dropped from 3 *and* 4*.*

Proof.

1. Let $\mathcal{J} \subseteq \psi^{-1}(\mathcal{M})$ be nonempty, and let $x_0 \in \mathcal{J}$ be chosen so that $\psi(x_0)$ is minimal in $\psi(\mathcal{J}) \subseteq \mathcal{M}$. We claim x_0 is minimal in $\mathcal{J}$.

   If $y \in \mathcal{J}$ and $y \leq x_0$, then $\psi(y) \leq \psi(x_0)$, so, by the minimality of $\psi(x_0)$, we have $\psi(y) \geq \psi(x_0)$. Since ψ is strictly increasing, this implies $x_0 \leq y$.
2. Dually.
3. Let $\mathcal{J} \subseteq \psi(\mathcal{K})$ be nonempty and let x_0 be minimal in $\psi^{-1}(\mathcal{J})$. We claim that $\psi(x_0)$ is minimal in $\mathcal{J}$.

   If $y \in \mathcal{J}$ and $y \leq \psi(x_0)$ then $y \in \{\leq \psi(x_0)\} \subseteq \psi(\{\leq x_0\})$, so there is some $x \leq x_0$ such that $y = \psi(x)$. But $x \in \psi^{-1}(\mathcal{J})$, so by the minimality of x_0, we get $x \geq x_0$ and, since ψ is increasing, $y = \psi(x) \geq \psi(x_0)$.
4. Dually

If $\mathcal{K}$ is partially ordered, then in the last line of the proof of 3 we get, by the minimality of x_0, that $x = x_0$ and so $y = \psi(x_0)$. □

Corollary 2.18. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a function between preordered classes.*

1. *If ψ is strictly increasing, then* $\psi^{-1}(\operatorname{Arad}\mathcal{L}) \subseteq \operatorname{Arad}\mathcal{K}$.
2. *If ψ is exact and increasing, then* $\psi(\operatorname{Arad}\mathcal{K}) \subseteq \operatorname{Arad}\mathcal{L}$.
3. *If ψ is exact, strictly increasing and injective, then* $\operatorname{Arad}\mathcal{K} = \psi^{-1}(\operatorname{Arad}\mathcal{L})$.
4. *If ψ is exact, strictly increasing and surjective, then* $\operatorname{Arad}\mathcal{L} = \psi(\operatorname{Arad}\mathcal{K})$.
5. *If $\mathcal{J}$ is a lower class in $\mathcal{L}$, then* $\operatorname{Arad}\mathcal{J} = (\operatorname{Arad}\mathcal{L}) \cap \mathcal{J}$.

Proof.

1. $\operatorname{Arad}\mathcal{L}$ is an Artinian lower class, so by 2.17 and 2.5, $\psi^{-1}(\operatorname{Arad}\mathcal{L})$ is an Artinian lower class in $\mathcal{K}$.
2. $\operatorname{Arad}\mathcal{K}$ is an Artinian lower class, so by 2.17 and 2.6, $\psi(\operatorname{Arad}\mathcal{K})$ is an Artinian lower class in $\mathcal{L}$.
3. Using 1 and 2 we get
$$\operatorname{Arad}\mathcal{K} = \psi^{-1}(\psi(\operatorname{Arad}\mathcal{K})) \subseteq \psi^{-1}(\operatorname{Arad}\mathcal{L}) \subseteq \operatorname{Arad}\mathcal{K}.$$
   Thus $\operatorname{Arad}\mathcal{K} = \psi^{-1}(\operatorname{Arad}\mathcal{L})$.
4. Using 1 and 2 we get
$$\operatorname{Arad}\mathcal{L} = \psi(\psi^{-1}(\operatorname{Arad}\mathcal{L})) \subseteq \psi(\operatorname{Arad}\mathcal{K}) \subseteq \operatorname{Arad}\mathcal{L}.$$
   Thus $\operatorname{Arad}\mathcal{L} = \psi(\operatorname{Arad}\mathcal{K})$.
5. The inclusion $\psi\colon \mathcal{J} \to \mathcal{L}$ is exact, strictly increasing and injective, so from 3 we get $\operatorname{Arad}\mathcal{J} = \psi^{-1}(\operatorname{Arad}\mathcal{L}) = (\operatorname{Arad}\mathcal{L}) \cap \mathcal{J}$.

□

Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be an increasing function between preordered classes. From 2.17 we have that, if $\mathcal{K}$ is Artinian and ψ is exact, then $\psi(\mathcal{K})$ is Artinian. We will see next that, with the stronger hypothesis on $\mathcal{K}$ that $\Downarrow\mathcal{K}$ is Artinian, and the weaker hypothesis on ψ that ψ is increasing, we can get the same conclusion.

Though everything here applies, via 2.11, to preordered classes, it will be convenient to discuss only poclasses. We introduce some standard concepts for poclasses:

Definition 2.19.

- *A* **chain** *is a subclass* $\mathcal{C}$ *of a poclass which is totally ordered, that is,* $c_1, c_2 \in \mathcal{C}$ *implies* $c_1 \le c_2$ *or* $c_2 \le c_1$.
- *An* **antichain** *is a subclass* $\mathcal{A}$ *of a poclass such that if* $a_1, a_2 \in \mathcal{A}$ *and* $a_1 \le a_2$*, then* $a_1 = a_2$.

A chain is Artinian and Noetherian if and only if it is finite. More generally we have

Lemma 2.20. *A poclass* $\mathcal{L}$ *is both Artinian and Noetherian if and only if all chains in* $\mathcal{L}$ *are finite.*

Proof. If $\mathcal{L}$ is both Artinian and Noetherian, then any chain is Artinian and Noetherian and so is finite.

Conversely, suppose all chains in $\mathcal{L}$ are finite. Then the image of any decreasing or increasing sequence in $\mathcal{L}$ is finite. Thus $\mathcal{L}$ is Artinian and Noetherian. □

We now consider the condition that $\Downarrow\mathcal{L}$ is Artinian for a poclass $\mathcal{L}$:

Proposition 2.21. [17] *Let* $\mathcal{L}$ *be a poclass. Then the following are equivalent:*

1. $\Downarrow\mathcal{L}$ *is Artinian.*
2. $\mathcal{L}$ *is Artinian and contains no infinite antichains.*
3. *Every nonempty subclass of* $\mathcal{L}$ *has a nonzero finite number of minimal elements.*
4. *Every infinite subclass of* $\mathcal{L}$ *contains an infinite strictly increasing sequence.*
5. *Every infinite sequence* $a_1, a_2, a_3, \dots$ *in* $\mathcal{L}$ *whose image is infinite, contains a strictly increasing subsequence.*
6. *Every nonempty upper class in* $\mathcal{L}$ *is finitely generated.*
7. $\Uparrow\mathcal{L}$ *is Noetherian.*

Proof.

$1 \Rightarrow 2$ $\mathcal{L}$ is embedded in the Artinian poclass $\Downarrow\mathcal{L}$ via the map $x \mapsto \{\le x\}$. Hence $\mathcal{L}$ is Artinian.

Suppose $\mathcal{L}$ contained an infinite antichain $\mathcal{A} = \{a_1, a_2, \dots\}$. Then the sequence of lower classes

$$\downarrow\mathcal{A} \supseteq \downarrow(\mathcal{A} \setminus \{a_1\}) \supseteq \downarrow(\mathcal{A} \setminus \{a_1, a_2\}) \supseteq \dots,$$

is strictly decreasing, so $\Downarrow\mathcal{L}$ could not be Artinian.

$2 \Rightarrow 3$ Since $\mathcal{L}$ is Artinian, any nonempty subclass of $\mathcal{L}$ has a minimal element, and since the minimal elements form an antichain, there must be a finite number of them.

$3 \Rightarrow 4$ Suppose $\mathcal{M}_1 \subseteq \mathcal{L}$ is infinite. Let $\mathcal{A}_1$ be the set of minimal elements of $\mathcal{M}_1$. Then

$$\mathcal{M}_1 = \bigcup_{a \in \mathcal{A}_1} (\{\ge a\} \cap \mathcal{M}_1).$$

Since $\mathcal{M}_1$ is infinite and $\mathcal{A}_1$ is finite, there must be some $a_1 \in \mathcal{A}_1$ such that $\{\ge a_1\} \cap \mathcal{M}_1$ is infinite. Set $\mathcal{M}_2 = \{> a_1\} \cap \mathcal{M}_1$. Then $\mathcal{M}_2$ is infinite, and we can use the above process to get some $a_2 \in \mathcal{M}_2$ such that $\mathcal{M}_3 = \{> a_2\} \cap \mathcal{M}_2$ is infinite. Repetition of this process then gives a strictly increasing sequence $a_1 < a_2 < a_3 < \dots$ in $\mathcal{M}_1$.

$4 \Rightarrow 5$ Let $a_1, a_2, a_3, \ldots$ be an infinite sequence in $\mathcal{L}$ such that $\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$ is infinite. By hypothesis, there is an injective function $\sigma\colon \mathbb{N} \to \mathbb{N}$ such that

$$a_{\sigma(1)} < a_{\sigma(2)} < a_{\sigma(3)} < \ldots$$

is a strictly increasing sequence. We define another function $\psi\colon \mathbb{N} \to \mathbb{N}$ inductively by $\psi(1) = 1$ and

$$\psi(n+1) = \min\{m \in \mathbb{N} \mid \sigma(m) > \sigma(\psi(n)) \text{ and } a_{\sigma(m)} > a_{\sigma(\psi(n))}\}$$

for $n \in \mathbb{N}$. This is well defined since for a fixed n the sets $\{m \mid \sigma(m) \leq \sigma(\psi(n))\}$ and $\{m \mid a_{\sigma(m)} \leq a_{\sigma(\psi(n))}\}$ are finite, so there is some $N \in \mathbb{N}$ such that $\sigma(m) > \sigma(\psi(n))$ and $a_{\sigma(m)} > a_{\sigma(\psi(n))}$ for all $m \geq N$.

By construction, $\sigma(\psi(1)) < \sigma(\psi(2)) < \sigma(\psi(2)) < \ldots$, and so

$$a_{\sigma(\psi(1))} < a_{\sigma(\psi(2))} < a_{\sigma(\psi(3))} < \ldots$$

is a strictly increasing subsequence of the original sequence.

$5 \Rightarrow 6$ Suppose to the contrary that $\mathcal{J} \neq \emptyset$ is an upper class of $\mathcal{L}$ which is not finitely generated. We proceed inductively to construct a sequence $a_1, a_2, \ldots$ of distinct elements in $\mathcal{J}$ as follows:

Let $a_1 \in \mathcal{J}$ be arbitrary. For the induction step, suppose we have already chosen $a_1, a_2, \ldots, a_n \in \mathcal{J}$. Since $\mathcal{J}$ is not finitely generated, $\mathcal{J} \setminus (\uparrow \{a_1, a_2, \ldots, a_n\})$ is nonempty, and we pick a_{n+1} to be any element of this subclass.

By construction, $\{\geq a_n\}$ does not contain a_m for any $m > n$. In particular, for any $n \in \mathbb{N}$, $\{\geq a_n\}$ contains at most a finite number of elements from the sequence. Hence $a_1, a_2, \ldots$ has no infinite strictly increasing subsequence.

$6 \Rightarrow 7$ Let $\mathcal{J}_1 \subseteq \mathcal{J}_2 \subseteq \mathcal{J}_3 \subseteq \ldots$ be an increasing sequence of upper classes. Set $\mathcal{J} = \bigcup_{n \in \mathbb{N}} \mathcal{J}_n$, then $\mathcal{J}$ is also an upper class so is finitely generated. Each of these generators is in some $\mathcal{J}_n$, so there must be some index N such that $\mathcal{J}_N$ contains all the generators. Thus $\mathcal{J}_N = \mathcal{J}$ and $\mathcal{J}_n = \mathcal{J}_N$ for all $n \geq N$.

$7 \Rightarrow 1$ $\Downarrow\mathcal{L}$ and $\Uparrow\mathcal{L}$ are anti-isomorphic.

□

Note that any totally ordered Artinian poclass satisfies 2 of this proposition. In particular, $\Downarrow(\mathbb{Z}^+)$ is Artinian.

Combining 2.20 and this proposition we get immediately

Corollary 2.22. *Let $\mathcal{L}$ be a poclass. Then the following are equivalent:*

1. *$\mathcal{L}$ is finite.*
2. *$\Downarrow\mathcal{L}$ is Artinian and $\mathcal{L}$ is Noetherian.*
3. *$\mathcal{L}$ has no infinite chains or infinite antichains.*

Corollary 2.23. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be an increasing function between poclasses. If $\Downarrow\mathcal{K}$ is Artinian then so are $\Downarrow\psi(\mathcal{K})$ and $\psi(\mathcal{K})$.*

Proof. From 2.5, ψ^{-1} maps lower classes of $\psi(\mathcal{K})$ to lower classes of $\mathcal{K}$. This map is strictly increasing. By 2.17, $\Downarrow\mathcal{K}$ being Artinian then implies that $\Downarrow\psi(\mathcal{K})$ is Artinian. From 2.21, $\Downarrow\psi(\mathcal{K})$ Artinian implies $\psi(\mathcal{K})$ is Artinian. □

Corollary 2.24. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a map between poclasses.*

1. *If $\Downarrow\mathcal{K}$ is Artinian, $\mathcal{L}$ Noetherian and ψ increasing then $\psi(\mathcal{K})$ is finite.*
2. *If $\Downarrow\mathcal{K}$ is Artinian, $\mathcal{L}$ Artinian and ψ decreasing then $\psi(\mathcal{K})$ is finite.*

Proof.

1. From 2.23, $\Downarrow\psi(\mathcal{K})$ is Artinian. We also have $\psi(\mathcal{K})$ is Noetherian, so, by 2.22, $\psi(\mathcal{K})$ is finite.
2. Dually.

□

Item 2 of this corollary can be considered as a generalization of the fact that the image of any decreasing sequence (that is, the image of $\mathbb{Z}^+$ under a decreasing function) in an Artinian poclass is finite.

Proposition 2.25. *Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be poclasses. Then $\Downarrow(\mathcal{L}_1 \times \mathcal{L}_2)$ is Artinian if and only if $\Downarrow\mathcal{L}_1$ and $\Downarrow\mathcal{L}_2$ are Artinian.*

Proof. The poclasses $\Downarrow\mathcal{L}_1$ and $\Downarrow\mathcal{L}_2$ are easily seen to embed in $\Downarrow(\mathcal{L}_1 \times \mathcal{L}_2)$. Thus if $\Downarrow(\mathcal{L}_1 \times \mathcal{L}_2)$ is Artinian, then so are $\Downarrow\mathcal{L}_1$ and $\Downarrow\mathcal{L}_2$.

Conversely, suppose $\Downarrow\mathcal{L}_1$ and $\Downarrow\mathcal{L}_2$ are Artinian. Let $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots$ be an infinite sequence in $\mathcal{L}_1 \times \mathcal{L}_2$ such that $\mathcal{A} = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots\} \subseteq \mathcal{L}_1 \times \mathcal{L}_2$ is infinite. We will show that this sequence has a strictly increasing subsequence...

Since $\mathcal{A}$ is infinite, at least one of the projections $\pi_1(\mathcal{A}) \subseteq \mathcal{L}_1$ and $\pi_2(\mathcal{A}) \subseteq \mathcal{L}_2$ must be infinite. Without loss of generality, we will assume that $\pi_1(\mathcal{A}) \subseteq \mathcal{L}_1$ is infinite. Using 2.21.5, there is a subsequence of $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots$ in which the first components are strictly increasing. Call the image of this new sequence $\mathcal{A}'$. We get two cases:

- Suppose $\pi_2(\mathcal{A}')$ is infinite. Then, using 2.21.5 again, there is a subsequence in $\mathcal{A}'$ with the second components strictly increasing. Since the first components are also strictly increasing, this provides a strictly increasing subsequence in $\mathcal{A}$.
- Suppose $\pi_2(\mathcal{A}')$ is finite. Since $\mathcal{A}'$ is infinite, there must be some $y_0 \in \pi_2(\mathcal{A}')$ such that
$$\{x \in \mathcal{L}_1 \mid (x, y_0) \in \mathcal{A}'\}$$
is infinite. The restriction of the subsequence in $\mathcal{A}'$ to those elements whose second component is y_0, is then a strictly increasing subsequence in $\mathcal{A}$.

□

We have already noted that $\Downarrow(\mathbb{Z}^+)$ is Artinian, so using this proposition inductively, we get that $\Downarrow(\mathbb{Z}^+)^n$ is Artinian for all $n \in \mathbb{N}$.

Corollary 2.26. *Let $\mathcal{L}$ be a preordered class, $n \in \mathbb{N}$ and $\psi\colon (\mathbb{Z}^+)^n \to \mathcal{L}$ an increasing function. Then the image of ψ is an Artinian subclass of $\mathcal{L}$.*

Proof. Composing ψ with the canonical map from $\mathcal{L}$ to $\overline{\mathcal{L}}$ gives an increasing map from $(\mathbb{Z}^+)^n$ to a poclass. Applying 2.23, we get that $\overline{\psi((\mathbb{Z}^+)^n)}$ is an Artinian subclass of $\overline{\mathcal{L}}$. Thus $\psi((\mathbb{Z}^+)^n)$ is an Artinian subclass of $\mathcal{L}$. □

This result is crucial in proving that finitely generated submonoids are Artinian in 12.7.

Let $\mathcal{A}$ be an infinite subclass of a poclass $\mathcal{L}$. From 2.21, we get that if $\Downarrow\mathcal{A}$ is Artinian then any infinite subclass of $\mathcal{A}$ contains an infinite strictly increasing sequence. This is an unnecessarily strong condition on $\mathcal{A}$ if one wants to know only that $\mathcal{A}$ itself contains a strictly increasing sequence.

In the next proposition we consider the dual case. This proposition gives a condition on $\mathcal{A}$ that ensures that it contains an infinite strictly decreasing sequence. The significance of this condition will become apparent in the following corollary and in 12.12.

Theorem 2.27. *Let $\mathcal{A}$ be an infinite subclass of a poclass $\mathcal{L}$ such that $\mathcal{A} = \bigcup_{n\in\mathbb{N}} \mathcal{A}_n$, where $\mathcal{A}_n$ is a finite antichain for all $n \in \mathbb{N}$ and*

$$(\forall n \in \mathbb{N})(\forall a \in \mathcal{A}_{n+1})\ (\exists a' \in \mathcal{A}_n \text{ such that } a \leq a').$$

Then $\mathcal{A}$ contains an infinite strictly decreasing sequence.

Proof. Set

$$\mathcal{A}^* = \{(n, a_n) \mid n \in \mathbb{N} \text{ and } a_n \in \mathcal{A}_n\},$$

the disjoint union of the $\mathcal{A}_n$. Let $\pi\colon \mathcal{A}^* \to \mathcal{A}$ be the projection onto $\mathcal{A}$: $\pi(n, a_n) = a_n$. Define a relation on $\mathcal{A}^*$ by $(m, a_m) \leq (n, a_n)$ if $m \geq n$ and there is a chain,

$$a_m \leq a_{m-1} \leq \ldots \leq a_n,$$

such that $a_i \in \mathcal{A}_i$ for $i = m-1, m-2, \ldots, n+1$. It is easy to check that $\leq$ is a partial order on $\mathcal{A}^*$, and that π is an increasing function.

Also easy to see are the following:

1. If $m \geq n$ and $a_m \in \mathcal{A}_m$, then there is some $a_n \in \mathcal{A}_n$ such that $(m, a_m) \leq (n, a_n)$.
2. If $l \geq m \geq n$, $a_l \in \mathcal{A}_l$ and $a_n \in \mathcal{A}_n$ such that $(l, a_l) \leq (n, a_n)$, then there is some $a_m \in \mathcal{A}_m$ such that $(l, a_l) \leq (m, a_m) \leq (n, a_n)$

We will construct a decreasing sequence $(1, a_1) \geq (2, a_2) \geq (3, a_3) \geq \ldots$ in $\mathcal{A}^*$ such that $\pi(\{\leq (n, a_n)\})$ is infinite for all $n \in \mathbb{N}$:

From 1, every element of $\mathcal{A}^*$ is in $\{\leq (1, a)\}$ for some $a \in \mathcal{A}_1$, that is,

$$\mathcal{A}^* = \bigcup_{a\in\mathcal{A}_1} \{\leq (1, a)\}.$$

Since $\pi(\mathcal{A}^*) = \mathcal{A}$ is infinite and $\mathcal{A}_1$ is finite, there is some $a_1 \in \mathcal{A}_1$ such that $\pi(\{\leq (1, a_1)\})$ is infinite.

We continue by induction: Suppose we have $(n, a_n) \in \mathcal{A}^*$ such that $\pi(\{\leq (n, a_n)\})$ is infinite. Note that, since $\mathcal{A}_n$ is an antichain, $(n, a'_n) \in \{\leq (n, a_n)\}$ if and only if $a'_n = a_n$. Let

$$\mathcal{B}_{n+1} = \mathcal{A}_{n+1} \cap \pi(\{\leq (n, a_n)\}).$$

Then, from 2, every element of $\{\leq (n, a_n)\}$, except (n, a_n), is in $\{\leq (n+1, a)\}$ for some $a \in \mathcal{B}_{n+1}$, that is,

$$\{\leq (n, a_n)\} = \{(n, a_n)\} \cup \bigcup_{a\in\mathcal{B}_{n+1}} \{\leq (n+1, a)\}.$$

Since $\pi(\{\leq (n, a_n)\})$ is infinite and $\mathcal{B}_{n+1}$ is finite, there must be some $a_{n+1} \in \mathcal{B}_{n+1} \subseteq \mathcal{A}_{n+1}$ such that $\pi(\{\leq (n+1, a_{n+1})\})$ is infinite.

We will show that sequence $a_1 \geq a_2 \geq \ldots$ in $\mathcal{A}$ has no minimum element. . .

Suppose to the contrary that there is some $N \in \mathbb{N}$ such that $a_n = a_N$ for all $n \geq N$. This means, in particular, that $a_N \in \mathcal{A}_n$ for all $n \geq N$.

If $(n, a'_n) \leq (N, a_N)$, then $n \geq N$, $a'_n \leq a_N$ and $a'_n, a_N \in \mathcal{A}_n$. Since $\mathcal{A}_n$ is an antichain, this implies $a'_n = a_N$ and hence $\pi(n, a'_n) = a_N$.

This is true for all $(n, a'_n) \in \{\leq (N, a_N)\}$, and so we have $\pi(\{\leq (N, a_N\}) = \{a_N\}$. This, of course, contradicts $\pi(\{\leq (N, a_N)\})$ being infinite. □

Since strictly decreasing sequences in Artinian poclasses are not possible, we get immediately

Corollary 2.28. *Let $\mathcal{A}$ be an subclass of an Artinian poclass $\mathcal{L}$ such that $\mathcal{A} = \bigcup_{n \in \mathbb{N}} \mathcal{A}_n$, where $\mathcal{A}_n$ is a finite antichain for all $n \in \mathbb{N}$ and*

$$(\forall n \in \mathbb{N})(\forall a \in \mathcal{A}_{n+1})\ (\exists a' \in \mathcal{A}_n \text{ such that } a \leq a').$$

Then $\mathcal{A}$ is finite.

Corollary 2.29. [3, page 183] *Let $\mathcal{L}$ be an Artinian poclass. Then the class of all finitely generated lower classes of $\mathcal{L}$ ordered by inclusion is Artinian.*

Proof. Let $\mathcal{J}_1 \supseteq \mathcal{J}_2 \supseteq \ldots$ be a decreasing sequence of finitely generated lower classes of $\mathcal{L}$. For $n \in \mathbb{N}$, let $\mathcal{A}_n$ be a finite set of generators of $\mathcal{J}_n$. By deleting redundant generators from this set, we can assume that $\mathcal{A}_n$ is an antichain. Further, since

$$\mathcal{A}_{n+1} \subseteq \mathcal{J}_{n+1} \subseteq \mathcal{J}_n = \bigcup_{a \in \mathcal{A}_n} \{\leq a\},$$

for every element $a \in \mathcal{A}_{n+1}$, there is some $a' \in \mathcal{A}_n$ such that $a \leq a'$. Thus $\mathcal{A} = \bigcup_{n \in \mathbb{N}} \mathcal{A}_n$ satisfies the hypothesis of 2.28 and $\mathcal{A}$ is finite.

For all $n \in \mathbb{N}$, let $\mathcal{K}_n = \mathcal{J}_n \cap \mathcal{A}$. Note that $\mathcal{K}_n \in \Downarrow \mathcal{A}$, and that $\mathcal{J}_n$ is generated by $\mathcal{K}_n$ as a lower class of $\mathcal{L}$. Then $\mathcal{K}_1 \supseteq \mathcal{K}_2 \supseteq \ldots$ is a decreasing sequence in $\Downarrow \mathcal{A}$, so there is some $N \in \mathbb{N}$ such that $\mathcal{K}_n = \mathcal{K}_N$ for all $n \geq N$. Hence $\mathcal{J}_n = \mathcal{J}_N$ for all $n \geq N$. □

In our discussion of the order in module categories and the order of submodules within modules, we will need to have at hand some information about lattices. The standard reference for lattice theory is G. Birkhoff, *Lattice Theory, 3rd ed.* [3]. For the reasons explained at the beginning of this section, we will allow the possibility that a lattice may be a proper class:

Definition 2.30. *A* **lattice** *is a poclass $\mathcal{L}$ such that every pair of elements, $x, y \in \mathcal{L}$, has a supremum, $x \vee y$, and an infimum, $x \wedge y$.*

Note that $y = y \wedge (x \vee y) = y \vee (x \wedge y)$ in a lattice.

A **bounded lattice** is a lattice which has a maximum element $\top$ and a minimum element $\perp$. Note that if a lattice has a maximal (minimal) element, then this element is maximum (minimum). Thus any Artinian lattice has a minimum element.

A lattice $\mathcal{L}$ is **distributive** if

$$(\forall x, y, z \in \mathcal{L})\ (x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z))$$

and

$$(\forall x, y, z \in \mathcal{L})\ (x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)).$$

These two properties are equivalent and to prove distributivity it suffices to show

$$x \wedge (y \vee z) \leq (x \wedge y) \vee (x \wedge z).$$

A lattice $\mathcal{L}$ is **modular** if

$$(\forall y_1, y_2, x \in \mathcal{L})\ (y_1 \leq y_2 \implies y_1 \vee (x \wedge y_2) = (y_1 \vee x) \wedge y_2).$$

Any distributive lattice is modular.

Lemma 2.31. *Let $\mathcal{L}$ be a lattice and $x \in \mathcal{L}$. Define the map $\lambda \colon \mathcal{L} \to \{\leq x\} \times \{\geq x\}$ by $\lambda(y) = (x \wedge y, x \vee y)$ for all $y \in \mathcal{L}$.*

1. *If $\mathcal{L}$ is modular then λ is strictly increasing.*
2. *If $\mathcal{L}$ is distributive then λ is an embedding.*

Proof. λ is clearly an increasing function.

1. Suppose $y_1 \leq y_2$ in $\mathcal{L}$ such that $\lambda(y_1) = \lambda(y_2)$. Then $x \wedge y_1 = x \wedge y_2$ and $x \vee y_1 = x \vee y_2$, so
$$y_1 = y_1 \vee (x \wedge y_1) = y_1 \vee (x \wedge y_2) = (y_1 \vee x) \wedge y_2 = (y_2 \vee x) \wedge y_2 = y_2.$$
Thus if $y_1 < y_2$ then we must have $\lambda(y_1) < \lambda(y_2)$.
2. Suppose $y_1, y_2 \in \mathcal{L}$ such that $\lambda(y_1) \leq \lambda(y_2)$. Then $x \wedge y_1 \leq x \wedge y_2$ and $x \vee y_1 \leq x \vee y_2$, so
$$\begin{aligned} y_1 &= y_1 \vee (x \wedge y_1) \leq y_1 \vee (x \wedge y_2) = (y_1 \vee x) \wedge (y_1 \vee y_2) \\ &\leq (y_2 \vee x) \wedge (y_1 \vee y_2) = (y_1 \wedge x) \vee y_2 \leq (y_2 \wedge x) \vee y_2 \\ &= y_2. \end{aligned}$$
Thus $y_1 \leq y_2$ if and only if $\lambda(y_1) \leq \lambda(y_2)$.

$\square$

The most important property of modular lattices is the following:

Lemma 2.32. [3, Theorem 13, page 13] *Let $\mathcal{L}$ be a modular lattice and $a, b \in \mathcal{L}$. Then the maps $\phi \colon [b, a \vee b] \to [a \wedge b, a]$ given by $x \mapsto x \wedge a$, and $\psi \colon [a \wedge b, a] \to [b, a \vee b]$ given by $y \mapsto y \vee b$ are inverse isomorphisms.*

Proof. The functions ψ and ϕ are clearly increasing.

Let $x \in [b, a \vee b]$. Applying ϕ to the inequality $b \leq x \leq a \vee b$ we get

$$\phi(b) = b \wedge a \leq \phi(x) \leq \phi(a \vee b) = (a \vee b) \wedge a = a.$$

Thus ϕ maps into $[a \wedge b, a]$ as claimed.

Further,

$$\psi(\phi(x)) = (x \wedge a) \vee b = (a \vee b) \wedge x = x.$$

The second equality is due to the modularity of $\mathcal{L}$ and the inequality $b \leq x$. The last equality comes from the fact that $x \leq a \vee b$. Thus $\psi \circ \phi$ is the identity map on $[b, a \vee b]$.

The remainder of the proof is done similarly. $\square$

3 The Length Function on Ordered Classes

In this section we develop the concept of length functions on poclasses. These are particularly useful when applied to Artinian modular lattices, such as the lattice of submodules of an Artinian or Noetherian module. This we will do in Section 4.

The results to be presented in these two sections are extensions and abstractions of results in T. H. Gulliksen, *A Theory of Length for Noetherian Modules* [13]. With the greater abstraction used in this section we can put Noetherian and Artinian modules on the same footing. In particular, using the lattice theory developed in this section, we will construct in Section 4, the Krull length function on the module categories R-**Noeth** and R-**Art** for a ring R.

This section is outside the main subject of this dissertation and can be skipped by readers who are prepared to accept the existence of the Krull length function. For a simpler development of the Krull length function for Noetherian modules see [6, Section 4].

This section depends heavily on the arithmetic of the ordinal numbers. For the details of ordinal arithmetic see W. Sierpinski, *Cardinal and Ordinal Numbers* [28] or M. D. Potter, *Sets, An Introduction* [25]. We collect here a few of those facts that are relevant:

Notation: We will use lowercase Greek letters for elements of **Ord**, the class of ordinal numbers. The first infinite ordinal is written ω.

- Ordinal addition is associative but not commutative. For example, $\omega+1 \neq 1+\omega = \omega$.
- Ordinal addition is cancellative on the left: $\alpha+\beta = \alpha+\gamma \implies \beta = \gamma$. Also $\alpha+\beta \leq \alpha+\gamma \implies \beta \leq \gamma$.
- For a fixed ordinal α, the map from **Ord** to **Ord** given by $\beta \mapsto \alpha+\beta$ is strictly increasing.
- If $\alpha \leq \beta$, then $\beta-\alpha$ is the unique ordinal γ such that $\beta = \alpha+\gamma$, hence $\beta = \alpha+(\beta-\alpha)$. For any $\alpha, \beta \in$ **Ord**, we have $\beta = (\alpha+\beta)-\alpha$.
- For a fixed ordinal α, the map from $\{\geq \alpha\} \subseteq$ **Ord** to **Ord** given by $\beta \mapsto \beta-\alpha$ is strictly increasing.
- An ordinal $\alpha \neq 0$ has the property that $\beta+\alpha = \alpha$ for all $\beta < \alpha$ if and only if it has the form $\alpha = \omega^\gamma$ for some ordinal γ.
- Any nonzero ordinal can be expressed uniquely in the **normal** form

$$\omega^{\gamma_1} + \omega^{\gamma_2} + \cdots + \omega^{\gamma_n}$$

  where $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$ are ordinals. This same form can be written

$$\omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$$

  where $\gamma_1 > \gamma_2 > \cdots > \gamma_n$ and $n_1, n_2, \ldots, n_n \in \mathbb{N}$. To add two ordinals in either of these forms one needs only to use the rule that $\omega^\gamma + \omega^\delta = \omega^\delta$ if $\gamma < \delta$. For example, $(\omega^\omega + \omega^3 + \omega 2 + 1) + (\omega^3 + \omega) = \omega^\omega + \omega^3 2 + \omega$.

Definition 3.1. *Let $\mathcal{L}$ be a poclass. Then a* **length function** *on $\mathcal{L}$ is a strictly increasing function from $\mathcal{L}$ to* **Ord**.

Proposition 3.2. *If a poclass $\mathcal{L}$ has a length function then $\mathcal{L}$ is Artinian.*

Proof. Since **Ord** is Artinian, we can apply 2.17.1. □

We will shortly prove that the converse of this proposition is true if $\mathcal{L}$ a set. But first, we note that if $\mathcal{L}$ has any length function, then there is a smallest length function on $\mathcal{L}$:

Define $\lambda_{\mathcal{L}} \colon \mathcal{L} \to \mathbf{Ord}$ by

$$\lambda_{\mathcal{L}}(x) = \min\{\lambda(x) \mid \lambda \text{ is a length function on } \mathcal{L}\}$$

for all $x \in \mathcal{L}$. If $x < y$ in $\mathcal{L}$, then there is some length function $\lambda \colon \mathcal{L} \to \mathbf{Ord}$ such that $\lambda(y) = \lambda_{\mathcal{L}}(y)$, so $\lambda_{\mathcal{L}}(x) \le \lambda(x) < \lambda(y) = \lambda_{\mathcal{L}}(y)$. Thus $\lambda_{\mathcal{L}}$ is a length function.

Definition 3.3. *If $\mathcal{L}$ has a length function, then the length function $\lambda_{\mathcal{L}}$ as constructed above will be called the* **minimum length function** *on $\mathcal{L}$. If $\mathcal{L}$ has a maximum element $\top$ then we define the* **length** *of $\mathcal{L}$ by $\operatorname{len} \mathcal{L} = \lambda_{\mathcal{L}}(\top)$.*

Let $\mathcal{L}$ be a poclass with a length function and $\mathcal{M} \subseteq \mathcal{L}$ a nonempty subclass. Then $\lambda_{\mathcal{L}}$ restricts to a length function on $\mathcal{M}$, so for any $x \in \mathcal{M}$ we have $\lambda_{\mathcal{M}}(x) \le \lambda_{\mathcal{L}}(x)$. If $\mathcal{M}$ is, in addition, a lower class of $\mathcal{L}$ we will get the converse inequality. To show this, we consider first the case $\mathcal{M} = \{\le x\}$ for some $x \in \mathcal{L}$.

Proposition 3.4. *If a poclass $\mathcal{L}$ has a length function and $x \in \mathcal{L}$, then $\operatorname{len}\{\le x\} = \lambda_{\mathcal{L}}(x)$.*

Proof. Set $\mathcal{M} = \{\le x\}$ so $\lambda_{\mathcal{M}}(x) = \operatorname{len} \mathcal{M}$. If $\mathcal{M} = \mathcal{L}$ then $x = \top$ and the result is immediate.

Otherwise we have the case where $\mathcal{K} = \mathcal{L} \setminus \mathcal{M}$ is not empty.

Since $\mathcal{K}$ has a length function, we can define $\lambda \colon \mathcal{L} \to \mathbf{Ord}$ by

$$\lambda(y) = \begin{cases} \lambda_{\mathcal{M}}(y) & y \in \mathcal{M} \\ \lambda_{\mathcal{M}}(x) + 1 + \lambda_{\mathcal{K}}(y) & y \in \mathcal{K} \end{cases}$$

for all $y \in \mathcal{L}$. One can easily check that λ is strictly increasing, so that $\lambda_{\mathcal{L}}(x) \le \lambda(x) = \lambda_{\mathcal{M}}(x) = \operatorname{len} \mathcal{M}$.

From the discussion preceding the proposition we have also $\lambda_{\mathcal{L}}(x) \ge \lambda_{\mathcal{M}}(x) = \operatorname{len} \mathcal{M}$, so finally $\operatorname{len}\{\le x\} = \lambda_{\mathcal{L}}(x)$. □

Note that the order of the summation in the definition of λ is crucial to its claimed properties.

Corollary 3.5. *Let $\mathcal{L}$ be a poclass with a length function and $\mathcal{M} \subseteq \mathcal{L}$ a lower class of $\mathcal{L}$. Then $\lambda_{\mathcal{M}}(x) = \lambda_{\mathcal{L}}(x)$ for all $x \in \mathcal{M}$.*

Proof. Given $x \in \mathcal{M}$, we have $\{\le x\} \subseteq \mathcal{M}$. Thus $\lambda_{\mathcal{M}}(x) = \operatorname{len}\{\le x\} = \lambda_{\mathcal{L}}(x)$. □

It can be easily checked that we have defined $\lambda_{\mathcal{L}}$ so that $\lambda_{\mathbf{Ord}}(\alpha) = \operatorname{len}\{\le \alpha\} = \alpha$ for all ordinals α, and $\operatorname{len}[\alpha, \beta] = \beta - \alpha$ for all ordinals $\alpha \le \beta$.

Proposition 3.6. *Let $\mathcal{L}$ be a poclass with a length function and a maximum element.*

1. *$\lambda_{\mathcal{L}}$ is exact.*
2. *If $\alpha \le \operatorname{len} \mathcal{L}$, then there is some $x \in \mathcal{L}$ such that $\lambda_{\mathcal{L}}(x) = \alpha$.*

3. *For all* $x \in \mathcal{L}$, $\operatorname{len}\{\leq x\} + \operatorname{len}[x, \top] \leq \operatorname{len} \mathcal{L}$.
4. *If* $x \in \mathcal{L}$, *then* $\operatorname{len}\{\leq x\} = \operatorname{len} \mathcal{L}$ *if and only if* $\operatorname{len}[x, \top] = 0$.

Proof.

1. We need to show that $\{\leq \lambda_{\mathcal{L}}(x)\} \subseteq \lambda_{\mathcal{L}}(\{\leq x\})$ for all $x \in \mathcal{L}$...
   Suppose to the contrary, that there is an ordinal $\alpha \leq \lambda_{\mathcal{L}}(x)$, such that no $y \leq x$ satisfies $\lambda_{\mathcal{L}}(y) = \alpha$. Clearly $\alpha \neq \lambda_{\mathcal{L}}(x)$ so we have $\alpha < \lambda_{\mathcal{L}}(x)$. Set
   $$\mathcal{K} = \{y \leq x \mid \alpha < \lambda_{\mathcal{L}}(y)\}.$$
   We have $x \in \mathcal{K}$, so $\mathcal{K}$ is not empty. Let $y_0 \in \mathcal{K}$ be chosen so that $\lambda_{\mathcal{L}}(y_0)$ is minimum in $\lambda_{\mathcal{L}}(\mathcal{K}) \subseteq \mathbf{Ord}$. Define $\lambda\colon \mathcal{L} \to \mathbf{Ord}$ by
   $$\lambda(y) = \begin{cases} \lambda_{\mathcal{L}}(y) & y \neq y_0 \\ \alpha & y = y_0 \end{cases}$$
   for all $y \in \mathcal{L}$. It is not hard to show that λ is a length function on $\mathcal{L}$ which is smaller that $\lambda_{\mathcal{L}}$. This contradicts the definition of $\lambda_{\mathcal{L}}$.
2. Since $\lambda_{\mathcal{L}}$ is exact we have $\{\leq \lambda_{\mathcal{L}}(\top)\} \subseteq \lambda_{\mathcal{L}}(\{\leq \top\})$, and so $\{\leq \operatorname{len} \mathcal{L}\} \subseteq \lambda_{\mathcal{L}}(\mathcal{L})$.
3. Define $\lambda\colon [x, \top] \to \mathbf{Ord}$ by $\lambda(y) = \lambda_{\mathcal{L}}(y) - \lambda_{\mathcal{L}}(x)$. The function λ is well defined and strictly increasing, so $\operatorname{len}[x, \top] \leq \lambda(\top) = \lambda_{\mathcal{L}}(\top) - \lambda_{\mathcal{L}}(x) = \operatorname{len} \mathcal{L} - \operatorname{len}\{\leq x\}$. Hence $\operatorname{len}\{\leq x\} + \operatorname{len}[x, \top] \leq \operatorname{len} \mathcal{L}$.
4. Easy.

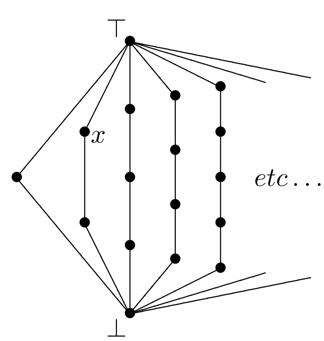
□

From this proposition, $\lambda_{\mathcal{L}}$ is an exact strictly increasing function from $\mathcal{L}$ to $\mathbf{Ord}$. In fact, we will see in 3.11 that $\lambda_{\mathcal{L}}$ is the only such function.

A simple example illustrating 2 and 3 of the proposition is the poset $\mathcal{L} = \{\leq \omega\} \subseteq \mathbf{Ord}$ with $\top = \omega$. For any $n \in \mathbb{N} \subseteq \mathcal{L}$ we have $\operatorname{len}\{\leq n\} = n$ and $\operatorname{len}[n, \top] = \omega$. In particular, $\operatorname{len}\{\leq n\} + \operatorname{len}[n, \top] = n + \omega = \omega = \operatorname{len} \mathcal{L}$, whereas $\operatorname{len}[n, \top] + \operatorname{len}\{\leq n\} = \omega + n > \omega = \operatorname{len} \mathcal{L}$. Thus the order of addition in 3.6.3 is crucial. Notice also that for any $x \in \mathcal{L}$, $\operatorname{len}[x, \top]$ is either 0 or ω, that is, there are only a finite number of possible values for $\operatorname{len}[x, \top]$. This is, in fact, always the case for modular lattices as we will see in 3.35.

Another useful example to have at hand is

Example 3.7. *Let $\mathcal{L}$ be the poset illustrated below which contains chains of all finite lengths but no infinite chain: (Here "etc." means that there are chains of length 6, 7, 8, etc., which are connected between $\top$ and $\perp$).*



Clearly $\operatorname{len}\mathcal{L} = \omega$. *Also,* $\operatorname{len}\{\leq x\} = \lambda_{\mathcal{L}}(x) = 2$ *and* $\operatorname{len}[x, \top] = 1$*, so one can see that, in general, knowing* $\operatorname{len}\{\leq x\}$ *and* $\operatorname{len}[x, \top]$ *does not give any upper bound on* $\operatorname{len}\mathcal{L}$*.*

Proposition 3.8. *If a poclass* $\mathcal{L}$ *has a length function, then* $\lambda_{\mathcal{L}}$ *satisfies*

$$\lambda_{\mathcal{L}}(x) = \begin{cases} 0 & \textit{if } x \textit{ is minimal in } \mathcal{L} \\ \sup\{\lambda_{\mathcal{L}}(y) + 1 \mid y < x\} & \textit{otherwise} \end{cases}$$

for all $x \in \mathcal{L}$*.*

Proof. From 3.4, $\lambda_{\mathcal{L}}(x)$ is determined by $\{\leq x\}$. If x is minimal in $\mathcal{L}$, then $\{\leq x\} = \{x\}$ so $\lambda_{\mathcal{L}}(x) = 0$.

If x is not minimal, then $\{\leq x\} = \{x\} \cup \{< x\}$ with $\{< x\}$ not empty. Since $\lambda_{\mathcal{L}}$ is strictly increasing, we have $\lambda_{\mathcal{L}}(x) > \lambda_{\mathcal{L}}(y)$ for all $y < x$. Indeed, from 3.6.2, $\lambda_{\mathcal{L}}(x)$ must be the smallest such ordinal.

Equivalently, $\lambda_{\mathcal{L}}(x)$ is the smallest ordinal such that $\lambda_{\mathcal{L}}(x) \geq \lambda_{\mathcal{L}}(y) + 1$ for all $y < x$. Thus $\lambda_{\mathcal{L}}(x) = \sup\{\lambda_{\mathcal{L}}(y) + 1 \mid y < x\}$ □

If $\mathcal{L}$ is an Artinian po<u>set</u> then above description of $\lambda_{\mathcal{L}}$ can be used to define $\lambda_{\mathcal{L}}$. Slightly more generally we have the following:

Proposition 3.9. *Let* $\mathcal{L}$ *be an Artinian poclass such that* $\{\leq x\}$ *is a set for all* $x \in \mathcal{L}$*. Then* $\mathcal{L}$ *has a length function, and* $\lambda_{\mathcal{L}}$ *is defined inductively by*

$$\lambda_{\mathcal{L}}(x) = \begin{cases} 0 & \textit{if } x \textit{ is minimal in } \mathcal{L} \\ \sup\{\lambda_{\mathcal{L}}(y) + 1 \mid y < x\} & \textit{otherwise} \end{cases}$$

for all $x \in \mathcal{L}$*.*

Proof. Let $\lambda' \colon \mathcal{L} \to \mathbf{Ord}$ be defined by

$$\lambda'(x) = \begin{cases} 0 & \text{if } x \text{ is minimal in } \mathcal{L} \\ \sup\{\lambda'(y) + 1 \mid y < x\} & \text{otherwise} \end{cases}$$

for all $x \in \mathcal{L}$. It is a standard induction argument to show that λ' is well defined by the above equation, the key point being that any sub<u>set</u> of **Ord** has a supremum, [15, Section 20], so $\sup\{\lambda'(y) + 1 \mid y < x\}$ is defined whenever $\lambda'(y)$ is defined for all $y < x$.

We check that λ' is a length function...

Suppose $x_1 < x_2$ in $\mathcal{L}$. Then $\lambda'(x_2) = \sup\{\lambda'(y) + 1 \mid y < x_2\} \geq \lambda'(x_1) + 1$, so $\lambda'(x_1) < \lambda'(x_2)$. Thus λ' is strictly increasing and a length function.

It remains to show that λ' is the smallest length function and so $\lambda' = \lambda_{\mathcal{L}}$...

Suppose $\lambda \colon \mathcal{L} \to \mathbf{Ord}$ is strictly increasing. Set $\mathbb{B} = \{x \in \mathcal{L} \mid \lambda(x) < \lambda'(x)\}$. If $\mathbb{B} \neq \emptyset$ then it must have a minimal element, x_0. x_0 can not be minimal in $\mathcal{L}$ since $\lambda(x_0) < \lambda'(x_0) = 0$ is not possible. Hence $\lambda'(x_0) = \sup\{\lambda'(y) + 1 \mid y < x_0\}$.

If $y < x_0$ then $y \notin \mathbb{B}$ so we get $\lambda'(y) \leq \lambda(y) < \lambda(x_0)$. Thus $\lambda'(y) + 1 \leq \lambda(x_0)$ for any $y < x_0$. But then we would get $\lambda'(x_0) \leq \lambda(x_0)$, contradicting $x_0 \in \mathbb{B}$.

Therefore we have $\mathbb{B} = \emptyset$ and $\lambda'(x) \leq \lambda(x)$ for all $x \in \mathcal{L}$. □

The most important properties of strictly increasing functions and exact functions are given in the next proposition, which should be compared with 2.17:

Proposition 3.10. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a function between poclasses.*

1. *If ψ is strictly increasing and $\psi(\mathcal{K})$ has a length function, then $\mathcal{K}$ has a length function and $\lambda_{\mathcal{K}}(x) \le \lambda_{\psi(\mathcal{K})}(\psi(x))$ for all $x \in \mathcal{K}$.*
2. *If ψ is exact and $\mathcal{K}$ has a length function, then $\psi(\mathcal{K})$ has a length function and $\lambda_{\psi(\mathcal{K})}(\psi(x)) \le \lambda_{\mathcal{K}}(x)$ for all $x \in \mathcal{K}$.*
3. *If ψ is strictly increasing and exact, then $\psi(\mathcal{K})$ has a length function if and only if $\mathcal{K}$ has a length function, and in that case, $\lambda_{\mathcal{K}}(x) = \lambda_{\psi(\mathcal{K})}(\psi(x))$ for all $x \in \mathcal{K}$.*

Proof.

1. $\lambda_{\psi(\mathcal{K})} \circ \psi\colon \mathcal{K} \to \mathbf{Ord}$ is a length function on $\mathcal{K}$. Thus, for all $x \in \mathcal{K}$, we have $\lambda_{\mathcal{K}}(x) \le \lambda_{\psi(\mathcal{K})}(\psi(x))$.
2. Define $\lambda\colon \psi(\mathcal{K}) \to \mathbf{Ord}$ by $\lambda(y) = \min(\lambda_{\mathcal{K}} \circ \psi^{-1}(y))$ for all $y \in \psi(\mathcal{K})$. Note that $\lambda(\psi(x)) \le \lambda_{\mathcal{K}}(x)$ for all $x \in \mathcal{K}$.

   We will show that λ is strictly increasing. Let $y_1 < y_2$ in $\psi(\mathcal{K})$ and $x_2 \in \mathcal{K}$ be such that $\psi(x_2) = y_2$ and $\lambda_{\mathcal{K}}(x_2) = \lambda(y_2)$. Then $y_1 \in \{\le \psi(x_2)\} \subseteq \psi(\{\le x_2\})$, so there is some $x_1 \le x_2$ such that $\psi(x_1) = y_1$. Since $y_1 \ne y_2$ we must have $x_1 < x_2$ and hence $\lambda(y_1) = \lambda(\psi(x_1)) \le \lambda_{\mathcal{K}}(x_1) < \lambda_{\mathcal{K}}(x_2) = \lambda(y_2)$.

   Since λ is a length function on $\psi(\mathcal{K})$, we have $\lambda_{\psi(\mathcal{K})}(y) \le \lambda(y)$ for all $y \in \psi(\mathcal{K})$. Thus for $x \in \mathcal{K}$, we get $\lambda_{\psi(\mathcal{K})}(\psi(x)) \le \lambda(\psi(x)) \le \lambda_{\mathcal{K}}(x)$.
3. Immediate from 1 and 2.

□

If $\mathcal{L}$ is given with a length function then this proposition can be put in a simpler form:

Corollary 3.11. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a function between poclasses such that $\mathcal{L}$ has a length function.*

1. *If ψ is strictly increasing, then $\mathcal{K}$ has a length function and $\lambda_{\mathcal{K}}(x) \le \lambda_{\mathcal{L}}(\psi(x))$ for all $x \in \mathcal{K}$.*
2. *If ψ is exact and $\mathcal{K}$ has a length function, then $\lambda_{\mathcal{L}}(\psi(x)) \le \lambda_{\mathcal{K}}(x)$ for all $x \in \mathcal{K}$.*
3. *If ψ is strictly increasing and exact, then $\mathcal{K}$ has a length function and $\lambda_{\mathcal{K}}(x) = \lambda_{\mathcal{L}}(\psi(x))$ for all $x \in \mathcal{K}$.*

Proof.

1. This follows from 3.10.1 and the fact that $\lambda_{\psi(\mathcal{K})}(y) \le \lambda_{\mathcal{L}}(y)$ for all $y \in \psi(\mathcal{K})$.
2. This follows from 3.10.2 and the fact that, since ψ is exact, $\psi(\mathcal{K})$ is a lower class in $\mathcal{L}$, and so, from 3.5, $\lambda_{\psi(\mathcal{K})}(y) = \lambda_{\mathcal{L}}(y)$ for all $y \in \psi(\mathcal{K})$.
3. Immediate from 1 and 2.

□

This proposition has the immediate corollary that if $\mathcal{K}$ is a poclass with a length function, then the only strictly increasing exact function from $\mathcal{K}$ to $\mathbf{Ord}$ is $\lambda_{\mathcal{K}}$.

Also immediate from 3.11 is

Corollary 3.12. *Let $\psi\colon \mathcal{K} \to \mathcal{L}$ be a function between poclasses with length functions and maximum elements.*

1. *If ψ is strictly increasing then* $\operatorname{len} \mathcal{K} \le \operatorname{len} \mathcal{L}$.

2. *If ψ is exact then* $\operatorname{len}\mathcal{K} \geq \operatorname{len}\psi(\mathcal{K})$.
3. *If ψ is exact and strictly increasing then* $\operatorname{len}\mathcal{K} = \operatorname{len}\psi(\mathcal{K})$.

Consider the poclass $\mathbf{Ord}\times\mathbf{Ord}$. This poclass is Artinian and for any $(\alpha,\beta) \in \mathbf{Ord}\times\mathbf{Ord}$, $\{\leq (\alpha,\beta)\} \cong \{\leq \alpha\} \times \{\leq \beta\}$ is a set. Thus, from 3.9, $\mathbf{Ord}\times\mathbf{Ord}$ has a length function. The function $\lambda_{\mathbf{Ord}\times\mathbf{Ord}}$ gives a new operation on ordinals that will be the key to the rest of this section:

Definition 3.13. *Define the operation* **natural sum**, $\oplus$, *on ordinals by*

$$\alpha \oplus \beta = \lambda_{\mathbf{Ord}\times\mathbf{Ord}}(\alpha,\beta)$$

for ordinals α and β. Note that $\alpha \oplus \beta = \operatorname{len}\{\leq (\alpha,\beta)\} = \operatorname{len}(\{\leq \alpha\} \times \{\leq \beta\})$.

The natural sum of ordinals was originally defined by G. Hessenberg [16, pages 591-594] as in Definition 3.17 (see also [28, page 363]). In 3.18, we will show that these two definitions for the natural sum are equivalent.

We will shortly prove many properties of the natural sum, but its existence is sufficient to prove the following:

Proposition 3.14. *Let $\mathcal{K}$ and $\mathcal{L}$ be poclasses.*

1. *$\mathcal{K}\times\mathcal{L}$ has a length function if and only if $\mathcal{K}$ and $\mathcal{L}$ have length functions.*
2. *If $\mathcal{K}\times\mathcal{L}$ has a length function, then $\lambda_{\mathcal{K}\times\mathcal{L}}(x,y) = \lambda_{\mathcal{K}}(x)\oplus\lambda_{\mathcal{L}}(y)$ for all $(x,y) \in \mathcal{K}\times\mathcal{L}$.*
3. *If $\mathcal{K}$ and $\mathcal{L}$ have length functions and maximum elements then*

$$\operatorname{len}(\mathcal{K}\times\mathcal{L}) = \operatorname{len}\mathcal{K} \oplus \operatorname{len}\mathcal{L}.$$

Proof. If $\mathcal{K}\times\mathcal{L}$ has a length function, then since $\mathcal{K}$ and $\mathcal{L}$ can be embedded in $\mathcal{K}\times\mathcal{L}$, they have length functions.

Conversely, if $\mathcal{K}$ and $\mathcal{L}$ have length functions, set $\lambda = \lambda_{\mathcal{K}} \times \lambda_{\mathcal{L}}$. Thus λ maps $\mathcal{K}\times\mathcal{L}$ to $\mathbf{Ord}\times\mathbf{Ord}$ and is defined by $\lambda(x,y) = (\lambda_{\mathcal{K}}(x), \lambda_{\mathcal{L}}(y))$ for all $x \in \mathcal{K}$ and $y \in \mathcal{L}$. Since $\lambda_{\mathcal{K}}$ and $\lambda_{\mathcal{L}}$ are strictly increasing and exact, so is λ.

From 3.11.3, $\mathcal{K}\times\mathcal{L}$ has a length function and

$$\lambda_{\mathcal{K}\times\mathcal{L}}(x,y) = \lambda_{\mathbf{Ord}\times\mathbf{Ord}}(\lambda(x,y)) = \lambda_{\mathcal{K}}(x) \oplus \lambda_{\mathcal{L}}(y)$$

for all $(x,y) \in \mathcal{K}\times\mathcal{L}$.

3 follows directly from 2. □

Proposition 3.15. *The natural sum is a commutative and associative operation.*

Proof. This follows from the rule $\alpha \oplus \beta = \operatorname{len}(\{\leq \alpha\} \times \{\leq \beta\})$, and the commutativity and associativity of the direct product operation on poclasses. □

Proposition 3.16. *Let $\alpha, \beta, \alpha_1, \beta_1, \ldots$ be ordinals. Then*

1. $\alpha \oplus 0 = 0 \oplus \alpha = \alpha$
2. $(\alpha_1 \oplus \beta) + \alpha_2 \leq (\alpha_1 + \alpha_2) \oplus \beta$
3. $\alpha + \beta \leq \alpha \oplus \beta, \quad \beta + \alpha \leq \alpha \oplus \beta$
4. $\alpha_1 + \beta_1 + \alpha_2 + \cdots + \alpha_n + \beta_n \leq (\alpha_1 + \alpha_2 + \cdots + \alpha_n) \oplus (\beta_1 + \beta_2 + \cdots + \beta_n)$

Proof.

1. $\alpha \oplus 0 = \operatorname{len}(\{\leq \alpha\} \times \{0\}) = \operatorname{len}\{\leq \alpha\} = \alpha$

2. Set $\mathcal{L} = \{\leq (\alpha_1 + \alpha_2, \beta)\} \subseteq \mathbf{Ord} \times \mathbf{Ord}$, and $x = (\alpha_1, \beta)$. Then $\operatorname{len}\{\leq x\} = \alpha_1 \oplus \beta$ and $\operatorname{len}[x, (\alpha_1 + \alpha_2, \beta)] = \operatorname{len}[\alpha_1, \alpha_1 + \alpha_2] = \alpha_2$. We can now apply 3.6.3 to get $\operatorname{len}\{\leq x\} + \operatorname{len}[x, (\alpha_1 + \alpha_2, \beta)] \leq \operatorname{len} \mathcal{L}$, that is, $(\alpha_1 \oplus \beta) + \alpha_2 \leq (\alpha_1 + \alpha_2) \oplus \beta$.
3. These are special cases of 2 with $\alpha_1 = 0$.
4. By induction from 2.

$\square$

Consider the natural sum of two ordinals which are given in normal form, for example, $\alpha = \omega^\omega + \omega^3 + \omega 2 + 1$ and $\beta = \omega^3 + \omega$. Using 3 of this proposition, we get the inequalities $\alpha + \beta = \omega^\omega + \omega^3 2 + \omega \leq \alpha \oplus \beta$, and $\beta + \alpha = \alpha = \omega^\omega + \omega^3 + \omega 2 + 1 \leq \alpha \oplus \beta$.

Using 4 of this proposition, we can interleave the terms of these two normal forms and add them to get a greater lower bound for $\alpha \oplus \beta$. There is a unique way of doing this so that no cancellations occur in adjacent terms of this sum, namely: Write down the terms gathered from both the normal forms in decreasing order and then add. In the example, we have the six terms $\omega^\omega, \omega^3, \omega^3, \omega 2, \omega, 1$ so

$$\omega^\omega + \omega^3 + \omega^3 + \omega 2 + \omega + 1 = \omega^\omega + \omega^3 2 + \omega 3 + 1 \leq \alpha \oplus \beta.$$

We will show that this method actually gives us the natural sum of α and β, not just a lower bound for it. But first we need to formalize this construction:

Definition 3.17. *Let α and β be nonzero ordinals. With suitable re-labeling, the normal forms for these ordinals can be written using the same strictly decreasing set of exponents $\gamma_1 > \gamma_2 > \cdots > \gamma_n$:*

$$\alpha = \omega^{\gamma_1} m_1 + \omega^{\gamma_2} m_2 + \cdots + \omega^{\gamma_n} m_n$$
$$\beta = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$$

where $n_i, m_i \in \mathbb{Z}^+$, that is we allow m_i, n_i to be zero.

Now we define the operation $\oplus'$ by

$$\alpha \oplus' \beta = \omega^{\gamma_1}(m_1 + n_1) + \omega^{\gamma_2}(m_2 + n_2) + \cdots + \omega^{\gamma_n}(m_n + n_n).$$

This is a well defined operation because of the uniqueness of the normal forms for ordinals. In addition, we define $0 \oplus' \alpha = \alpha \oplus' 0 = \alpha$, and $0 \oplus' 0 = 0$.

Proposition 3.18. *The operations $\oplus$ and $\oplus'$ are identical.*

Proof. From 3.16.4, $\alpha \oplus' \beta \leq \alpha \oplus \beta$ for any ordinals α and β. To show the opposite inequality we need only show that $\oplus'$: $\mathbf{Ord} \times \mathbf{Ord} \to \mathbf{Ord}$ is strictly increasing:

Since $\oplus'$ is commutative and increasing, it suffices to show only that $\alpha \oplus' (\beta + 1) > \alpha \oplus' \beta$. But this follows easily from the definition of $\oplus'$, in fact, $\alpha \oplus' (\beta + 1) = (\alpha \oplus' \beta) + 1$. $\square$

Corollary 3.19. *The operation $\oplus$ is cancellative, that is, $\alpha \oplus \beta = \alpha \oplus \gamma$ implies $\beta = \gamma$.*

Proof. Using the uniqueness of normal forms, it is easy to see that $\oplus'$ is cancellative, so this corollary follows from the previous proposition. $\square$

Notice that $\oplus$ is cancellative on both sides, unlike ordinary ordinal addition.

We next consider the special case when $\mathcal{L}$ is a modular lattice.

Recall from 3.6, that if $\mathcal{L}$ is a poclass with a length function and a maximum element, and $x \in \mathcal{L}$, then $\operatorname{len}\{\leq x\} + \operatorname{len}[x, \top] \leq \operatorname{len} \mathcal{L}$. In contrast to this result, there is, in general, no function of $\operatorname{len}\{\leq x\}$ and $\operatorname{len}[x, \top]$ which gives an upper bound for $\operatorname{len} \mathcal{L}$. This situation changes if $\mathcal{L}$ is a modular lattice . . .

Proposition 3.20. *Let $\mathcal{L}$ be a bounded modular lattice with a length function, and $x \in \mathcal{L}$. Then*

$$\operatorname{len}[\perp, x] + \operatorname{len}[x, \top] \leq \operatorname{len} \mathcal{L} \leq \operatorname{len}[\perp, x] \oplus \operatorname{len}[x, \top].$$

In addition, $\operatorname{len}[\perp, x] = \operatorname{len} \mathcal{L}$ *if and only if* $\operatorname{len}[x, \top] = 0$.

Proof. The first inequality and the last claim we have from 3.6. For the second inequality, we define $\lambda\colon \mathcal{L} \to [\perp, x] \times [x, \top]$ as in 2.31:

$$\lambda(y) = (x \wedge y, x \vee y)$$

for all $y \in \mathcal{L}$.

Since $\mathcal{L}$ is modular, λ is strictly increasing, and we can use 3.12.1 and 3.14.3 to get

$$\operatorname{len} \mathcal{L} \leq \operatorname{len}([\perp, x] \times [x, \top]) = \operatorname{len}[\perp, x] \oplus \operatorname{len}[x, \top].$$

□

For modular lattices, this proposition reduces many questions about the relationship between $\operatorname{len}[\perp, x]$, $\operatorname{len}[x, \top]$ and $\operatorname{len} \mathcal{L}$ to ordinal arithmetic.

The most important example of this arises from the observation that for any nonzero ordinals α and β, $\alpha + \beta$ and $\alpha \oplus \beta$ have the same leading term in their normal forms. Further this leading term depends only on the leading terms of α and β. For example, if $\alpha = \omega^\omega + \omega^3 + \omega 2 + 1$ and $\beta = \omega^3 + \omega$, then $\alpha + \beta = \omega^\omega + \omega^3 2 + \omega$ and $\alpha \oplus \beta = \omega^\omega + \omega^3 2 + \omega 3 + 1$, both having the leading term ω^ω.

This suggests that we define a map, to be called the Krull length, on ordinals which picks out the leading term of a normal form. Since 0 has no normal form we have to make a special case for it.

First we define the range of this map:

Definition 3.21. *We define* **Krull** $= (\mathbf{Ord} \times \mathbb{N}) \cup \{0\}$ *with operation* $+$ *given by*

1. $0 + 0 = 0$
2. $0 + (\gamma, n) = (\gamma, n) + 0 = (\gamma, n)$ *for all* $(\gamma, n) \in \mathbf{Ord} \times \mathbb{N}$
3. $$(\gamma_1, n_1) + (\gamma_2, n_2) = \begin{cases} (\gamma_1, n_1) & \text{if } \gamma_2 < \gamma_1 \\ (\gamma_2, n_2) & \text{if } \gamma_1 < \gamma_2 \\ (\gamma_1, n_1 + n_2) & \text{if } \gamma_1 = \gamma_2 \end{cases}$$

   for all $(\gamma_1, n_1), (\gamma_2, n_2) \in \mathbf{Ord} \times \mathbb{N}$.

In Section 5, we will consider $(\mathbf{Krull}, +)$ to be a monoid, but here we simply consider it to be a class with a binary operation.

Definition 3.22. *For an ordinal $\alpha = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$ in normal form, we define the* **Krull length** *of α by* $\operatorname{Klen}(\alpha) = (\gamma_1, n_1) \in \mathbf{Krull}$. *We also define* $\operatorname{Klen}(0) = 0$.

It will also be useful to have a name for the map which picks out the exponent of the leading term for normal forms:

Definition 3.23. *Let* $\mathbf{Ord}^* = \mathbf{Ord} \cup \{-1\}$.

We define a map κ: $\mathbf{Krull} \to \mathbf{Ord}^*$ *by* $\kappa(\gamma, n) = \gamma$ *for* $(\gamma, n) \in \mathbf{Ord} \times \mathbb{N}$, *and* $\kappa(0) = -1$.

For an ordinal α, *we define the* **Krull dimension** *of* α *by* $\mathrm{Kdim}(\alpha) = \kappa(\mathrm{Klen}(\alpha))$.

We have defined Kdim such that, for an ordinal $\alpha = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$ in normal form, we have $\mathrm{Kdim}(\alpha) = \gamma_1$. Also $\mathrm{Kdim}(0) = -1$.

It is easy to confirm the following:

Proposition 3.24. *If* α *and* β *are ordinals then*

1. $\mathrm{Klen}(\alpha + \beta) = \mathrm{Klen}(\alpha \oplus \beta) = \mathrm{Klen}(\alpha) + \mathrm{Klen}(\beta)$
2. $\mathrm{Kdim}(\alpha + \beta) = \mathrm{Kdim}(\alpha \oplus \beta) = \max\{\mathrm{Kdim}(\alpha), \mathrm{Kdim}(\beta)\}$

For a poclass with a length function and a maximum element we use these same function names as follows:

Definition 3.25. *Let* $\mathcal{L}$ *be a poclass with a length function and a maximum element. Then we define the* **Krull length** *of* $\mathcal{L}$ *by*

$$\mathrm{Klen}\,\mathcal{L} = \mathrm{Klen}(\mathrm{len}\,\mathcal{L})$$

and the **Krull dimension** *of* $\mathcal{L}$ *by*

$$\mathrm{Kdim}\,\mathcal{L} = \mathrm{Kdim}(\mathrm{len}\,\mathcal{L}).$$

Finally, we can put 3.20, 3.24 and this definition together to get

Proposition 3.26. *Let* $\mathcal{L}$ *be a bounded modular lattice with a length function and* $x \in \mathcal{L}$. *Then*

1. $\mathrm{Klen}\,\mathcal{L} = \mathrm{Klen}[\bot, x] + \mathrm{Klen}[x, \top]$
2. $\mathrm{Kdim}\,\mathcal{L} = \max\{\mathrm{Kdim}[\bot, x], \mathrm{Kdim}[x, \top]\}$

We next consider the following question: If $\mathcal{L}$ is a bounded modular lattice with a length function, what are the possible values of $\mathrm{len}[x, \top]$ for elements $x \in \mathcal{L}$? The surprising answer is that there are only a finite number of possibilities. This is to be contrasted with the possible values of $\mathrm{len}[\bot, x]$ which, by 3.6, include all ordinals $\alpha \leq \mathrm{len}\,\mathcal{L}$, and with Example 3.7 which is a non-modular lattice such that $\mathrm{len}[x, \top]$ takes on all natural numbers.

Definition 3.27. *A poclass* $\mathcal{L}$ *with a length function and a maximum element is* **critical** *if* $\mathrm{len}[x, \top] = \mathrm{len}\,\mathcal{L}$ *for all* $x < \top$.

Note that the trivial poclass with one element and length 0 is critical.

Proposition 3.28. *Let* $\mathcal{L}$ *be a poclass with a length function and a maximum element. If* $\mathcal{L}$ *is critical then* $\mathrm{len}\,\mathcal{L} = 0$ *or* $\mathrm{len}\,\mathcal{L} = \omega^\gamma$ *for some ordinal* γ. *If* $\mathcal{L}$ *is, in addition, a modular lattice then the converse is true.*

Proof. Suppose $\mathcal{L}$ is critical with $\mathrm{len}\,\mathcal{L} = \alpha > 0$. By 3.6.2, for every $\beta < \alpha$ there is some $x \in \mathcal{L}$ such that $\mathrm{len}\{\leq x\} = \beta$. By assumption, $\mathrm{len}[x, \top] = \alpha$, so from 3.6.3, $\beta + \alpha \leq \alpha$. That is, $\beta + \alpha = \alpha$ for all $\beta < \alpha$. This implies that $\alpha = \omega^\gamma$ for some $\gamma \in \mathbf{Ord}$.

Suppose $\mathcal{L}$ is a modular lattice, $\mathrm{len}\,\mathcal{L} = \omega^\gamma$ and $x < \top$. Then $\mathrm{len}[\bot, x] < \omega^\gamma$ so $\mathrm{Kdim}[\bot, x] < \gamma$. From 3.26, $\mathrm{Kdim}\,\mathcal{L} = \max\{\mathrm{Kdim}[\bot, x], \mathrm{Kdim}[x, \top]\}$ so we must have

$\operatorname{Kdim}[x,\top] = \operatorname{Kdim}\mathcal{L} = \gamma$, that is, $\operatorname{len}[x,\top] \geq \omega^\gamma$. On the other hand, since $[x,\top]$ is a sublattice of $\mathcal{L}$ we have $\operatorname{len}[x,\top] \leq \operatorname{len}\mathcal{L} = \omega^\gamma$.

□

We will say $\mathcal{L}$ is **γ-critical** if it is critical and $\operatorname{len}\mathcal{L} = \omega^\gamma$.

See Example 3.7 for a lattice $\mathcal{L}$ which is not critical even though $\operatorname{len}\mathcal{L} = \omega$.

Definition 3.29. *A* **critical series** *for a bounded poclass $\mathcal{L}$, is a sequence*

$$\bot = x_0 < x_1 < \cdots < x_n = \top$$

in $\mathcal{L}$ such that $[x_{i-1}, x_i]$ is γ_i-critical for $i = 1, 2, \ldots, n$, and $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$.

As an example we consider again the modular lattice $\mathcal{L} = \{\leq \omega\} \subseteq \mathbf{Ord}$ and the series $0 < 1 < 2 < \omega$ in $\mathcal{L}$. Each of the factors of this series is critical, $\operatorname{len}[0,1] = \operatorname{len}[1,2] = 1 = \omega^0$, $\operatorname{len}[2,\omega] = \omega$, but the series itself is not, since the factors are not decreasing in length. There are of course many other series in $\mathcal{L}$ whose factors are all critical, but there is only one critical series, namely $0 < \omega$.

In more generality, we will show that any bounded modular lattice with a length function has a critical series, and the lengths of the factors in any such series are uniquely determined by the length of the whole lattice, though the series itself is not unique.

Lemma 3.30. *Let $\mathcal{L}$ be a bounded modular lattice with a length function, and $\alpha, \beta \in$* **Ord** *such that $\alpha + \beta = \alpha \oplus \beta$.*

1. *If $\operatorname{len}\mathcal{L} = \alpha + \beta$, then there is some $x \in \mathcal{L}$ such that $\operatorname{len}[\bot, x] = \alpha$ and $\operatorname{len}[x,\top] = \beta$.*
2. *If there is $x \in \mathcal{L}$ such that $\operatorname{len}[\bot, x] = \alpha$ and $\operatorname{len}[x,\top] = \beta$ then $\operatorname{len}\mathcal{L} = \alpha + \beta$.*

Proof.

1. From 3.6, there is some $x \in \mathcal{L}$ such that $\operatorname{len}[\bot, x] = \alpha$. We will show $\operatorname{len}[x,\top] = \beta$.
   From 3.20, $\alpha + \operatorname{len}[x,\top] \leq \alpha + \beta = \alpha \oplus \beta \leq \alpha \oplus \operatorname{len}[x,\top]$. Cancellation in the first inequality gives $\operatorname{len}[x,\top] \leq \beta$. Cancellation in the second inequality gives $\beta \leq \operatorname{len}[x,\top]$.
2. This follows directly from 3.20.

□

The obvious question to ask here is: For what ordinals α and β does $\alpha + \beta = \alpha \oplus \beta$?

Lemma 3.31. *Suppose $\alpha + \beta = \alpha \oplus \beta = \omega^{\gamma_1} + \omega^{\gamma_2} + \cdots + \omega^{\gamma_n}$ in normal form where $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$. Then $\alpha = 0$ or $\beta = 0$, or there is some $i \in \{1, 2, \ldots, n-1\}$ such that $\alpha = \omega^{\gamma_1} + \omega^{\gamma_2} + \cdots + \omega^{\gamma_i}$ and $\beta = \omega^{\gamma_{i+1}} + \omega^{\gamma_{i+2}} + \cdots + \omega^{\gamma_n}$.*

Proof. Easy ordinal arithmetic. □

For example, if $\operatorname{len}\mathcal{L} = \omega^\omega + \omega^3 2 + 1$, then the previous two lemmas guarantee the existence of an $x \in \mathcal{L}$ such that $\operatorname{len}[x,\top]$ is any of the following ordinals:

$$0,\ 1,\ \omega^3 + 1,\ \omega^3 2 + 1,\ \omega^\omega + \omega^3 2 + 1.$$

More generally,

Proposition 3.32. *Suppose $\mathcal{L}$ is a bounded modular lattice with a length function, and $\operatorname{len}\mathcal{L} = \omega^{\gamma_1} + \omega^{\gamma_2} + \cdots + \omega^{\gamma_n}$ in normal form where $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$. Then for any ordinal of the form $\beta = \omega^{\gamma_i} + \omega^{\gamma_{i+1}} + \cdots + \omega^{\gamma_n}$ with $i \in \{1, 2, \ldots, n\}$, there is an element $x \in \mathcal{L}$ such that $\operatorname{len}[x,\top] = \beta$.*

Lemma 3.33. *Let $\mathcal{L}$ be a bounded modular lattice with a length function. Then the following are equivalent*

1. $\operatorname{len}\mathcal{L} = \omega^{\gamma_1} + \omega^{\gamma_2} + \cdots + \omega^{\gamma_n}$ *in normal form with* $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$.
2. $\mathcal{L}$ *has a critical series* $\perp = x_0 < x_1 < \cdots < x_n = \top$ *with* $[x_{i-1}, x_i]$ γ_i*-critical for* $i = 1, 2, \ldots, n$.

Proof. This is a simple induction from 3.30 and 3.31. □

Proposition 3.34. *Suppose $\mathcal{L}$ is a bounded modular lattice with a length function, and $\mathcal{L}$ has the critical series $\perp = x_0 < x_1 < \cdots < x_n = \top$ with $[x_{i-1}, x_i]$ γ_i-critical. Let $y \in \mathcal{L}$ and set $y_i = y \vee x_i$ for $i = 0, 1, 2, \ldots, n$. Then $\operatorname{len}[y_{i-1}, y_i]$ is ω^{γ_i} or zero. Further, the sequence $y = y_0 \leq y_2 \leq \cdots \leq y_n = \top$, after removal of duplicate entries, forms a critical series for $[y, \top]$.*

Proof. We remind the reader that for any a and b in a modular lattice, $[b, a \vee b]$ is isomorphic to $[a \wedge b, a]$. In this proof we set $a = x_i$ and $b = y_{i-1}$. Then $a \vee b = x_i \vee y_{i-1} = x_i \vee (x_{i-1} \vee y) = x_i \vee y = y_i$, and $a \wedge b = x_i \wedge y_{i-1} = x_i \wedge (x_{i-1} \vee y) = x_{i-1} \vee (x_i \wedge y)$. The last equality uses the modularity of the lattice. Thus $[y_{i-1}, y_i] \cong [x_{i-1} \vee (x_i \wedge y), x_i]$. But $x_{i-1} \leq x_{i-1} \vee (x_i \wedge y) \leq x_i$, so, $[y_{i-1}, y_i]$ is isomorphic to a final segment of $[x_{i-1}, x_i]$. Since $[x_{i-1}, x_i]$ is γ_i-critical, either $\operatorname{len}[y_{i-1}, y_i] = 0$ or $\operatorname{len}[y_{i-1}, y_i] = \omega^{\gamma_i}$.

The claim that $y = y_0 \leq y_2 \leq \cdots \leq y_n = \top$ forms a critical series for $[y, \top]$ is then clear. □

Corollary 3.35. *Suppose $\mathcal{L}$ is a bounded modular lattice with a length function, and $\operatorname{len}\mathcal{L} = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$ in normal form with $\gamma_1 > \gamma_2 > \cdots > \gamma_n$. Then for any $x \in \mathcal{L}$,*

$$\operatorname{len}[x, \top] = \omega^{\gamma_1} m_1 + \omega^{\gamma_2} m_2 + \cdots + \omega^{\gamma_n} m_n$$

for some $m_i \in \mathbb{Z}^+$ such that $m_i \leq n_i$ for all i.

Continuing with the example $\operatorname{len}\mathcal{L} = \omega^\omega + \omega^3 2 + 1 \ldots$

If $x \in \mathcal{L}$, then $\operatorname{len}[x, \top]$ is one of the following ordinals:

$$0,\ 1,\ \omega^3,\ \omega^3 + 1,\ \omega^3 2,\ \omega^3 2 + 1,\ \omega^\omega,\ \omega^\omega + 1,$$
$$\omega^\omega + \omega^3,\ \omega^\omega + \omega^3 + 1,\ \omega^\omega + \omega^3 2,\ \omega^\omega + \omega^3 2 + 1$$

It is perhaps useful to collect in one place everything we now know from 3.6, 3.30, 3.31 and 3.35 about the sublattices $[\perp, x]$ and $[x, \top]$ of a bounded modular lattice:

Proposition 3.36. *Suppose $\mathcal{L}$ is a bounded modular lattice with a length function. We write $\operatorname{len}\mathcal{L} = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$ in normal form with $\gamma_1 > \gamma_2 > \cdots > \gamma_n$.*

- *For every ordinal $\alpha \leq \operatorname{len}\mathcal{L}$ there exists some $x \in \mathcal{L}$ such that $\operatorname{len}[\perp, x] = \alpha$.*
- *For every ordinal α of the form*

$$\alpha = \omega^{\gamma_i} m_i + \omega^{\gamma_{i+1}} n_{i+1} + \cdots + \omega^{\gamma_n} n_n$$

where $1 \leq i \leq n$ and $m_i \leq n_i$, there exists some $x \in \mathcal{L}$ such that $\operatorname{len}[x, \top] = \alpha$. Notice the change of normal form from 3.32.

- *For any* $x \in \mathcal{L}$,

$$\operatorname{len}[x, \top] = \omega^{\gamma_1} m_1 + \omega^{\gamma_2} m_2 + \cdots + \omega^{\gamma_n} m_n$$

for some $m_i \in \mathbb{Z}^+$ *such that* $m_i \leq n_i$ *for all* i.

4 The Krull Length of Artinian and Noetherian Modules

In this section we apply the results of our study of length functions to Noetherian and Artinian modules. In particular, we construct the Krull length function on the module categories R-**Noeth** and R-**Art** for a ring R. This function is a generalization of both the usual composition length function on modules of finite length, and the Krull dimension function on Noetherian modules. The Krull length function is critical in proving the cancellation properties of the corresponding monoids, $M(R\text{-}\mathbf{Noeth})$ and $M(R\text{-}\mathbf{Art})$.

We will also define a Krull dimension for Artinian and Noetherian modules and show that, for Noetherian modules, this new definition coincides with the standard definition of Krull dimension given by R. Gordon and J. C. Robson [12].

Throughout this section, R will be a fixed ring. If $A \in R$-**Mod**, we will write $\mathcal{L}(A)$ for the poset of submodules of A ordered by set inclusion, and $\mathcal{L}^\circ(A)$ for the poset of submodules of A ordered by reverse set inclusion, that is, the dual of $\mathcal{L}(A)$.

$\mathcal{L}(A)$ is a modular lattice with $\top = A$ and $\perp = 0$ in which $A_1 \wedge A_2 = A_1 \cap A_2$ and $A_1 \vee A_2 = A_1 + A_2$ for all $A_1, A_2 \in \mathcal{L}(A)$. $\mathcal{L}^\circ(A)$ is a modular lattice with $\top = 0$ and $\perp = A$ in which $A_1 \wedge A_2 = A_1 + A_2$ and $A_1 \vee A_2 = A_1 \cap A_2$ for all $A_1, A_2 \in \mathcal{L}^\circ(A)$. For the details of these claims about $\mathcal{L}(A)$, see L. Rowen, *Ring Theory, Volume 1*, [27, pages 7-9].

If A is an Artinian module, then $\mathcal{L}(A)$ is an Artinian poset, so using 3.25, we can define the **length**, **Krull length** and **Krull dimension** of A by

$$\begin{aligned} \operatorname{len}_\circ A &= \operatorname{len} \mathcal{L}(A) \\ \operatorname{Klen}_\circ A &= \operatorname{Klen} \mathcal{L}(A) \\ \operatorname{Kdim}_\circ A &= \operatorname{Kdim} \mathcal{L}(A). \end{aligned}$$

Similarly, if A is a Noetherian module, then $\mathcal{L}^\circ(A)$ is an Artinian poset, and we can define the **length**, **Krull length** and **Krull dimension** of A by

$$\begin{aligned} \operatorname{len}^\circ A &= \operatorname{len} \mathcal{L}^\circ(A) \\ \operatorname{Klen}^\circ A &= \operatorname{Klen} \mathcal{L}^\circ(A) \\ \operatorname{Kdim}^\circ A &= \operatorname{Kdim} \mathcal{L}^\circ(A). \end{aligned}$$

In the Artinian case $\operatorname{len}_\circ A = 0$ if and only if $A = 0$. In the Noetherian case, $\operatorname{len}^\circ A = 0$ if and only if $A = 0$.

This duplication of nomenclature among Artinian and Noetherian modules will not lead to confusion since if A happens to be both Artinian and Noetherian, then A has finite length and its composition length coincides with both $\operatorname{len}_\circ A$ and $\operatorname{len}^\circ A$.

The key to using our information about Artinian lattices to study modules is the following: Let $0 \to A \to B \to C \to 0$ be an exact sequence in R-**Mod**, and A' the image of A in B. Then $\mathcal{L}(C) \cong [A', B] \subseteq \mathcal{L}(B)$, and $\mathcal{L}(A) \cong [0, A'] \subseteq \mathcal{L}(B)$. So, from 3.20 and 3.24, we get the main result of this section:

Proposition 4.1. *Let* $0 \to A \to B \to C \to 0$ *be an exact sequence in* R-**Mod**.

- *If* $A, B, C \in R$-**Art**, *then*

$$\operatorname{len}_\circ A + \operatorname{len}_\circ C \le \operatorname{len}_\circ B \le \operatorname{len}_\circ A \oplus \operatorname{len}_\circ C$$
$$\operatorname{Klen}_\circ B = \operatorname{Klen}_\circ A + \operatorname{Klen}_\circ C$$
$$\operatorname{Kdim}_\circ B = \max\{\operatorname{Kdim}_\circ A, \operatorname{Kdim}_\circ C\}$$

  Further, $A \cong B$ *if and only if* $\operatorname{len}_\circ C = 0$ *if and only if* $\operatorname{len}_\circ A = \operatorname{len}_\circ B$.

- [13, 2.1] *If* $A, B, C \in R$-**Noeth**, *then*

$$\operatorname{len}^\circ C + \operatorname{len}^\circ A \le \operatorname{len}^\circ B \le \operatorname{len}^\circ C \oplus \operatorname{len}^\circ A$$
$$\operatorname{Klen}^\circ B = \operatorname{Klen}^\circ A + \operatorname{Klen}^\circ C$$
$$\operatorname{Kdim}^\circ B = \max\{\operatorname{Kdim}^\circ A, \operatorname{Kdim}^\circ C\}$$

  Further, $C \cong B$ *if and only if* $\operatorname{len}^\circ A = 0$ *if and only if* $\operatorname{len}^\circ C = \operatorname{len}^\circ B$.

As a simple application of this proposition we deduce a standard property of Artinian and Noetherian modules:

Corollary 4.2. *Let* ϕ: $B \to B$ *be an module endomorphism.*

1. *If* $B \in R$-**Art** *and* ϕ *is injective, then* ϕ *is surjective.*
2. *If* $B \in R$-**Noeth** *and* ϕ *is surjective, then* ϕ *is injective.*

Proof.

1. If ϕ is injective then we get the exact sequence $0 \to B \xrightarrow{\phi} B \to \operatorname{coker}\phi \to 0$ in R-**Art**. From 4.1, $\operatorname{coker}\phi = 0$, so ϕ is surjective.
2. By duality.

□

Corollary 4.3.

1. *If* $A, B \in R$-**Art**, *then* $\operatorname{len}_\circ(A \oplus B) = \operatorname{len}_\circ A \oplus \operatorname{len}_\circ B$.
2. *If* $A, B \in R$-**Noeth**, *then* $\operatorname{len}^\circ(A \oplus B) = \operatorname{len}^\circ A \oplus \operatorname{len}^\circ B$.

Proof.

1. The canonical short exact sequence $0 \to A \to A \oplus B \to B \to 0$ gives the inequality $\operatorname{len}_\circ(A \oplus B) \le \operatorname{len}_\circ A \oplus \operatorname{len}_\circ B$. To prove the converse inequality, we consider the map ψ: $\mathcal{L}(A) \times \mathcal{L}(B) \to \mathcal{L}(A \oplus B)$ defined by $\psi(A', B') = A' + B'$. This map is strictly increasing, so we can use 3.12 and 3.14 to get

$$\begin{aligned}\operatorname{len}_\circ A \oplus \operatorname{len}_\circ B &= \operatorname{len}_\circ \mathcal{L}(A) \oplus \operatorname{len}_\circ \mathcal{L}(B)\\ &= \operatorname{len}_\circ(\mathcal{L}(A) \times \mathcal{L}(B))\\ &\le \operatorname{len}_\circ \mathcal{L}(A \oplus B)\\ &= \operatorname{len}_\circ(A \oplus B).\end{aligned}$$

2. The function ψ defined above is also strictly increasing as a map from $\mathcal{L}^\circ(A) \times \mathcal{L}^\circ(B)$ to $\mathcal{L}^\circ(A \oplus B)$. So a similar argument works for Noetherian modules.

□

Using the cancellative property of $\oplus$ (3.19) we get

Corollary 4.4. *Let $A, B, C \in R$-**Mod** such that $A \oplus B \cong A \oplus C$.*

1. *If $A, B, C \in R$-**Art**, then $\operatorname{len}_\circ B = \operatorname{len}_\circ C$.*
2. *If $A, B, C \in R$-**Noeth**, then $\operatorname{len}^\circ B = \operatorname{len}^\circ C$.*

Corollary 4.5. [13, 2.11] *If A and B are submodules of a Noetherian module then $A + B$ is a direct sum if and only if $\operatorname{len}^\circ(A + B) = \operatorname{len}^\circ A \oplus \operatorname{len}^\circ B$.*

Proof. Define $\phi\colon A \oplus B \to A + B$ by $\phi(a, b) = a + b$ for $a \in A$ and $b \in B$. This map is surjective and has kernel isomorphic to $A \cap B$. Using the exact sequence

$$0 \to A \cap B \to A \oplus B \xrightarrow{\phi} A + B \to 0$$

and 4.1, we get $A \cap B = 0$ if and only if $\operatorname{len}^\circ(A \cap B) = 0$ if and only if $\operatorname{len}^\circ(A \oplus B) = \operatorname{len}^\circ(A + B)$. Thus $A + B$ is a direct sum if and only if $\operatorname{len}^\circ(A + B) = \operatorname{len}^\circ A \oplus \operatorname{len}^\circ B$. □

We next rewrite 3.36 as it applies to Artinian and Noetherian modules:

Proposition 4.6. *Let A be an Artinian module and write*

$$\operatorname{len}_\circ A = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$$

in normal form with $\gamma_1 > \gamma_2 > \cdots > \gamma_n$.

1. *For every ordinal $\alpha \leq \operatorname{len}_\circ A$ there exists a submodule $A' \leq A$ such that $\operatorname{len}_\circ A' = \alpha$.*
2. *For every ordinal α of the form*
$$\alpha = \omega^{\gamma_i} m_i + \omega^{\gamma_{i+1}} n_{i+1} + \cdots + \omega^{\gamma_n} n_n$$
where $1 \leq i \leq n$ and $m_i \leq n_i$, there exists a submodule $A' \leq A$ with $\operatorname{len}_\circ(A/A') = \alpha$.
3. *For any submodule $A' \leq A$,*
$$\operatorname{len}_\circ(A/A') = \omega^{\gamma_1} m_1 + \omega^{\gamma_2} m_2 + \cdots + \omega^{\gamma_n} m_n$$
for some $m_i \in \mathbb{Z}^+$ such that $m_i \leq n_i$ for all i.

Proposition 4.7. *Let A be a Noetherian module and write*

$$\operatorname{len}^\circ A = \omega^{\gamma_1} n_1 + \omega^{\gamma_2} n_2 + \cdots + \omega^{\gamma_n} n_n$$

in normal form with $\gamma_1 > \gamma_2 > \cdots > \gamma_n$.

1. *For every ordinal $\alpha \leq \operatorname{len}^\circ A$ there exists a submodule $A' \leq A$ with $\operatorname{len}^\circ(A/A') = \alpha$.*
2. *For every ordinal α of the form*
$$\alpha = \omega^{\gamma_i} m_i + \omega^{\gamma_{i+1}} n_{i+1} + \cdots + \omega^{\gamma_n} n_n$$
where $1 \leq i \leq n$ and $m_i \leq n_i$, there exists a submodule $A' \leq A$ such that $\operatorname{len}^\circ A' = \alpha$.
3. *For any submodule $A' \leq A$,*
$$\operatorname{len}^\circ A' = \omega^{\gamma_1} m_1 + \omega^{\gamma_2} m_2 + \cdots + \omega^{\gamma_n} m_n$$
for some $m_i \in \mathbb{Z}^+$ such that $m_i \leq n_i$ for all i.

For the remainder of this section we will discuss only Noetherian modules. The reason for this is that we need to show that, for Noetherian modules, $\operatorname{Kdim}^\circ$ is the same as the Krull dimension in the sense of Gordon and Robson [12].

Definition 4.8. *The* **Krull dimension (in the sense of Gordon and Robson)** [12], [11, Chapter 13]*, of a module $A \in R$-**Mod**, which we will denote by* $\operatorname{Kdim} A \in \mathbf{Ord}^*$*, is defined inductively as follows:*

- $\operatorname{Kdim} A = -1$ *if and only if* $A = 0$.
- *Consider an ordinal γ. Assume that we have defined which modules have* Kdim *equal to δ for every $\delta < \gamma$. Then* $\operatorname{Kdim} A = \gamma$ *if and only if*
  (a) *A does not have* Kdim *less than γ, and*
  (b) *for every countable decreasing chain $A_1 \geq A_2 \geq \ldots$ of submodules of A,* $\operatorname{Kdim}(A_i/A_{i+1}) < \gamma$ *for all but finitely many indices.*

This definition does not provide a Kdim for all modules. However, any Noetherian module has a Kdim. See [11, 13.3].

Lemma 4.9. *Let $A \in R$-**Noeth** with* $\operatorname{Kdim}^\circ A = \gamma$*. Then for any ordinal $\delta < \gamma$, there is a decreasing sequence $A_1 \geq A_2 \geq \ldots$ in A such that* $\operatorname{Kdim}^\circ(A_{i-1}/A_i) = \delta$ *for all i.*

Proof. Set $A_1 = A$. Since $\operatorname{Kdim}^\circ A_1 = \gamma$, we have $\operatorname{len}^\circ A_1 \geq \omega^\gamma > \omega^\delta$. Thus, by 4.7.1, there is some submodule $A_2 \leq A_1$ such that $\operatorname{len}^\circ(A_1/A_2) = \omega^\delta$. We have $\operatorname{Kdim}^\circ(A_1/A_2) = \delta$ and $\operatorname{Kdim}^\circ A_1 = \max\{\operatorname{Kdim}^\circ A_2, \operatorname{Kdim}^\circ(A_1/A_2)\}$, so $\operatorname{Kdim}^\circ A_2 = \gamma$, and we can repeat the process to get $A_3, A_4 \ldots$ as required. □

In the next lemma we use the easily proved fact that if $\alpha, \beta \in \mathbf{Ord}$ with $\alpha > 0$, then $\beta + \alpha \leq \alpha$ if and only if $\beta + \alpha = \alpha$ if and only if $\operatorname{Kdim} \beta < \operatorname{Kdim} \alpha$.

Lemma 4.10. *Let $A_1 \geq A_2 \geq \ldots$ be a decreasing sequence of submodules of a nonzero Noetherian module A with* $\operatorname{Kdim}^\circ A = \gamma$*. Then* $\operatorname{Kdim}^\circ(A_{i-1}/A_i) < \gamma$ *for all but a finite number of indices.*

Proof. Since the sequence of ordinals $\operatorname{len}^\circ A_1 \geq \operatorname{len}^\circ A_2 \geq \operatorname{len}^\circ A_3 \geq \ldots$ is decreasing, there is some $n \in \mathbb{N}$ such that $\operatorname{len}^\circ A_i = \operatorname{len}^\circ A_n$ for all $i \geq n$. If A_n is the zero module, then $\operatorname{Kdim}^\circ(A_{i-1}/A_i) = -1 < \gamma$ for all $i > n$ and we are done.

Otherwise, if A_n is nonzero, we have $\operatorname{len}^\circ A_n > 0$. Suppose $i > n$. Then, using the short exact sequence $0 \to A_i \to A_n \to A_n/A_i \to 0$, we have $\operatorname{len}^\circ(A_n/A_i) + \operatorname{len}^\circ A_n = \operatorname{len}^\circ(A_n/A_i) + \operatorname{len}^\circ A_i \leq \operatorname{len}^\circ A_n$. Hence $\operatorname{Kdim}^\circ(A_n/A_i) < \operatorname{Kdim}^\circ A_n$, and, since also $A_{i-1}/A_i \leq A_n/A_i$ and $A_n \leq A$, we get $\operatorname{Kdim}^\circ(A_{i-1}/A_i) < \operatorname{Kdim}^\circ A = \gamma$. □

Proposition 4.11. *For all $A \in R$-**Noeth**,* $\operatorname{Kdim}^\circ A = \operatorname{Kdim} A$*.*

Proof. Suppose the claim is not true. Among all modules in R-**Noeth** which serve as counterexamples, let A have minimum $\operatorname{Kdim}^\circ$. Set $\gamma = \operatorname{Kdim}^\circ A$. Then $\gamma > -1$ and for any module B with $\operatorname{Kdim}^\circ B < \gamma$ we have $\operatorname{Kdim}^\circ B = \operatorname{Kdim} B$.

Using Lemma 4.9, we see that, for any $\delta < \gamma$, the module A fails part (b) of the definition of having Kdim equal to δ. Thus A does not have Kdim less than γ, and A satisfies part (a) of the definition of having $\operatorname{Kdim} A = \gamma$.

Also, by Lemma 4.10, A satisfies part (b) of this definition. Therefore $\operatorname{Kdim} A = \gamma$, and A is not a counterexample. □

We should point out that if A is an Artinian module, then $\operatorname{Kdim} A = 0$. Therefore $\operatorname{Kdim} A = \operatorname{Kdim}_\circ A$ if and only if A has finite length.

5 Commutative Monoids

In this section we will present the basic definitions and properties of commutative monoids and semigroups. The basic references for the semigroup literature are A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups, Vol. I* [7], and J. M. Howie, *Fundamentals of Semigroup Theory* [18]. In these books, semigroup operations are written multiplicatively. In our discussion, we will use additive notation since we are only interested in commutative semigroups. Thus, to the extent that it does not conflict with a pre-existing usage, we will write $+$ for all semigroup/monoid operations and 0 for the identity element of monoids.

The other difference from these references, is that the semigroups and monoids we will discuss need not be sets. For example, the class of cardinal numbers with cardinal addition is a monoid by our definition. This situation is forced upon us by our interest in monoids which are derived from categories of modules, and the fact that such categories are often proper classes. See, in particular, 16.13.

Definition 5.1.

- *A* **semigroup** $(M, +)$ *is a nonempty class* M *with a binary operation* $+$ *which is associative. A* **commutative semigroup** *is a semigroup* $(M, +)$ *in which the operation* $+$ *is commutative.*
- *An element* 0 *of a semigroup* $(M, +)$ *is an* **identity** *if* $a = 0 + a = a + 0$ *for all* $a \in M$*. If a semigroup has an identity, it is unique.*
- *An element* ∞ *of a semigroup* $(M, +)$ *is called* **infinite** *if* $\infty + a = \infty$ *for all* $a \in M$*. If a semigroup has an infinite element, then it is unique.*
- *A* **monoid** $(M, +, 0)$ *is a semigroup* $(M, +)$ *with identity* 0*. A* **commutative monoid** *is a monoid* $(M, +, 0)$ *in which the operation* $+$ *is commutative.*
- *A* **subsemigroup** $(I, +)$ *of a monoid* $(M, +)$ *is a nonempty subclass* $I \subseteq M$ *which is itself a semigroup with the same operation as in* M*. That is,* I *is closed under the operation* $+$*. A subsemigroup of a commutative semigroup is commutative.*
- *A* **submonoid** $(I, +, 0)$ *of a monoid* $(M, +, 0)$ *is a nonempty subclass* $I \subseteq M$ *which is itself a monoid with the same identity and operation as in* M*. That is,* I *is closed under the operation* $+$ *and contains* 0*. A submonoid of a commutative monoid is commutative.*

From this point on, all semigroups and monoids will be assumed to be commutative. As is usual, we will write M rather than $(M, +)$ or $(M, +, 0)$ if the operation and identity are clear from the context.

Among the most frequently used monoids and semigroups are

- $(\mathbb{N}, +)$ – the natural numbers with addition.
- $(\mathbb{Z}^+, +)$ – the nonnegative numbers with addition.
- $(\mathbb{Z}, +)$ – the integers with addition.

By default, we will assume that $\mathbb{N}$, $\mathbb{Z}^+$ and $\mathbb{Z}$ are semigroups with addition as operation. Any of these sets could also be a semigroup with the operations multiplication, $\max\{\cdot,\cdot\}$ or $\min\{\cdot,\cdot\}$.

We also will use the following monoids:

- $\mathbb{Z}_n$ – the cyclic group of order $n \in \{2, 3, 4, \ldots\}$.
- $\{0, \infty\}$ – the two element monoid such that 0 is an identity and ∞ is an infinite element.

Other notational conventions for special monoids and semigroups can be found in the introduction.

Given a semigroup M, we can construct a monoid M^0 by adjoining a new element 0 to M such that $0 + a = a$ for all $a \in M^0$. Similarly, we can construct a semigroup M^∞ with an infinite element by adjoining a new element ∞ to M such that $\infty + a = \infty$ for all $a \in M^\infty$.

Any semigroup M has a module-like action of $\mathbb{N}$ on its elements given by $na = \sum_{i=1}^n a$ for any $n \in \mathbb{N}$ and $a \in M$. Thus $1a = a$, $2a = a + a$, $3a = a + a + a$, etc. If M is a monoid then we will extend this to a $\mathbb{Z}^+$ action by defining $0a = 0$ for all $a \in M$.

Proposition 5.2. *Let M be a semigroup (monoid).*

1. *If $\{I_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of subsemigroups (submonoids) of M, then $I = \cap_{\alpha\in\mathcal{I}} I_\alpha$, if nonempty, is a subsemigroup (submonoid) of M.*
2. *If $\{I_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of subsemigroups (submonoids) of M which is totally ordered by inclusion, then $I = \cup_{\alpha\in\mathcal{I}} I_\alpha$ is a subsemigroup (submonoid) of M.*

Proof. Routine. □

Let Y be a nonempty subclass of a semigroup M. Using 1 of this proposition, we define the subsemigroup generated by Y to be the intersection of all subsemigroups which contain Y. For example, the subsemigroup generated by a single element $a \in M$ is the set $\{a, 2a, 3a, \ldots\}$. The elements of the sequence $(a, 2a, 3a, \ldots)$ may not be distinct, so the subsemigroup generated by a could be finite. See [18, Section 1.2] for a discussion of all possible semigroups generated by a single element. In the general case, the subsemigroup generated by Y consists all finite sums of elements of Y.

Similarly, if M is a monoid, then we define the submonoid generated by a subclass $Y \subseteq M$ to be the intersection of all submonoids which contain Y. The submonoid generated by a single element $a \in M$ is $\{0, a, 2a, 3a, \ldots\}$. In general, the submonoid generated by Y consists of 0 and all finite sums of elements of Y.

If $\mathcal{K} = \{I_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of submonoids then we will write

$$\sum_{\alpha\in\mathcal{I}} I_\alpha \text{ or } \sum \mathcal{K}$$

for the submonoid generated by the union $\cup_{\alpha\in\mathcal{I}} I_\alpha = \cup\mathcal{K}$. If $\mathcal{I} = \{1, 2, \ldots, n\}$, we will use the alternative notation

$$\sum_{i=1}^n I_i = I_1 + I_2 + \ldots + I_n.$$

Since I_i is a monoid for $i = 1, 2, \ldots, n$, we have

$$I_1 + I_2 + \ldots + I_n = \{a_1 + a_2 + \ldots + a_n \mid a_i \in I_i \text{ for } i = 1, 2, \ldots, n\}.$$

Definition 5.3.

1. *A map* $\phi\colon M \to N$ *between semigroups is a* **semigroup homomorphism** *if*
$$(\forall a, b \in M)\ (\phi(a+b) = \phi(a) + \phi(b))$$
2. *A map* $\phi\colon M \to N$ *between monoids is a* **monoid homomorphism** *if it is a semigroup homomorphism and* $\phi(0) = 0$*.*
3. *Two semigroups,* M *and* N*, are* **isomorphic***, written* $M \cong N$*, if there is a bijective semigroup homomorphism* $\phi\colon M \to N$*. In this case,* ϕ *is called an* **isomorphism** *and its inverse,* ϕ^{-1}*, is also a bijective semigroup homomorphism.*

If M and N are monoids which are isomorphic as semigroups as above, then ϕ and ϕ^{-1} are necessarily monoid homomorphisms. So it is not necessary to define two types of isomorphisms, one for semigroups and one for monoids.

The inclusion map of a subsemigroup (submonoid) in a semigroup (monoid) is a semigroup (monoid) homomorphism.

If $\phi\colon M \to N$ is a semigroup or monoid homomorphism, then we will write $\operatorname{im}\phi = \phi(M)$ for the image. It is easy to see that $\operatorname{im}\phi$ is a subsemigroup of N.

If ϕ is a monoid homomorphism then we will write $\ker\phi = \{m \in M \mid \phi(m) = 0\}$ for the kernel of ϕ. Also easy to see is that $\ker\phi$ is a submonoid of M and $\operatorname{im}\phi$ is a submonoid of N.

Notice that these homomorphisms respect the $\mathbb{N}$ and $\mathbb{Z}^+$ actions on semigroups and monoids respectively. Specifically, if $\phi\colon M \to N$ is a semigroup homomorphism, then $\phi(na) = n\phi(a)$ for all $a \in M$ and $n \in \mathbb{N}$, and if $\phi\colon M \to N$ is a monoid homomorphism, then $\phi(na) = n\phi(a)$ for all $a \in M$ and $n \in \mathbb{Z}^+$.

Definition 5.4. *A* **congruence** *on a semigroup* M *is an equivalence relation* $\sim$ *on* M *such that*
$$(\forall a, b, c \in M)\ (a \sim b \implies a + c \sim b + c).$$
We will write $[a] = \{b \in M \mid b \sim a\}$ *for the congruence class containing* a*, and* $M/\!\sim$ *for the class of congruence classes. If we have several congruences,* $\sim_I$*,* $\sim_\alpha$*, etc., to consider simultaneously, then we will write* $[a]_I$*,* $[a]_\alpha$*, etc., for the corresponding congruence classes.*

If $\sim$ is a congruence on a semigroup M, then $M/\sim$ is a semigroup when given the operation $+$ defined by
$$[a] + [b] = [a + b]$$
for all $a, b \in M$. The map $\sigma\colon M \to M/\!\sim$ defined by $\sigma(a) = [a]$ is a semigroup homomorphism. We will call $M/\!\sim$ the **quotient** of M by $\sim$, and σ the **quotient homomorphism** from M to $M/\!\sim$.

If M is a monoid, then $M/\!\sim$ is a monoid with identity $[0]$ and σ is a monoid homomorphism.

The above construction provides a homomorphism from M for every congruence on M. In the other direction, one easily checks that if $\phi\colon M \to N$ is any semigroup or monoid homomorphism, then the equivalence $\sim$ defined by
$$a \sim b \iff \phi(a) = \phi(b)$$

is a congruence on M, and $(M/\sim) \cong \mathrm{im}(\phi)$. Thus we see that there is a bijection between congruences on M and homomorphic images of M (up to isomorphism).

The class of congruences on a semigroup or monoid has a natural partial order: If $\sim_\alpha$ and $\sim_\beta$ are two congruences on a semigroup or monoid M, then we write $\sim_\alpha \subseteq \sim_\beta$ if

$$(\forall a, b \in M)\ (a \sim_\alpha b \implies a \sim_\beta b)$$

or, equivalently, if

$$\{(a,b) \in M \times M \mid a \sim_\alpha b\} \subseteq \{(a,b) \in M \times M \mid a \sim_\beta b\}.$$

If $\sim_\alpha \subseteq \sim_\beta$ there is a surjective (semigroup or monoid) homomorphism from $M/\sim_\alpha$ to $M/\sim_\beta$ given by $[a]_\alpha \mapsto [a]_\beta$ for all $a \in M$.

The main property of this partial order that we will use is that any family of congruences has an infimum: Let $\{\sim_\alpha \mid \alpha \in \mathcal{I}\}$ be a family of congruences on M for some index class $\mathcal{I}$. Define the relation $\sim$ by

$$a \sim b \iff (\forall \alpha \in \mathcal{I})\ (a \sim_\alpha b)$$

for all $a, b \in M$. This relation is easily seen to be a congruence and the infimum of $\{\sim_\alpha \mid \alpha \in \mathcal{I}\}$ with respect to the partial order described above.

The existence of infima of families of congruences allows us to define congruences using generators. This is best explained with an example: Consider the congruence $\sim$ on $\mathbb{Z}^+$ generated by $3 \sim 4$. By this statement we mean that the congruence $\sim$ is the infimum of all congruences $\sim_\alpha$ on $\mathbb{Z}^+$ such that $3 \sim_\alpha 4$. A calculation shows that we get $\sim$-congruence classes $\{0\}$, $\{1\}$, $\{2\}$ and $\{3, 4, 5, \ldots\}$, so $(\mathbb{Z}^+/\sim) = \{[0], [1], [2], [3]\}$ with $[3]$ infinite.

If we had been discussing groups, rings or modules, we would have that the image of a homomorphism ϕ: $M \to N$ is isomorphic to $M/\ker\phi$. Even though monoid homomorphisms have kernels, there is no similar rule for monoid homomorphisms. Nonetheless, given a monoid M and homomorphism ϕ, we construct below a monoid, to be called $M/\ker\phi$, which is as large as possible among homomorphic images of M with $\ker\phi$ as kernel.

Definition 5.5. *Let I be a submonoid of a monoid M.*

1. *Define a congruence $\sim_I$ on M by*

$$m \sim_I m' \iff (\exists\, a, a' \in I \text{ such that } m + a = m' + a').$$

We will write M/I for $M/\sim_I$, the quotient monoid. Elements of M/I are written in the form $[m]_I^M$, or $[m]_I$. See the notation convention following 5.7.

2. *We will say I is* **normal** *in M if*

$$(\forall a \in I)\ (\forall m \in M)\ (a + m \in I \implies m \in I).$$

One readily checks that if σ: $M \to M/I$ is the quotient homomorphism, then $I \subseteq \ker(\sigma)$ with equality if and only if I is normal. In fact, I is the kernel of a monoid homomorphism from M if and only if it is normal. Most submonoids we will discuss in subsequent sections will be normal, indeed they have the stronger property of being order ideals. See 6.12.

Quotients of monoids by submonoids are useful because of their universal property:

Proposition 5.6. *Let $\phi\colon M \to N$ be a monoid homomorphism and I be a submonoid of M such that $I \subseteq \ker(\phi)$. Then ϕ factors uniquely through M/I, that is, there exists a unique monoid homomorphism $\hat{\phi}\colon M/I \to N$ such that the following diagram commutes:*

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[d, "\sigma"'] & N \\ M/I \arrow[ur, "\hat{\phi}"'] & \end{tikzcd}$$

Here $\sigma\colon M \to M/I$ is the quotient homomorphism.

Proof. Straight forward. □

By this proposition, M/I is, in a sense, the "largest" homomorphic image of M with kernel containing I.

This universal property of quotient monoids suffices to allow us to prove theorems which are similar to those of modules, so long as we consider only those homomorphisms which arise from taking quotients by submonoids:

Suppose we have submonoids $A \subseteq B \subseteq C$. From these we get quotient monoids B/A, C/A and C/B. We have, for $b \in B$,

$$[b]_A^C = \{b' \in C \mid \exists\, a, a' \in A \text{ such that } b + a = b' + a'\} \in C/A$$

and

$$[b]_A^B = \{b' \in B \mid \exists\, a, a' \in A \text{ such that } b + a = b' + a'\} \in B/A.$$

and, as we will see in Example 5.9, $[b]_A^C$ and $[b]_A^B$ can be distinct. One can easily check that if B is normal, then $[b]_A^C = [b]_A^B$. Even without B being normal, we still get monoid analogs of some standard homomorphism theorems for modules:

Proposition 5.7. *Let $A \subseteq B \subseteq C$ be submonoids of a monoid M. Then the map $\psi_1\colon B/A \to C/A$ given by $\psi_1([b]_A^B) = [b]_A^C$ is an injective monoid homomorphism. Thus B/A embeds in C/A.*

Proof. Consider the following diagram where σ_1 and σ_2 are the quotient homomorphisms and ι is inclusion.

$$\begin{tikzcd} B \arrow[r, "\iota"] \arrow[d, "\sigma_1"] & C \arrow[d, "\sigma_2"] \\ B/A \arrow[r, "\psi_1"] & C/A \end{tikzcd}$$

Since $A \subseteq \ker(\sigma_2\iota)$, there is a unique homomorphism, ψ_1, making the diagram commute (that is, $\psi_1([b]_A^B) = [b]_A^C$ for all $b \in B$). It is then a simple task to show that ψ_1 is injective. □

This proposition allows us to identify B/A with its image in C/A whenever appropriate. In making this identification, we are ignoring the distinction between $[b]_A^B \in B/A$ and $[b]_A^C \in C/A$. With this abuse of notation we no longer need to put superscripts on these elements, that is, we will write $[b]_A$ for both $[b]_A^B$ and $[b]_A^C$.

Since we can now consider B/A as a submonoid of C/A, we can discuss the relationship between C/B and $(C/A)/(B/A)$. The next proposition shows that these two monoids are in fact isomorphic.

Proposition 5.8. *Let $A \subseteq B \subseteq C$ be submonoids of a monoid M. Then the map $\psi_3\colon (C/A)/(B/A) \to C/B$ given by $\psi_3([[c]_A]_{B/A}) = [c]_B$ for all $c \in C$ is an isomorphism.*

Proof. Consider the following extension of the diagram from 5.7:

$$\begin{tikzcd}
B \arrow[r, "\iota"] \arrow[d, "\sigma_1"] & C \arrow[r, "\sigma_3"] \arrow[d, "\sigma_2"] & C/B \arrow[d, bend left, "\psi_4"] \\
B/A \arrow[r, "\psi_1"] & C/A \arrow[r, "\sigma_4"] \arrow[ur, "\psi_2"] & (C/A)/(B/A) \arrow[u, bend left, "\psi_3"]
\end{tikzcd}$$

Here $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ are all quotient homomorphisms, and ψ_1 is as in 5.7. We have $\sigma_3\iota = 0$ and $\sigma_4\psi_1 = 0$.

Now $A \subseteq B \subseteq \ker(\sigma_3)$ so there is a unique homomorphism $\psi_2\colon C/A \to C/B$ such that $\psi_2\sigma_2 = \sigma_3$. B/A (as a submonoid of C/A) is in the kernel of ψ_2 because $\psi_2\psi_1\sigma_1 = \psi_2\sigma_2\iota = \sigma_3\iota = 0$. So there is a unique homomorphism $\psi_3\colon (C/A)/(B/A) \to C/B$ such that $\psi_2 = \psi_3\sigma_4$. We note that, from the commutativity of the diagram, ψ_3 is given by $\psi_3([[c]_A]_{B/A}) = [c]_B$ for all $c \in C$.

Also, $B \subseteq \ker(\sigma_4\sigma_2)$ since $\sigma_4\sigma_2\iota = \sigma_4\psi_1\sigma_1 = 0$, so there is a unique homomorphism $\psi_4\colon C/B \to (C/A)/(B/A)$ such that $\psi_4\sigma_3 = \sigma_4\sigma_2$.

Finally, we will show that ψ_3 and ψ_4 are inverses of each other...

We have $\psi_3\psi_4\sigma_3 = \psi_3\sigma_4\sigma_2 = \psi_2\sigma_2 = \sigma_3$. But σ_3 is surjective, so $\psi_3\psi_4$ is the identity on C/B. Similarly, $\psi_4\psi_3\sigma_4\sigma_2 = \psi_4\psi_2\sigma_2 = \psi_4\sigma_3 = \sigma_4\sigma_2$ and $\sigma_4\sigma_2$ is surjective so $\psi_4\psi_3$ is the identity on $(C/A)/(B/A)$. $\square$

As an example of the preceding two propositions we have

Example 5.9. *Let $C = \mathbb{Z}^+$, $B = \{0, 6, 8, 10, \ldots\} = \{0\} \cup \{6 + 2n \mid n \in \mathbb{Z}^+\} \subseteq C$ and $A = \{0, 6, 12, \ldots\} = \{6n \mid n \in \mathbb{Z}^+\} \subseteq B$. One readily sees that A is a normal submonoid of C and B is a submonoid of C which is not normal. We will temporarily readopt the notation which includes superscripts. A calculation then shows*

$$[0]^C_A = \{0, 6, 12, \ldots\} \qquad [1]^C_A = \{1, 7, 13, \ldots\} \qquad [2]^C_A = \{2, 8, 14, \ldots\}$$
$$[3]^C_A = \{3, 9, 15, \ldots\} \qquad [4]^C_A = \{4, 10, 16, \ldots\} \qquad [5]^C_A = \{5, 11, 17, \ldots\}$$

so $C/A = \{[0]^C_A,\ [1]^C_A,\ [2]^C_A,\ [3]^C_A,\ [4]^C_A,\ [5]^C_A\} \cong \mathbb{Z}_6$. Also

$$[0]^C_B = \{0, 2, 4, 6, \ldots\} \qquad [1]^C_B = \{1, 3, 5, 7, \ldots\}$$

so $C/B = \{[0]^C_B,\ [1]^C_B\} \cong \mathbb{Z}_2$,

$$[0]^B_A = \{0, 6, 12, \ldots\} \qquad [8]^B_A = \{8, 14, 20, \ldots\} \qquad [10]^B_A = \{10, 16, 22, \ldots\}$$

so $B/A = \{[0]^B_A,\ [8]^B_A,\ [10]^A_B\} \cong \mathbb{Z}_3$. Notice that $[8]^B_A \neq [8]^C_A = [2]^C_A$.

Let $\psi_1\colon B/A \to C/A$ be as in 5.7, then $\psi_1([0]^B_A) = [0]^C_A$, $\psi_1([8]^B_A) = [2]^C_A$ and $\psi_1([10]^B_A) = [4]^C_A$, so ψ_1 embeds B/A in C/A. Further, identifying B/A with its image in C/A, we get $(C/A)/(B/A) \cong \mathbb{Z}_2 \cong C/B$ as expected.

Proposition 5.10. *Let A, B and C be submonoids of a monoid M with $A, B \subseteq C$, and $\sigma\colon C \to C/A$ the quotient homomorphism. Then*

$$\sigma(B) = (A+B)/A.$$

Proof. Since $A \subseteq A+B \subseteq C$, we get a commutative diagram as in 5.7:

$$\begin{tikzcd} A+B \arrow[r, "\iota"] \arrow[d, "\sigma_1"] & C \arrow[d, "\sigma"] \\ (A+B)/A \arrow[r, "\psi"] & C/A \end{tikzcd}$$

with σ_1 and σ quotient homomorphisms, and ψ injective. Since σ_1 is surjective, we have $\psi((A+B)/A) = \sigma\iota(A+B) = \sigma(A+B)$. But $\sigma(A+B) = \sigma(B)$, so with the identification of $(A+B)/A$ with $\psi((A+B)/A)$, we get $\sigma(B) = (A+B)/A$. □

We next consider direct products of families of semigroups or monoids:

Given a family of semigroups $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ for some index class $\mathcal{I}$, we can form the Cartesian product $\prod_{\alpha\in\mathcal{I}} M_\alpha$ with projections $\pi_\alpha\colon \prod_{\alpha\in\mathcal{I}} M_\alpha \to M_\alpha$ for each $\alpha \in \mathcal{I}$. We will write elements of $\prod_{\alpha\in\mathcal{I}} M_\alpha$ in the form $(a_\alpha)_{\alpha\in\mathcal{I}}$ with $a_\alpha \in M_\alpha$ for each $\alpha \in \mathcal{I}$. We can then define an operation $+$ on $\prod_{\alpha\in\mathcal{I}} M_\alpha$ by

$$(a_\alpha)_{\alpha\in\mathcal{I}} + (b_\alpha)_{\alpha\in\mathcal{I}} = (a_\alpha + b_\alpha)_{\alpha\in\mathcal{I}}.$$

With this operation $\prod_{\alpha\in\mathcal{I}} M_\alpha$ is a semigroup and the projections are homomorphisms. Further, $\prod_{\alpha\in\mathcal{I}} M_\alpha$ has the universal property one expects for a direct product of semigroups:

Proposition 5.11. *Let $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ be a family of semigroups and $\prod_{\alpha\in\mathcal{I}} M_\alpha$ as above. Suppose there are homomorphisms $\phi_\alpha\colon N \to M_\alpha$ for some semigroup N and all $\alpha \in \mathcal{I}$. Then there is a unique homomorphism $\phi\colon N \to \prod_{\alpha\in\mathcal{I}} M_\alpha$ such that for each $\alpha \in \mathcal{I}$ the following diagram commutes:*

$$\begin{tikzcd} N \arrow[r, "\phi"] \arrow[dr, "\phi_\alpha"'] & \prod_{\alpha\in\mathcal{I}} M_\alpha \arrow[d, "\pi_\alpha"] \\ & M_\alpha \end{tikzcd}$$

Proof. Since $\prod_{\alpha\in\mathcal{I}} M_\alpha$ is the Cartesian product of $\{M_\alpha \mid \alpha \in \mathcal{I}\}$, there is a unique <u>function</u> $\phi\colon N \to \prod_{\alpha\in\mathcal{I}} M_\alpha$ making the diagrams commute. It remains to check only that ϕ is a homomorphism. □

If M_α is a monoid for each $\alpha \in \mathcal{I}$, then $\prod_{\alpha\in\mathcal{I}} M_\alpha$ is a monoid with identity $(0_\alpha)_{\alpha\in\mathcal{I}}$ where 0_α is the identity element of M_α. The proposition then remains true if N is a monoid and ϕ_α is a monoid homomorphism for each $\alpha \in \mathcal{I}$, and provides a unique monoid homomorphism ϕ which makes the diagrams above commute.

Following the conventions of category theory, we will call $\prod_{\alpha\in\mathcal{I}} M_\alpha$, or any isomorphic semigroup, "the" **direct product** of $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ without specifying explicitly the projection homomorphisms. If $\mathcal{I} = \{1, 2, \ldots, n\}$, we will write $M_1 \times M_2 \times \ldots \times M_n$ for the direct product.

Within the category of monoids and monoid homomorphisms, we can construct a direct sum for any family of monoids. In defining this direct sum we follow the pattern used in defining the direct sum for modules...

Let $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ be a family of monoids. If $a = (a_\alpha)_{\alpha\in\mathcal{I}} \in \prod_{\alpha\in\mathcal{I}} M_\alpha$, then we will say $X \subseteq \mathcal{I}$ **supports** a if, for all $\alpha \in \mathcal{I}$, $a_\alpha \neq 0$ implies $\alpha \in X$. We will say a has **finite support** if there is a finite set X which supports a. Define also the submonoid

$$\bigoplus_{\alpha\in\mathcal{I}} M_\alpha = \{a \in \prod_{\alpha\in\mathcal{I}} M_\alpha \mid a \text{ has finite support}\},$$

and monoid homomorphisms $\iota_\alpha\colon M_\alpha \to \bigoplus_{\alpha\in\mathcal{I}} M_\alpha$ for each $\alpha \in \mathcal{I}$ by $\iota_\alpha(a) = (a_\beta)_{\beta\in\mathcal{I}}$ where

$$a_\beta = \begin{cases} 0_\beta & \beta \neq \alpha \\ a & \beta = \alpha \end{cases}$$

Note that $\pi_\alpha \circ \iota_\alpha$ is the identity homomorphism on M_α for all $\alpha \in \mathcal{I}$, and if $a \in \bigoplus_{\alpha\in\mathcal{I}} M_\alpha$ with finite support X, then

$$a = \sum_{\alpha\in X} \iota_\alpha(\pi_a(a)).$$

Proposition 5.12. *Let $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ be a family of monoids and $\bigoplus_{\alpha\in\mathcal{I}} M_\alpha$ as above. Suppose there are monoid homomorphisms $\phi_\alpha\colon M_\alpha \to N$ for some monoid N and all $\alpha \in \mathcal{I}$. Then there is a unique monoid homomorphism $\phi\colon \bigoplus_{\alpha\in\mathcal{I}} M_\alpha \to N$ such that for each $\alpha \in \mathcal{I}$ the following diagram commutes:*

$$\begin{tikzcd}
\bigoplus_{\alpha\in\mathcal{I}} M_\alpha \arrow[r, "\phi"] & N \\
M_\alpha \arrow[u, "\iota_\alpha"] \arrow[ur, "\phi_\alpha"'] &
\end{tikzcd}$$

Proof. For $a = (a_\alpha)_{\alpha\in\mathcal{I}} \in \bigoplus_{\alpha\in\mathcal{I}} M_\alpha$ with finite support X, define

$$\phi(a) = \sum_{\alpha\in X} \phi_\alpha(\pi_\alpha(a)).$$

It is then easy to check that ϕ is a monoid homomorphism and the unique one which makes the diagrams commute. □

Once again, following the conventions of category theory, we will call $\bigoplus_{\alpha\in\mathcal{I}} M_\alpha$, or any isomorphic monoid, "the" **direct sum** of $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ without specifying explicitly the injective homomorphisms ι_α. If the index set $\mathcal{I}$ is finite, then the direct sum and direct product coincide so we will use the finite direct product notation $M_1 \times M_2 \times \ldots \times M_n$ for finite direct sums.

If all M_α are the same, that is $M_\alpha = M$ for all $\alpha \in \mathcal{I}$, then we will write $M^{\mathcal{I}}$ for the direct product and $M^{(\mathcal{I})}$ for the direct sum.

Our next task is to show that, similar to the case of free Abelian groups, any free monoid is of the form $(\mathbb{Z}^+)^{(\mathcal{I})}$ for some index class $\mathcal{I}$:

Definition 5.13. *A monoid F is* **free** *with* **basis** *$B \subseteq F$ if any map from B to a monoid M extends uniquely to a monoid homomorphism from F to M. Specifically, if $\Lambda\colon B \to M$ is any map, then there is a unique monoid homomorphism ϕ which makes the following diagram commute:*

$$\begin{tikzcd} F \arrow[r, "\phi"] & M \\ B \arrow[u, "\iota"] \arrow[ur, "\Lambda"'] & \end{tikzcd}$$

where ι is the inclusion map.

Consider the monoid $(\mathbb{Z}^+)^{(\mathcal{I})}$ for an index class $\mathcal{I}$. As described above, this monoid comes with homomorphisms $\iota_\alpha\colon \mathbb{Z}^+ \to (\mathbb{Z}^+)^{(\mathcal{I})}$ and $\pi_\alpha\colon (\mathbb{Z}^+)^{(\mathcal{I})} \to \mathbb{Z}^+$ for all $\alpha \in \mathcal{I}$. We will write $b_\alpha = \iota_\alpha(1) \in (\mathbb{Z}^+)^{(\mathcal{I})}$ for all $\alpha \in \mathcal{I}$, and $B = \{b_\alpha \mid \alpha \in \mathcal{I}\}$. Note that for any $a \in (\mathbb{Z}^+)^{(\mathcal{I})}$ with finite support X, we have

$$a = \sum_{\alpha \in X} \iota_\alpha(\pi_\alpha(a)) = \sum_{\alpha \in X} \pi_\alpha(a)\iota_\alpha(1) = \sum_{\alpha \in X} \pi_\alpha(a) b_\alpha,$$

so B generates $(\mathbb{Z}^+)^{(\mathcal{I})}$.

Proposition 5.14. *The monoid $(\mathbb{Z}^+)^{(\mathcal{I})}$ is free with basis B.*

Proof. Let M be a monoid and $\Lambda\colon B \to M$ a function. Then for $a \in (\mathbb{Z}^+)^{(\mathcal{I})}$ with finite support X we define

$$\phi(a) = \sum_{\alpha \in X} \pi_\alpha(a)\Lambda(b_\alpha).$$

It is then easy to check that ϕ is a monoid homomorphism which makes the diagram commute and the unique such homomorphism. $\square$

Notice that $\mathcal{I}$ is in one-to-one correspondence with B. In fact, $(\mathbb{Z}^+)^{(B)}$ is <u>the</u> free monoid with basis B in the sense of the next proposition:

Proposition 5.15. *If F is a free monoid with basis $B \subseteq F$, then $F \cong (\mathbb{Z}^+)^{(B)}$.*

Proof. We have a bijection between the basis of F and the basis of $(\mathbb{Z}^+)^{(B)}$, so using the freeness of these two monoids, there are monoid homomorphisms from F to $(\mathbb{Z}^+)^{(B)}$ and vice versa. Using the uniqueness part of the definition of freeness one shows, in the standard way, that the compositions of these two homomorphisms are identity maps on F and $(\mathbb{Z}^+)^{(B)}$. $\square$

The existence of free monoids allows us to define monoids using generators and relations. We consider as an example, the universal monoid M generated by two elements p and q such that $3p = 2p$, $3q = 2q$ and $2p + q = 2q + p$. (This monoid will appear again in 11.21.) The monoid M is constructed from the free monoid $F = (\mathbb{Z}^+)^2$ as follows: Set $p' = (1, 0) \in F$

and $q' = (0,1) \in F$, and let $\sim$ be the congruence on F generated by $3p' \sim 2p'$, $3q' \sim 2q'$ and $2p' + q' \sim 2q' + p'$. Then set $p = [p']$, $q = [q']$ and $M = F/\sim$. The monoid M we have constructed is generated by the elements p and q which satisfy the required equations.

The monoid M is conventionally called "the monoid generated by two elements p and q such that $3p = 2p$, $3q = 2q$ and $2p + q = 2q + p$". That is, the word "universal" is dropped. However without the universality condition, the monoid M is not uniquely determined (up to isomorphism). Indeed a one element monoid satisfies the other conditions. The universal property of M is that any monoid satisfying the other conditions is a homomorphic image of M:

Let N be a monoid generated by two elements $\bar{p}$ and $\bar{q}$ satisfying the relations $3\bar{p} = 2\bar{p}$, $3\bar{q} = 2\bar{q}$ and $2\bar{p} + \bar{q} = 2\bar{q} + \bar{p}$. The monoid F is free with basis $B = \{p', q'\}$, so the map from B to N defined by $p' \mapsto \bar{p}$ and $q' \mapsto \bar{q}$, extends to a monoid homomorphism $\phi\colon F \to N$ such that $\phi(p') = \bar{p}$ and $\phi(q') = \bar{q}$.

Let $\bar{\sim}$ be the congruence on F which corresponds to ϕ, namely, $x \mathrel{\bar{\sim}} y \iff \phi(x) = \phi(y)$ for $x, y \in F$. We have by hypothesis that $3p' \mathrel{\bar{\sim}} 2p'$, $3q' \mathrel{\bar{\sim}} 2q'$ and $2p' + q' \mathrel{\bar{\sim}} 2q' + p'$, so $\bar{\sim}$ is one of those congruences on F whose infimum is the congruence $\sim$. In particular, $\sim \subseteq \bar{\sim}$, so there is a surjective homomorphism from $M = F/\sim$ to $F/\bar{\sim}$. Since $\bar{p}$ and $\bar{q}$ generate N, ϕ is surjective and $(F/\bar{\sim}) \cong N$, so there is a surjective monoid homomorphism from M to N.

Of course, a similar argument applies to any monoid defined using generators and relations and provides a universal property for such monoids.

6 Order and Monoids

Vital to our application of monoid theory to module categories, is the order structure that is built into the monoids we will construct from such categories. Thus in this section we investigate the general properties of monoids with respect to their order.

Definition 6.1. *Let M be a semigroup. We define a relation $\leq$ on M by*

$$a \leq b \iff \exists c \in M \text{ such that } a + c = b$$

for elements $a, b \in M$.

Some simple properties of this relation are collected in this proposition:

Proposition 6.2. *Let a, b, c be elements of a semigroup M, and $\phi\colon M \to M'$ a semigroup homomorphism. Then*

1. $a \leq b \leq c \implies a \leq c$
2. $a \leq b \implies a + c \leq b + c$
3. $a \leq b \implies \phi(a) \leq \phi(b)$

If M is a monoid, then, in addition, we have

4. $0 \leq a$
5. $a \leq a$

Items 1 and 5 imply that $\leq$ is a preorder on monoids. Following Wehrung [31], we will call this the **minimum preorder**. Some authors [5, Definition 2.1.1] use the name “algebraic preorder”.

For semigroups, the relation $\leq$ may not be a preorder since there is no certainty that $a \leq a$. An example of this is the semigroup $\mathbb{N}$ in which the relation $\leq$, as defined above, coincides with $<$ with its usual meaning.

We will adopt the notation used in Section 2 for preordered classes. In particular, for $a \in M$ we write

$$\{\leq a\} = \{b \in M \mid b \leq a\}.$$

Though, in general, monoids are preordered, many of the monoids we will work with are, in fact, partially ordered by $\leq$. These we will call **partially ordered monoids**. Warning: This name is used by many authors simply to mean a monoid with a partial order satisfying 6.2.2.

Examples of partially ordered monoids are $\mathbb{Z}^+$, $\{0, \infty\}$. In $\mathbb{Z}^+$ the order $\leq$ coincides with the usual one. In $\{0, \infty\}$ we have $0 \leq \infty$ but $\infty \not\leq 0$. Any Abelian group, $\mathbb{Z}$ for example, has $a \leq b$ for all elements a and b, so these monoids represent another extreme case.

An example of an intermediate case is the monoid $(\mathbb{Z}, \cdot)$, the integers with multiplication as operation. Here we have $a \leq b$ if and only if a divides b, so $\leq$ is not the usual order on $\mathbb{Z}$. Since $-a \leq a \leq -a$ for any $a \in \mathbb{Z}$, this monoid is not partially ordered.

The order in a submonoid A of a monoid M may not be the same as in M: Though $a \leq b$ in A implies $a \leq b$ in M, the converse is not true. For example, in $\mathbb{Z}$ we have $2 \leq 3 \leq 2$;

in the submonoid $\mathbb{Z}^+$ we have $2 \leq 3 \not\leq 2$; and in the submonoid $\{0, 2, 3, 4, \ldots\} = \mathbb{Z}^+ \setminus \{1\}$ we have $2 \not\leq 3 \not\leq 2$. This complication does not occur for submonoids which are also order ideals (soon to be defined).

It is easy to check that if M_1 and M_2 are monoids, then the minimum preorder of the monoid $M_1 \times M_2$ coincides with the order of $M_1 \times M_2$ thought of as a product of preordered classes as in Section 2.

From 2.2, any preordered class $\mathcal{L}$ has associated with it a universal poclass $\overline{\mathcal{L}}$. This same construction applied to a monoid M with its minimum preorder, gives not just a poclass, but a partially ordered monoid:

Definition 6.3. *Let M be a semigroup. We will write $\equiv$ for the relation on M defined by*

$$a \equiv b \iff a \leq b \leq a$$

for $a, b \in M$.

If M is a monoid, then using 6.2.2 and 6.2.5, this relation is easily seen to be a congruence. In this circumstance, we will use the notation

$$\{\equiv a\} = \{b \in M \mid b \leq a \leq b\},$$

for the $\equiv$-congruence class containing $a \in M$. We define $\overline{M} = M/\equiv$.

The monoid $\overline{M}$ is partially ordered and is the largest partially ordered monoid which is a homomorphic image of M:

Proposition 6.4. *Let ϕ: $M \to N$ be a homomorphism between monoids. If N is partially ordered, then there exists a unique monoid homomorphism $\bar{\phi}$ making the following diagram commute.*

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[d, "\{\equiv\ \}"'] & N \\ \overline{M} \arrow[ur, "\bar{\phi}"'] & \end{tikzcd}$$

Proof. Straight forward. □

If ϕ: $M \to N$ is a monoid homomorphism, then by this proposition, there is a unique induced monoid homomorphism $\bar{\phi}$: $\overline{M} \to \overline{N}$ making the following diagram commute:

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[d, "\{\equiv\ \}"'] & N \arrow[d, "\{\equiv\ \}"] \\ \overline{M} \arrow[r, "\bar{\phi}"'] & \overline{N} \end{tikzcd}$$

We will need to define a second relation on semigroups which relates to the order structure:

Definition 6.5. *Let M be a semigroup. We define a relation $\ll$ on M by*

$$a \ll b \iff a + b \leq b$$

for $a, b \in M$. Of course, $a \ll b \iff a + b \equiv b$.

F. Wehrung [31] uses the same notation $a \ll b$ for the stronger relation $a + b = b$. Of course, for a partially ordered monoid these definitions are the same.

The relation $\ll$ is not reflexive so is not a preorder. Also, $\ll$ is not necessarily compatible with the semigroup operation, that is, if $a \ll b$ then it may not be true that $a + c \ll b + c$.

Proposition 6.6. *Let $a, b, c, a_1, a_2, b_1, b_2$ be elements of a semigroup M, and $\phi: M \to M'$ a semigroup homomorphism. Then*

1. $a \ll b \implies a \leq b$
2. $a \ll b \leq c \implies a \ll c$
3. $a \leq b \ll c \implies a \ll c$
4. $a \ll b \ll c \implies a \ll c$
5. $a_1, a_2 \ll b \implies a_1 + a_2 \ll b$
6. $a_1 \ll b_1$ and $a_2 \ll b_2 \implies a_1 + a_2 \ll b_1 + b_2$
7. $a \ll b \implies \phi(a) \ll \phi(b)$

Proof. Trivial. □

As mentioned already, $\ll$ is not reflexive, so an element e of a semigroup such that $e \ll e$ is special:

Definition 6.7.

- *An element e of a semigroup is* **regular** *if any of the following equivalent conditions is true:*
    1. $e \ll e$
    2. $2e \leq e$
    3. $2e \equiv e$

  A **regular semigroup** *is a semigroup whose elements are regular.*

  In a monoid, any element of $\{\leq 0\}$ is regular. Regular elements which are not in $\{\leq 0\}$ we will call **proper**.
- *An element e of a semigroup is an* **idempotent** *if $2e = e$.*

  In a monoid, 0 is an idempotent, and, in fact, the only idempotent in $\{\leq 0\}$. Nonzero idempotents we will call **proper**.

This definition comes from that of regular elements in rings: An element e of a ring is (Von-Neumann) regular if there is some element x such that $e = exe$. When this condition is converted to additive notation with a commutative operation we get $e = 2e + x$, that is, $2e \leq e$.

One readily checks that if M is a semigroup with no regular elements, then M^0 is a monoid with no proper regular elements.

Proposition 6.8. *Let M be a monoid and $\overline{M}$ its associated partially ordered monoid. If $e \in M$, then*

1. *e is regular in $M \iff \{\equiv e\}$ is regular in $\overline{M}$*
2. *e is an idempotent in $M \implies \{\equiv e\}$ is an idempotent in $\overline{M}$*

Proof. Trivial. □

Regular elements and idempotents are closely related:

Proposition 6.9. *If e is a regular element in a monoid, then there is a unique idempotent in $\{\equiv e\}$.*

Proof. Since $2e \le e$, there is some $s \in M$ such that $2e + s = e$. Set $f = e + s$. Then, in addition, we have $e = 2e + s = f + s$, so $e \le f \le e$. Also $2f = 2e + 2s = e + s = f$.

To check the uniqueness of f, we suppose that $2f' = f'$ for some $f' \in \{\equiv e\} = \{\equiv f\}$. Then there are $u, v \in M$ such that $f = f' + u$ and $f' = f + v$, and hence

$$f = f' + u = 2f' + u = 2f + 2v + u = f + 2v + u = f' + u + v = f + v = f'.$$

□

This proposition can be thought of as a consequence of the similarly proved fact that $(\{\equiv e\}, +)$ is a group with identity f. This is a subject we will explore in detail in Section 10.

Proposition 6.10. *If e, e_1, e_2 are regular elements in a monoid M and $a \in M$, then*

1. *$e \le a \implies e \ll a$*
2. *$a \le e \implies a \ll e$*
3. *$e_1, e_2 \ll a \implies e_1 + e_2$ is a regular element such that $e_1 + e_2 \le a$.*

Proof. This follows directly from 6.6. □

We now consider certain submonoids which behave well with respect the partial order of monoids:

Proposition 6.11. *For a nonempty subclass $I \subseteq M$ of a monoid, the following are equivalent:*

1. *$(\forall x, y \in M)\ (x, y \in I \iff x + y \in I)$*
2. *I is a submonoid of M and $(\forall x, y \in M)\ (x \le y \in I \implies x \in I)$*
3. *I is both a submonoid and a lower class of M.*
4. *$I = \ker \phi$ for some homomorphism $\phi\colon M \to \{0, \infty\}$.*

Proof. Easy. □

Definition 6.12. *A subclass $I \in M$ of a monoid satisfying any of the conditions of this proposition is called an* **order ideal** [2] *of M. We will write $\mathcal{L}(M)$ for the class of order ideals of a monoid M, ordered by inclusion. Consistent with this definition, we will write $I \le M$ if I is an order ideal of M.*

A monoid is **conical** *if $\{0\}$ is an order ideal, that is, if $x \le 0$ implies $x = 0$ for all $x \in M$.*

An order ideal, $I \subseteq M$, is a subclass of a monoid which not only preserves the monoid operation, but also the order. More precisely, if $x, y \in I$ then $x \leq y$ in I if and only if $x \leq y$ in M.

An order ideal is a union of $\equiv$-congruence classes of M, so that I is an order ideal of M if and only if $\overline{I}$ is an order ideal of $\overline{M}$. In particular, $\mathcal{L}(M) \cong \mathcal{L}(\overline{M})$.

From 6.6, if $a \in M$ then $\{\ll a\}$ is always an order ideal, whereas $\{\leq a\}$ is an order ideal if and only if a is regular. In particular, $\{\leq 0\}$ is an order ideal and, since it is contained in any other order ideal, $\{\leq 0\}$ is the minimum element of $\mathcal{L}(M)$. The maximum element of $\mathcal{L}(M)$ is M itself.

Note also that order ideals are normal submonoids. So that, for all $a \in M$ and $I \leq M$ we have

$$[a]_I = 0 \iff a \in I.$$

Also, if $I \leq J \leq M$ then, for all $a \in M$, we have $[a]_I^M = [a]_I^J$.

The following gives some other easy properties of order ideals:

Proposition 6.13. *Let M be a monoid.*

1. *If $\{I_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of order ideals of M, then $I = \cap_{\alpha\in\mathcal{I}} I_\alpha$ is an order ideal of M.*
2. *If $\{I_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of order ideals of M which is totally ordered by inclusion, then $I = \cup_{\alpha\in\mathcal{I}} I_\alpha$ is an order ideal of M.*
3. *If $\phi\colon M \to N$ is a homomorphism and $B \leq N$ an order ideal, then $I = \phi^{-1}(B)$ is an order ideal in M.*
4. *If I is an order ideal of M then M/I is conical.*
5. *If $J \subseteq I \leq M$ are monoids then $J \leq I$ if and only if $J \leq M$.*

Proof. Easy. □

If Y is any nonempty subclass of a monoid M, then we can use 1 above to define the order ideal generated by Y, namely

$$\langle Y \rangle = \bigcap \{I \mid Y \subseteq I \leq M\}.$$

Of course, $\langle Y \rangle$ is the smallest order ideal containing Y. The order ideal generated by Y will contain the submonoid generated by Y and this inclusion is, in general, proper.

Proposition 6.14. *If Y is a nonempty subclass of a monoid M, then $y \in \langle Y \rangle$ if and only if $y \leq y_1 + y_2 + \cdots + y_n$ for some $y_1, y_2, \ldots, y_n \in Y$.*

Proof. Let

$$Z = \{y \in M \mid \exists y_1, y_2, \ldots, y_n \in Y \text{ such that } y \leq y_1 + y_2 + \cdots + y_n\}.$$

This is easily seen to be an order ideal which contains Y, so $\langle Y \rangle \subseteq Z$. On the other hand, all elements of form $y_1 + y_2 + \cdots + y_n$ with $y_1, y_2, \ldots, y_n \in Y$ are in $\langle Y \rangle$ because $\langle Y \rangle$ is a submonoid, and then all other elements of Z are in $\langle Y \rangle$ because $\langle Y \rangle$ is a lower set. Thus $Z \subseteq \langle Y \rangle$. □

Item 1 of 6.13 can be also reinterpreted as saying that every nonempty subclass $\mathcal{K} \subseteq \mathcal{L}(M)$ has an infimum:

$$\inf \mathcal{K} = \bigcap_{I\in\mathcal{K}} I.$$

This, and the fact that M itself is an order ideal, means that every nonempty subclass $\mathcal{K} \subseteq \mathcal{L}(M)$ also has a supremum, namely:

$$\sup \mathcal{K} = \langle \bigcup_{I \in \mathcal{K}} I \rangle.$$

We have shown by the above discussion that for any monoid M, the class of order ideals, $\mathcal{L}(M)$, is a complete lattice.

We will adopt the lattice theoretical notation for segments of lattices: if $I \subseteq J$ are order ideals of M then

$$[I, J] = \{K \in \mathcal{L}(M) \mid I \subseteq K \subseteq J\}.$$

This is, of course, a sublattice of $\mathcal{L}(M)$.

Let I be an order ideal in a monoid M and $\sigma\colon M \to M/I$ the quotient homomorphism. From 6.13.3, we know already that σ^{-1} maps order ideals of M/I to order ideals of M which necessarily contain I. Thus we can consider σ^{-1} to be map from $\mathcal{L}(M/I)$ to $[I, M]$. This map is actually a lattice isomorphism:

Proposition 6.15. *Let I be an order ideal in a monoid M and $\sigma\colon M \to M/I$, the quotient homomorphism. Then $\sigma^{-1}\colon \mathcal{L}(M/I) \to [I, M] \subseteq \mathcal{L}(M)$ is a lattice isomorphism. The inverse isomorphism is induced from σ.*

Proof. We show first that if $I \le J \le M$ then $\sigma(J) \le M/I \ldots$

Since J is a submonoid of M, $\sigma(J)$ is a submonoid of M/I, so we need only check that $\sigma(J)$ is a lower class. If $x \in M$ and $y \in J$ such that $\sigma(x) \le \sigma(y) \in \sigma(J)$, then there is some $u \in M$ such that $\sigma(x) + \sigma(u) = \sigma(y)$, that is, $[x + u]_I = [y]_I$. Thus there are $s, t \in I \le J$ such that $x + u + s = y + t$, and, in particular, $x \le y + t \in J$. J is an order ideal so $x \in J$ and $\sigma(x) \in \sigma(J)$. This shows that $\sigma(J)$ is an order ideal in M/I.

Next we show $J = \sigma^{-1}(\sigma(J))$ for order ideals J with $I \le J \le M \ldots$

The inclusion $J \subseteq \sigma^{-1}(\sigma(J))$ is true for any subclass J, so we need check only the reverse inclusion. Suppose $y \in \sigma^{-1}(\sigma(J))$ then there is $x \in J$ such that $\sigma(y) = \sigma(x)$, that is, $[y]_I = [x]_I$, Thus there are $s, t \in I \le J$ such that $y + s = x + t$, and, in particular, $y \le x + t \in J$. J is an order ideal so $y \in J$.

Finally, since σ is surjective, we have $K = \sigma(\sigma^{-1}(K))$ for any order ideal (indeed, for any subclass) K of M/I. Since we know also that $\sigma^{-1}(K)$ is in $[I, M]$, this then completes the proof. □

We will see in the next section that in decomposition and refinement monoids, there are other stronger relationships between sublattices of $\mathcal{L}(M)$.

It will be useful to have a notation for the relationship of membership in an order ideal generated by a single element:

Definition 6.16. *Let M be a monoid. We define a relation $\prec$ on M by*

$$a \prec b \iff a \in \langle \{b\} \rangle$$

for $a, b \in M$. This relation also gives us a new notation for the order ideal generated by a single element: $\{\prec a\} = \langle \{a\} \rangle$.

From 6.14 we have

$$a \prec b \iff (\exists n \in \mathbb{N} \text{ such that } a \leq nb).$$

We collect here some simple properties of this relation.

Proposition 6.17. *Let a, b, c be elements of a monoid M. Then*

1. $0 \prec a$
2. $a \prec a$
3. $a \prec b \prec c \implies a \prec c$
4. $a \prec b \implies a + c \prec b + c$

Proof.

1. Trivial.
2. Trivial.
3. Since b is in the order ideal $\{\prec c\}$, we get $\{\prec b\} \subseteq \{\prec c\}$, and so $a \in \{\prec b\} \subseteq \{\prec c\}$. That is, $a \prec c$.
4. Since $a \prec b$, there is some $n \in \mathbb{N}$ such that $a \leq nb$. But then $a + c \leq nb + c \leq n(b+c)$. Thus $a + c \prec b + c$.

$\square$

From 2 and 3 of this proposition we see that $\prec$ is a preorder on M.

It is easy to check that for a finite set, $\{a_1, a_2, \ldots, a_n\} \subseteq M$,

$$\langle\{a_1, a_2, \ldots, a_n\}\rangle = \{\prec a_1 + a_2 + \cdots + a_n\},$$

so that any finitely generated order ideal is, in fact, generated by a single element.

Using the preorder $\leq$ on M, we constructed a partially ordered monoid $\overline{M}$ with a universal property. This same process can be applied to the preorder $\prec$ with similar consequences:

Definition 6.18. *Let M be a monoid. We will write $\asymp$ for the relation on M defined by*

$$a \asymp b \iff a \prec b \prec a$$

for $a, b \in M$. Using 6.17, this relation is easily seen to be a congruence. We will use the notation

$$\{\asymp a\} = \{b \in M \mid b \prec a \prec b\},$$

for the $\asymp$-congruence class containing $a \in M$. We define $\widetilde{M} = M/\asymp$. The $\asymp$-congruence classes are called the **Archimedean components** [7, Chapter 4.3] *of M.*

Note that $\{\asymp a\} \leq \{\asymp b\}$ in $\widetilde{M}$ if and only if $a \prec b$ in M.

To discuss the universal property of $\widetilde{M}$, we need to define a new type of monoid:

Definition 6.19. *A poclass, $\mathcal{L}$, is a* **(join)-semilattice** [3, Page 9] *if for each pair of elements, $a, b \in \mathcal{L}$, the supremum $a \vee b$ exists.*

A semilattice $\mathcal{L}$ together with the operation $\vee$ is a semigroup in which $a \vee a = a$ for all $a \in \mathcal{L}$. In fact, this property characterizes semilattices among semigroups:

Proposition 6.20. *If M is a semigroup such that $a = 2a$ for all $a \in M$, then*

$$(\forall a, b \in M)\ (a \leq b \iff a + b = b \iff a \ll b),$$

and M with its minimum preorder is a semilattice in which $+$ and $\vee$ coincide.

Proof. [18, 1.3.2] If $a \leq b$ then there is some $c \in M$ such that $a+c=b$, so $a+b=a+a+c=a+c=b$. Thus $a \leq b \implies a+b=b$. The other implications above are trivial.

If $a \in M$, then $a=a+a$ implies that $a \leq a$. Thus $\leq$ is a preorder on M. If $a \leq b \leq a$ then $b=a+b=a$, so M is partially ordered by $\leq$.

It remains to show that $a+b=a \vee b$ for any $a, b \in M$...

If $a, b \leq x$ then $x+(a+b)=(x+a)+b=x+b=x$, so $a+b \leq x$. Thus $a+b$ is the supremum of a and b. $\square$

If a semilattice $\mathcal{L}$ has a minimum element $\perp$, then $(\mathcal{L}, \vee)$ is a monoid with identity $\perp$. Conversely, a monoid M such that $a=2a$ for all $a \in M$ is a semilattice with minimum element 0. Thus we have two ways of thinking of the same mathematical object: Either as a poclass with some special properties or as a semigroup with some special properties. We will call both $\mathcal{L}$ and M as described above **semilattice monoids**.

Proposition 6.21. *For any monoid M, $\widetilde{M}$ is a semilattice monoid.*

Proof. Let $\{\asymp a\} \in \widetilde{M}$ for some $a \in M$. Then $2\{\asymp a\}=\{\asymp 2a\}$, but a and $2a$ are in the same Archimedean component of M, so $\{\asymp 2a\}=\{\asymp a\}$. The claim then follows from the previous proposition. $\square$

The universal property of $\widetilde{M}$ can now be specified:

Proposition 6.22. *Let M and N be monoids and $\phi\colon M \to N$ a monoid homomorphism. If N is a semilattice monoid, then ϕ factors uniquely through $\widetilde{M}$.*

Proof. Straight forward. $\square$

Of course, since $\widetilde{M}$ is partially ordered, the quotient map from M to $\widetilde{M}$ factors through $\overline{M}$. So given the hypotheses of the proposition above, we get the following commutative diagram:

$$\begin{tikzcd} M \arrow[r] \arrow[drr, "\phi"'] & \overline{M} \arrow[r] & \widetilde{M} \arrow[d] \\ & & N \end{tikzcd}$$

where an element $a \in M$ maps to $\phi(a)$ via $a \mapsto \{\equiv a\} \mapsto \{\asymp a\} \mapsto \phi(a)$.

Finally we note that $(\mathcal{L}(M), \vee)$ is a semilattice monoid with minimum element $\{\leq 0\}$, and that $\widetilde{M}$ embeds in $\mathcal{L}(M)$:

Proposition 6.23. *Let M be a monoid. Then the map from $\widetilde{M}$ to $(\mathcal{L}(M), \vee)$ defined by $\{\asymp a\} \mapsto \{\prec a\}$ for $a \in M$ is an injective monoid homomorphism.*

Proof. Consider the map $\phi\colon M \to \mathcal{L}(M)$ given by $\phi(a)=\{\prec a\}$ for $a \in M$. If $a, b \in M$ then $\phi(a+b)=\{\prec a+b\}$ is an order ideal containing $\{\prec a\}$ and $\{\prec b\}$. Since any order ideal which contains $\{\prec a\}$ and $\{\prec b\}$ must also contain $a+b$, and hence $\{\prec a+b\}$, we must have $\{\prec a+b\}=\{\prec a\} \vee \{\prec b\}$, that is, $\phi(a+b)=\phi(a) \vee \phi(b)$.

Since also $\phi(0)=\{\prec 0\}=\{\leq 0\}$, ϕ is a monoid homomorphism from M to a semilattice monoid.

Using the universal property of $\widetilde{M}$, there is an induced monoid homomorphism defined by $\{\asymp a\} \mapsto \{\prec a\}$. It is easy to check that $\{\prec a\}=\{\prec b\}$ implies that a and b are in the same Archimedean component, and so this induced homomorphism is injective. $\square$

7 Refinement and Decomposition Monoids

In the last section we discussed order in monoids. This is one important ingredient in understanding the monoids which we will use to study module categories in 16.

The other main ingredient is the refinement property that is the subject of this section. This monoid property has appeared in other contexts and so has been studied before. See for example, Tarski [30], Wehrung [31], [32], Dobbertin [8], Ara, Goodearl, O'Meara and Pardo [2], or Goodearl [10].

Definition 7.1. *Let M be a semigroup. Then*

1. *M has* **refinement** *if for all $a_1, a_2, b_1, b_2 \in M$ with $a_1 + a_2 = b_1 + b_2$, there exist $c_{11}, c_{12}, c_{21}, c_{22} \in M$ such that*
$$a_1 = c_{11} + c_{12} \qquad a_2 = c_{21} + c_{22}$$
$$b_1 = c_{11} + c_{21} \qquad b_2 = c_{12} + c_{22}.$$
2. *M has* **(Riesz) decomposition** *if for all $a, b_1, b_2 \in M$ with $a \leq b_1 + b_2$, there exist $c_1, c_2 \in M$ such that $a_1 = c_1 + c_2$, $c_1 \leq b_1$ and $c_2 \leq b_2$.*

The main theme in later sections will be monoids with these properties, that is decomposition monoids and refinement monoids. In a few examples we will make use of refinement semigroups, primarily as a means of defining interesting refinement monoids. One readily checks that if M is a refinement semigroup, then adjoining an identity element to form M^0 yields a refinement monoid.

It is convenient to record refinements and decompositions using matrices: The refinement of $a_1 + a_2 = b_1 + b_2$ from the definition would be written

$$\begin{array}{cc} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \end{array}$$

This means that the sum of the entries in each row (column) equals the entry labeling the row (column).

The decomposition from the definition would be written as

$$\begin{array}{cc} & a \\ \begin{array}{c} b_1 \geq \\ b_2 \geq \end{array} & \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \end{array}$$

or

$$\begin{array}{cc} & \begin{array}{cc} \leq b_1 & \leq b_2 \end{array} \\ a & \begin{pmatrix} c_1 & c_2 \end{pmatrix} \end{array}$$

meaning $a_1 = c_1 + c_2$, $c_1 \leq b_1$, and $c_2 \leq b_2$.

Suppose we have the equation $a_1 + a_2 + a_3 = b_1 + b_2$ in a refinement semigroup. Using the refinement property we get the refinement matrix

$$\begin{array}{c c} & \begin{matrix} b_1 & b_2 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 + a_3 \end{matrix} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \end{array}$$

The equation $a_2 + a_3 = c_{21} + c_{22}$ can itself be refined:

$$\begin{array}{c c} & \begin{matrix} c_{21} & c_{22} \end{matrix} \\ \begin{matrix} a_2 \\ a_3 \end{matrix} & \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \end{array}$$

Thus we have a 3×2 refinement matrix:

$$\begin{array}{c c} & \begin{matrix} b_1 & b_2 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} & \begin{pmatrix} c_{11} & c_{12} \\ d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \end{array}$$

Notice that this refinement matrix has the same entries in the top row, c_{11} and c_{12}, as the original one. We will say that this refinement matrix has been obtained by **further refinement** from the original one.

Using further refinement and induction, the refinement and decomposition properties can be extended to equations and inequalities with more than two terms. For example, in a refinement semigroup, if $a_1 + a_2 + \cdots + a_m = b_1 + b_2 + \cdots + b_n$, then there is a refinement matrix of the following form:

$$\begin{array}{c c} & \begin{matrix} b_1 & b_2 & \dots & b_n \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{matrix} & \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix} \end{array}$$

Here, as in matrix theory, we use the symbol $*$ for entries in refinement matrices that do not need to be explicitly named.

It is also worth noting that in a <u>refinement</u> semigroup, if $a_1 + a_2 + \cdots + a_m \le b_1 + b_2 + \cdots + b_n$, then there are refinement matrices of the following forms:

$$\begin{array}{c c} & \begin{matrix} \le b_1 & \le b_2 & \dots & \le b_n \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{matrix} & \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix} \end{array} \qquad \begin{array}{c c} & \begin{matrix} b_1 & b_2 & \dots & b_n \end{matrix} \\ \begin{matrix} a_1 \le \\ a_2 \le \\ \vdots \\ a_m \le \end{matrix} & \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix} \end{array}$$

(In fact, in the matrix on the right, all but one of the $\le$ symbols could be replaced by equalities.) This property is not true in general for decomposition semigroups. See Example 7.4.

For a decomposition semigroup, we get the following: If $a \leq b_1 + b_2 + \cdots + b_n$, then there is a decomposition matrix of the form:

$$\begin{array}{cc} & \begin{array}{cccc} \leq b_1 & \leq b_2 & \ldots & \leq b_n \end{array} \\ a & \begin{pmatrix} * & * & \ldots & * \end{pmatrix} \end{array}$$

Proposition 7.2. *If M is a monoid, then*

$$\begin{aligned} M \text{ has refinement } &\Longrightarrow M \text{ has decomposition} \\ &\Longrightarrow \overline{M} \text{ has decomposition} \\ &\Longrightarrow \widetilde{M} \text{ has decomposition } \iff \widetilde{M} \text{ has refinement} \end{aligned}$$

Proof. We show first that if $\overline{M}$ has decomposition, then so does $\widetilde{M}$:

Suppose we have $\{\asymp a\} \leq \{\asymp b_1\} + \{\asymp b_2\}$ for some $a, b_1, b_2 \in M$. This implies that $a \prec b_1 + b_2$, that is, there is some $n \in \mathbb{N}$ such that $a \leq n(b_1 + b_2) = nb_1 + nb_2$. Since M has decomposition, there are $c_1, c_2 \in M$ such that $c_1 \leq nb_1$, $c_2 \leq nb_2$ and $a = c_1 + c_2$. Hence $\{\asymp c_1\} \leq \{\asymp b_1\}$, $\{\asymp c_2\} \leq \{\asymp b_2\}$ and $\{\asymp a\} = \{\asymp c_1\} + \{\asymp c_2\}$.

Since $\widetilde{M}$ is a semilattice, the claim that $\widetilde{M}$ has decomposition if and only if it has refinement is a special case of the next proposition. The remaining claims are all trivial. □

It is an open and interesting question whether there is a refinement monoid M such that $\overline{M}$ does not have refinement.

Proposition 7.3. *Let M be a semilattice. Then M has decomposition if and only if it has refinement.*

Proof. We will write $+$ for the operation in M, so that $2a = a$ for all $a \in M$ and

$$(\forall a, b \in M)\ (a \leq b \iff a + b = b).$$

Suppose that M has decomposition and there are $a_1, a_2, b_1, b_2 \in M$ such that $a_1 + a_2 = b_1 + b_2$. Since $a_1, a_2 \leq b_1 + b_2$, there are $c_{11}, c_{12}, c_{21}, c_{22} \in M$ such that $a_1 = c_{11} + c_{12}$, $a_2 = c_{21} + c_{22}$, $c_{11}, c_{21} \leq b_1$ and $c_{12}, c_{22} \leq b_2$. Similarly, there are $d_{11}, d_{12}, d_{21}, d_{22} \in M$ such that $b_1 = d_{11} + d_{21}$, $b_2 = d_{12} + d_{22}$, $d_{11}, d_{12} \leq a_1$ and $d_{21}, d_{22} \leq a_2$. It is then easy to check that

$$\begin{array}{cc} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} c_{11} + d_{11} & c_{12} + d_{12} \\ c_{21} + d_{21} & c_{22} + d_{22} \end{pmatrix} \end{array}$$

is a refinement of the original equation.

We confirm this for a_1: We have $a_1 \leq a_1 + d_{11} + d_{12} = c_{11} + d_{11} + c_{12} + d_{12} \leq 4a_1 = a_1$. Since M is partially ordered, this implies $c_{11} + d_{11} + c_{12} + d_{12} = a_1$.

The converse, as noted in the previous proposition, is trivial. □

Example 7.4. *A decomposition monoid without refinement:*

Let $M = \{0, 1, \infty\}$, where $1 + 1 = \infty$. It is a simple calculation to show that M is a decomposition monoid, but that the equation $1 + 1 = \infty + \infty$ has no refinement.

Also worth noting is that we have $\infty+\infty \leq 1+1$, but there is no decomposition in either of the forms

$$\begin{array}{c c} & \begin{array}{cc} \leq 1 & \leq 1 \end{array} \\ \begin{array}{c} \infty \\ \infty \end{array} & \begin{pmatrix} * & * \\ * & * \end{pmatrix} \end{array} \qquad \begin{array}{c c} & \begin{array}{cc} 1 & 1 \end{array} \\ \begin{array}{c} \infty \leq \\ \infty \leq \end{array} & \begin{pmatrix} * & * \\ * & * \end{pmatrix} \end{array}$$

One checks easily that $\{0,\infty\}$ has refinement.

Example 7.5. *Let M be an Abelian group written additively. Then if $a_1, a_2, b_1, b_2 \in M$ with $a_1 + a_2 = b_1 + b_2$, we get the refinement*

$$\begin{array}{c c} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} 0 & a_1 \\ b_1 & a_2 - b_1 \end{pmatrix} \end{array}$$

Thus Abelian groups are refinement monoids.

This example shows that $\mathbb{Z}$ is a refinement monoid. We show the same for $\mathbb{Z}^+$: With $a_1, a_2, b_1, b_2 \in \mathbb{Z}^+$ such that $a_1 + a_2 = b_1 + b_2$, we either have $a_2 - b_1 \in \mathbb{Z}^+$ or $b_1 - a_2 \in \mathbb{Z}^+$. In the first case we can make a refinement as in the example above. In the second case we have the refinement

$$\begin{array}{c c} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} b_1 - a_2 & b_2 \\ a_2 & 0 \end{pmatrix} \end{array}$$

with entries in $\mathbb{Z}^+$.

Even though $\mathbb{Z}$ and $\mathbb{Z}^+$ have refinement, $\mathbb{N}$ does not even have decomposition since there is no decomposition of $1 \leq 1+1$ in $\mathbb{N}$.

Similarly one shows that $\mathbb{R}$ and $\mathbb{R}^+$ are refinement monoids. Unlike $\mathbb{N}$, the semigroup $\mathbb{R}^{++}$ has refinement:

Example 7.6. *Let $a_1, a_2, b_1, b_2 \in \mathbb{R}^{++}$ be such that $a_1 + a_2 = b_1 + b_2$. Without loss of generality, we can assume that $a_1 \leq a_2, b_1, b_2$ in the usual order in $\mathbb{R}$. It is easily checked that all entries in the refinement matrix*

$$\begin{array}{c c} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} \frac{1}{2}a_1 & \frac{1}{2}a_1 \\ b_1 - \frac{1}{2}a_1 & b_2 - \frac{1}{2}a_1 \end{pmatrix} \end{array}$$

are in $\mathbb{R}^{++}$.

Thus $\mathbb{R}^{++}$ is a refinement semigroup.

Example 7.7. *Let $\mathcal{L}$ be a distributive lattice (see Section 2). Then $(\mathcal{L}, \vee)$ is a refinement semigroup: If $a_1, a_2, b_1, b_2 \in \mathcal{L}$ are such that $a_1 \vee a_2 = b_1 \vee b_2$, then*

$$\begin{array}{c c} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} a_1 \wedge b_1 & a_1 \wedge b_2 \\ a_2 \wedge b_1 & a_2 \wedge b_2 \end{pmatrix} \end{array}$$

is a refinement matrix.

Since distributivity is a self dual property, $(\mathcal{L}, \wedge)$ is also a refinement semigroup.

As a special case of this, let X be a class and $\mathcal{M}$ a nonempty family of subclasses of X which is closed under union and intersection. Then $(\mathcal{M}, \cup)$ is a distributive lattice, and so given $A_1, A_2, B_1, B_2 \in \mathcal{M}$ with $A_1 \cup A_2 = B_1 \cup B_2$ we get the refinement

$$\begin{array}{cc} & \begin{array}{cc} B_1 & B_2 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{pmatrix} A_1 \cap B_1 & A_1 \cap B_2 \\ A_2 \cap B_1 & A_2 \cap B_2 \end{pmatrix} \end{array}$$

The semigroup $(\mathcal{M}, \cap)$ is also a refinement semigroup.

We have already seen in the case $\mathbb{N} \subseteq \mathbb{Z}$, that subsemigroups of a refinement semigroup may not have refinement or decomposition. The same is true of submonoids of refinement monoids. See, for example, 8.8. For order ideals we have a much better situation:

Proposition 7.8. *Let I be an order ideal in a monoid M. Then*

1. *M has refinement $\Longrightarrow$ I and M/I have refinement.*
2. *M has decomposition $\Longrightarrow$ I and M/I have decomposition.*

Proof.

1. I has refinement: Let $a_1, a_2, b_1, b_2 \in I$ be such that $a_1 + a_2 = b_1 + b_2$. Refinement in M implies there is a refinement of the form

$$\begin{array}{cc} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \end{array}$$

The elements c_{ij} are bounded above by elements of I, so, since I is an order ideal, they are also in I. Thus I is a refinement monoid.
M/I has refinement: Let $a_1, a_2, b_1, b_2 \in M$ be such that $[a_1]_I + [a_2]_I = [b_1]_I + [b_2]_I$. Then there are $u_1, u_2 \in I$ such that $a_1 + a_2 + u_1 = b_1 + b_2 + u_2$. Since M is a refinement monoid, there is a refinement matrix of the form

$$\begin{array}{cc} & \begin{array}{ccc} b_1 & b_2 & u_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \\ u_1 \end{array} & \begin{pmatrix} c_{11} & c_{12} & u_3 \\ c_{21} & c_{22} & u_4 \\ u_5 & u_6 & u_7 \end{pmatrix} \end{array}$$

The elements u_3, u_4, u_5, u_6, u_7 are bounded above by either u_1 or u_2 which are in I. Since I is an order ideal, we have $u_i \in I$, and $[u_i]_I = 0$ for $i = 3, 4, \ldots, 7$. The matrix

$$\begin{array}{cc} & \begin{array}{cc} [b_1]_I & [b_2]_I \end{array} \\ \begin{array}{c} [a_1]_I \\ [a_2]_I \end{array} & \begin{pmatrix} [c_{11}]_I & [c_{12}]_I \\ [c_{21}]_I & [c_{22}]_I \end{pmatrix} \end{array}$$

is then a refinement of $[a_1]_I + [a_2]_I = [b_1]_I + [b_2]_I$ in M/I.
2. The proof for decomposition monoids is very similar to the proof for refinement monoids.

□

Let $\mathcal{K} \subseteq \mathcal{L}(M)$ be a family of order ideals of M. In general, it may be possible that the submonoid generated by $\mathcal{K}$, namely $\sum \mathcal{K}$, maybe smaller than the order ideal generated by $\mathcal{K}$, namely $\sup \mathcal{K}$. For a decomposition monoid, however, the submonoid and the order ideal generated by $\mathcal{K}$ are the same:

Proposition 7.9. *Let M be a decomposition monoid, and $\mathcal{K} \subseteq \mathcal{L}(M)$, a family of order ideals of M. Then*

$$\sup \mathcal{K} = \sum \mathcal{K}.$$

Proof. Since $\sum \mathcal{K}$ is a submonoid of $\sup \mathcal{K}$, it suffices to show that $\sum \mathcal{K}$ is a lower class…

Suppose then that we have $a \leq b$ with $b \in \sum \mathcal{K}$. Then $b = b_1 + b_2 + \cdots + b_n$ for some elements $b_1, b_2, \ldots b_n \in \bigcup \mathcal{K}$. Using the decomposition property, there are $c_1, c_2, \ldots, c_n$ such that $a = c_1 + c_2 + \cdots + c_n$ and $c_i \leq b_i$ for $i = 1, 2, \ldots, n$.

But then each c_i is in the same order ideal of $\mathcal{K}$ that contains b_i, in particular, $c_i \in \bigcup \mathcal{K}$. Thus a is a finite sum of elements of $\bigcup \mathcal{K}$ and $a \in \sum \mathcal{K}$. □

The above proposition says that if A and B are order ideals then so is $A + B$, and in the lattice $\mathcal{L}(M)$, $A \vee B = A + B$. We already know that $A \wedge B = A \cap B$.

Proposition 7.10. *For any decomposition monoid M, the lattice $\mathcal{L}(M)$ is distributive.*

Proof. From the discussion in Section 2, it suffices to show that $A \cap (B + C) \subseteq (A \cap B) + (A \cap C)$ for all $A, B, C \in \mathcal{L}(M)$…

Suppose $a \in A \cap (B + C)$, then there are $y \in B, z \in C$ such that $y + z = a \in A$. But A is an order ideal, so $y, z \in A$. Thus $y \in A \cap B$, $z \in A \cap C$ and $a \in (A \cap B) + (A \cap C)$. □

Proposition 7.11. *Let I be an order ideal in a decomposition monoid M. Then the quotient homomorphism $\sigma\colon M \to M/I$ induces a lattice homomorphism from $\mathcal{L}(M)$ to $\mathcal{L}(M/I)$. Specifically, if $A, B \in \mathcal{L}(M)$ then*

1. *$\sigma(A)$ is an order ideal.*
2. *$\sigma(A \cap B) = \sigma(A) \cap \sigma(B)$*
3. *$\sigma(A + B) = \sigma(A) + \sigma(B)$.*

In addition, $\sigma^{-1}(\sigma(A)) = A + I$.

Proof. Let $A \in \mathcal{L}(M)$. From 5.10, we have $\sigma(A) = (A + I)/I$ and so $\sigma(A) = \sigma(A + I)$. Since $A + I$ is an order ideal in $[I, M]$, 6.15 implies that $\sigma(A) = \sigma(A + I)$ is an order ideal in M/I and

$$\sigma^{-1}(\sigma(A)) = \sigma^{-1}(\sigma(A + I)) = A + I.$$

The map σ can be considered the composition of the maps, $\sigma'\colon \mathcal{L}(M) \to [I, M]$ and $\sigma\colon [I, M] \to \mathcal{L}(M/I)$, the restriction of σ to $[I, M]$, so that

$$A \xmapsto{\sigma'} A + I \xmapsto{\sigma} (A + I)/I.$$

The map $\sigma\colon [I, M] \to \mathcal{L}(M/I)$ is a lattice isomorphism by 6.15, so it remains to show that σ' is a lattice homomorphism…

If $A, B \leq M$, then

$$\sigma'(A + B) = A + B + I = (A + I) + (B + I) = \sigma'(A) + \sigma'(B)$$
$$\sigma'(A \cap B) = (A \cap B) + I = (A + I) \cap (B + I) = \sigma'(A) \cap \sigma'(B)$$

The last result used that $\mathcal{L}(M)$ is distributive. Thus σ' is a lattice homomorphism. □

For a decomposition monoid M, $\mathcal{L}(M)$ is modular so we get immediately

Proposition 7.12. *Let M be a decomposition monoid and $A, B \in \mathcal{L}(M)$. Then the map $\phi\colon [A \cap B, A] \to [B, A+B]$ defined by $\phi(I) = I + B$ is a lattice isomorphism with inverse given by $\phi^{-1}(J) = J \cap A$.*

This result we can strengthen considerably, especially for refinement monoids, where we have in fact a monoid isomorphism $A/(A \cap B) \cong (A+B)/B$:

Proposition 7.13. *Let A and B be two order ideals in a refinement monoid and let $\psi\colon A/(A \cap B) \to (A+B)/B$ be the map defined by $\psi([a]_{A\cap B}) = [a]_B$ for all $a \in A$. Then ψ is a monoid isomorphism of $A/(A \cap B)$ and $(A+B)/B$.*

Proof. Since $A \cap B \subseteq B$, the map $[a]_{A\cap B} \mapsto [a]_B$ is a well defined homomorphism from $M/(A \cap B)$ to M/B. If $a \in A$ then $[a]_B \in (A+B)/B$, so ψ is just the restriction of this map to $A/(A \cap B)$, and is itself a homomorphism. It remains to check only that ψ is a bijection from $A/(A \cap B)$ to $(A+B)/B$...

$\underline{\psi \text{ surjective:}}$ Let $[a+b]_B \in (A+B)/B$ with $a \in A$ and $b \in B$. Then $[a+b]_B = [a]_B = \psi([a]_{A\cap B}) \in \psi(A/(A \cap B))$.

$\underline{\psi \text{ injective:}}$ Suppose $\psi([a]_{A\cap B}) = \psi([a']_{A\cap B})$ for some $a, a' \in A$. Then $[a]_B = [a']_B$ so there are $b, b' \in B$ such that $a + b = a' + b'$. Using the refinement property we get the following refinement matrix:

$$\begin{array}{cc} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} a' \\ b' \end{array} & \begin{pmatrix} x & x' \\ y' & y \end{pmatrix} \end{array}.$$

But $y' \le a, b'$ so $y' \in A \cap B$. Similarly, $x' \in A \cap B$, and we get $[a]_{A\cap B} = [x+y']_{A\cap B} = [x]_{A\cap B} = [x+x']_{A\cap B} = [a']_{A\cap B}$. □

For decomposition monoids we do not get quite as strong a result as for refinement monoids:

Proposition 7.14. *Let A and B be two order ideals in a decomposition monoid and $\psi\colon \overline{A/(A \cap B)} \to \overline{(A+B)/B}$ as in the previous proposition. Then ψ induces an isomorphism of $\overline{A/(A \cap B)}$ and $\overline{(A+B)/B}$.*

Proof. As in the previous proposition, ψ is a surjective homomorphism from $\psi\colon A/(A \cap B)$ to $(A+B)/B$. Without refinement, ψ would not be injective. Instead we get that the induced map $\overline{\psi}\colon \overline{A/(A \cap B)} \to \overline{(A+B)/B}$ is injective:

Suppose $\psi([a]_{A\cap B}) \equiv \psi([a']_{A\cap B})$ for some $a, a' \in A$, then $[a]_B \equiv [a']_B$. In particular, $[a]_B \le [a']_B$. So there is $b' \in B$ such that $a \le a' + b'$. From the decomposition property, there are $x, y \in M$ such that $x \le a'$, $y \le b'$ and $a = x + y$. But then $y \le a, b'$, so $y \in A \cap B$ and $[a]_{A\cap B} = [x+y]_{A\cap B} = [x]_{A\cap B} \le [a']_{A\cap B}$. Similarly, $[a']_{A\cap B} \le [a]_{A\cap B}$, and so $[a]_{A\cap B} \equiv [a']_{A\cap B}$, that is, $[a]_{A\cap B}$ and $[a']_{A\cap B}$ represent the same element of $\overline{A/(A \cap B)}$. □

Among the easy consequences of this proposition is that for decomposition monoids, ψ induces a lattice isomorphism between $\mathcal{L}(A/(A \cap B))$ and $\mathcal{L}((A+B)/B)$. From 6.15, we have also isomorphisms $\sigma_{A\cap B}\colon [A \cap B, A] \to \mathcal{L}(A/A \cap B)$ and $\sigma_B\colon [B, A+B] \to \mathcal{L}(A+B/B)$. Combined with the isomorphism $\phi\colon [A \cap B, A] \to [B, A+B]$ from 7.12, we have the diagram

$$
\begin{tikzcd}
{[A \cap B, A]} \arrow[r, "\phi"] \arrow[d, "\sigma_{A \cap B}"] & {[B, A + B]} \arrow[d, "\sigma_B"] \\
\mathcal{L}(A/(A \cap B)) \arrow[r, "\psi"] & \mathcal{L}((A + B)/B)
\end{tikzcd}
$$

It is easy to check that this diagram commutes.

8 Cancellation and Separativity

In this section we will investigate the relationship $a + c = b + c$ for elements a, b, c of a monoid.

We begin by gathering in the next few propositions some simple facts about the relationship $a + c = b + c$ that we will use repeatedly in this section:

Proposition 8.1. *Let a, b, c, c' be elements of a monoid such that $a+c = b+c$. Then*

1. $c \leq c' \implies a + c' = b + c'$
2. $c \leq a \implies 2a = a + b$
3. $c \leq a, b \implies 2a = a + b = 2b$

Proof. To prove 1, let x be such that $c' = c+x$. Then $a+c' = a+c+x = b+c+x = b+c'$. Statements 2 and 3 are easy consequences of 1. □

Notice that the equation $2a = a+b$ is a special case of $a+c = b+c$ with $c \leq a$. A similar statement is not true in general about the two situations $2a = a+b = 2b$, and $a+c = b+c$ with $c \leq a, b$, unless the monoid has decomposition.

In this and later sections, we will make frequent use of 8.1.1 without reference. For example, we use it three times in the following proposition:

Proposition 8.2. *Let a, b, c be elements of a monoid M such that $a + c = b + c$.*

1. *If there is an element $d \in M$ such that $c \ll d \leq a, b$, then $a = b$.*
2. *If $c \ll c \leq a$, then there is some idempotent $e \equiv c$ such that $a = b + e$.*

Proof.

1. Let a', b' be such that $a = d + a'$ and $b = d + b'$. Then $a' + d + c = b' + d + c$. Since $c + d \leq d$, we get $a' + d = b' + d$, that is $a = b$.
2. We have $c \ll c$, so from 6.9, there is some idempotent $e \equiv c$. Since $c \leq e$ and $a + c = b + c$, we get $a + e = b + e$. And since $e \leq c \leq a$ we get $a + e = a$. Thus $a = b + e$.

□

Note that, in particular, if c is a regular element such that $c \leq a, b$, then $a + c = b + c$ implies $a = b$. Next we collect some simple consequences of having a refinement of the equation $a + c = b + c$.

Proposition 8.3. *Let a, b, c be elements of a monoid such that $a + c = b + c$ and let*

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

be a refinement of this equation. Then $a + b_1 = b + a_1$, $a_1 + c = b_1 + c$, and

1. $c \leq a \implies a + a_1 = b + a_1$

2. $c \leq a, b \implies a + a_1 = b + a_1 = a + b_1 = b + b_1$

Proof. The only non-trivial statement is that $c \leq a$ implies $a + a_1 = b + a_1$: We have $c_1 \leq c \leq a$ so by 8.1.1, $a_1 + c_1 = b_1 + c_1$ implies that $a + a_1 = a + b_1$. □

Notice that the existence of a refinement of the equation $a + c = b + c$ gives rise to a similar equation $a_1 + c_1 = b_1 + c_1$. In a refinement monoid this process can be repeated to produce a decreasing chain of such equations, $a_i + c_i = b_i + c_i$ with $i = 1, 2, 3, \ldots$, such that $a \geq a_1 \geq a_2 \geq \ldots$, $b \geq b_1 \geq b_2 \geq \ldots$ and $c \geq c_1 \geq c_2 \geq \ldots$. This fact will inspire the investigation of Artinian refinement monoids in Section 12.

In addition, we get the following proposition, which says, in effect, that if $a_0+c_0 = b_0+c_0$, then a_0 and b_0 differ by elements which can be chosen "arbitrarily small" in comparison to c_0. The quotation marks reflect the fact that the inequality $na_n \leq c_0$ which appears in the proposition, does not even forbid that $a_n \geq c_0$ if c_0 is regular.

Proposition 8.4. *Let M be a refinement monoid, $a_0, b_0, c_0 \in M$ such that $a_0+c_0 = b_0+c_0$ and $n \in \mathbb{N}$. Then there are $a_n, b_n, c_n, c_{n-1}, d_n \in M$ such that $na_n, nb_n, c_n \leq c_0$, and the following is a refinement matrix*

$$\begin{array}{c|cc} & b_0 & c_{n-1} \\ \hline a_0 & d_n & a_n \\ c_{n-1} & b_n & c_n \end{array}$$

In particular, $a_0 + c_n = b_0 + c_n$.

Proof. [31, 1.11] Define inductively $a_i, b_i, c_i, d'_i \in M$ for $i = 1, 2, \ldots, n$ by making refinements of the form

$$\begin{array}{c|cc} & b_i & c_i \\ \hline a_i & d'_{i+1} & a_{i+1} \\ c_i & b_{i+1} & c_{i+1} \end{array}$$

By an easy induction we get $a_0 = a_n + \sum_{i=1}^n d'_i$, $b_0 = b_n + \sum_{i=1}^n d'_i$, $c_0 = c_n + \sum_{i=1}^n a_i = c_n + \sum_{i=1}^n b_i$. Setting $d_n = \sum_{i=1}^n d'_i$ gives $a_0 = d_n + a_n$, and $b_0 = d_n + b_n$. And since $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$, we have $na_n \leq c_0$, $nb_n \leq c_0$. Note also that since $a_n + c_n = b_n + c_n$, we get $a_0 + c_n = b_0 + c_n$. □

It is worth noting one easy consequence of this lemma: If $c_0 \in M$ and $n \in \mathbb{N}$ are such that for any $x \in M$, $nx \leq c_0$ implies $x = 0$, then c_0 cancels from $a_0 + c_0 = b_0 + c_0$.

In the previous proposition, we started with the equation $a_0+c_0 = b_0+c_0$ in a refinement monoid. If, in addition, we have $c_0 \leq a_0$, we will get a similar but much stronger result, (8.6). To prove this, we need the following lemma:

Lemma 8.5. *Let M be a refinement monoid, $a_0, b_0, c_0 \in M$ such that $a_0 + c_0 = b_0 + c_0$.*

1. *If $c_0 \leq a_0$, then there is a refinement matrix*

$$\begin{array}{c|cc} & b_0 & c_0 \\ \hline a_0 & d_1 & a_1 \\ c_0 & b_1 & c_1 \end{array}$$

such that $c_1 \leq a_1$.

2. *If $c_0 \leq a_0, b_0$, then there is a refinement matrix as above such that $c_1 \leq a_1, b_1$.*

Proof.

1. From $a_0 + c_0 = b_0 + c_0$ we get a refinement matrix of the form

$$\begin{array}{c|cc} & b_0 & c_0 \\ \hline a_0 & d' & a' \\ c_0 & b' & c' \end{array}$$

Since $c' \leq c_0 \leq a_0 = d' + a'$, we can write $c' = d'' + c_1$ where $d'' \leq d'$ and $c_1 \leq a'$. Since $d'' \leq d'$, we can write $d' = d'' + d_1$, giving the following refinement:

$$\begin{array}{c|cc} & b_0 & c_0 \\ \hline a_0 & d'' + d_1 & a' \\ c_0 & b' & d'' + c_1 \end{array}$$

Moving the d'' terms gives refinement matrix

$$\begin{array}{c|cc} & b_0 & c_0 \\ \hline a_0 & d_1 & d'' + a' \\ c_0 & d'' + b' & c_1 \end{array}$$

Setting $a_1 = d'' + a'$ and $b_1 = d'' + b'$ gives the required refinement matrix. Further, we have $c_1 \leq a' \leq a' + d'' = a_1$. For use in the proof of 2, we note that $c_1 \leq c'$ and $b_1 \geq b'$.

2. Since $c_0 \leq b_0$ we can use 1 (with the roles of a_0 and b_0 interchanged) to get the refinement matrix

$$\begin{array}{c|cc} & b_0 & c_0 \\ \hline a_0 & d' & a' \\ c_0 & b' & c' \end{array}$$

with $c' \leq b'$.

Now a repetition of the argument of 1 using $c_0 \leq a_0$ gives a new refinement matrix

$$\begin{array}{c|cc} & b_0 & c_0 \\ \hline a_0 & d_1 & a_1 \\ c_0 & b_1 & c_1 \end{array}$$

with $c_1 \leq a_1$. In addition, we have $c_1 \leq c' \leq b' \leq b_1$.

□

Proposition 8.6. *Let M be a refinement monoid, $a_0, b_0, c_0 \in M$ such that $a_0 + c_0 = b_0 + c_0$, $c_0 \leq a_0$ and $n \in N$. Then there are $a_n, b_n, c_n, c_{n-1}, d_n \in M$ such that $na_n, nb_n, nc_n \leq c_0$ and $c_n \leq a_n$, and the following is a refinement matrix*

$$\begin{array}{c|cc} & b_0 & c_{n-1} \\ \hline a_0 & d_n & a_n \\ c_{n-1} & b_n & c_n \end{array}$$

In particular, $a_0 + c_n = b_0 + c_n$ with $nc_n \leq c_0$.

Proof. The proof proceeds exactly as the proof of 8.4, except that at each induction step we use 8.5.1 to give a refinement matrix such that $c_{i+1} \leq a_{i+1}$. Thus, in addition to the conclusions given in 8.4, we have $c_n \leq a_n$, and also $nc_n \leq na_n \leq c_0$. □

We turn away now from the few cancellation rules that are true in any refinement monoid, to cancellation properties that occur in the special classes of monoids (primely generated, Artinian and semi-Artinian) we will study later. There are five such cancellation properties we need: Three of these we will study in this section, the other two will be the subject of the next section.

The following properties are usefully defined for subclasses of semigroups (rather than only for monoids or order ideals):

Definition 8.7.

- *A subclass X of a semigroup is* **cancellative** *if*
$$(\forall a, b, c \in X)\ (a + c = b + c \implies a = b).$$
- *A subclass X of a semigroup is* **strongly separative** [2, Section 5] *if*
$$(\forall a, b \in X)\ (2a = a + b \implies a = b).$$
- *A subclass X of a semigroup is* **separative** [7, Chapter 4.3] *if*
$$(\forall a, b \in X)\ (2a = a + b = 2b \implies a = b).$$

Clearly, for any subclass X we have the following implications:
$$X \text{ cancellative} \implies X \text{ strongly separative} \implies X \text{ separative}.$$
Also, any subclass of a cancellative (strongly separative, separative) subclass is also cancellative (strongly separative, separative).

The monoid $\{0, \infty\}$ has refinement but is not cancellative. For the reversed situation we have:

Example 8.8. *A monoid which is cancellative but does not have decomposition or refinement:*

Let $M = \{0, 2, 3, 4, \ldots\}$ be the submonoid of $\mathbb{Z}^+$ obtained by deleting the number 1. Then, since M is a subset of $\mathbb{Z}^+$, it is cancellative. On the other hand, the inequality $2 \leq 3 + 3$ can not be decomposed in M since $2 \not\leq 3$.

If $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of monoids then it is straight forward to show that $\prod_{\alpha\in\mathcal{I}} M_\alpha$ and $\bigoplus_{\alpha\in\mathcal{I}} M_\alpha$ are cancellative (strongly separative, separative) if and only if M_α is cancellative (strongly separative, separative) for each $\alpha \in \mathcal{I}$.

One can check that if M is a separative semigroup then M^0 is a separative monoid. A similar statement is not true about either cancellative or strongly separative semigroups. For example, if M is the trivial one element semigroup, then M is trivially cancellative and so strongly separative, but $M^0 \cong \{0, \infty\}$ is neither cancellative nor strongly separative.

In the next few lemmas, we give other properties equivalent to separativity and strong separativity. First we will consider separative monoids:

Proposition 8.9. *For a monoid M, the following are equivalent:*

1. $(\forall a, b, c \in M)$ $(a + c = b + c$ and $c \leq a, b \implies a = b)$
2. $(\forall a, b, c \in M)(\forall n \in \mathbb{N})$ $(a + nc = b + nc \implies a + c = b + c)$
3. $(\forall a, b, c, d \in M)$ $(a + c = b + c$ and $c \prec d \leq a, b \implies a = b)$

Proof.

1 ⇒ 2 The $n = 1$ case is trivial, so assume $n \geq 2$, then $a+nc = b+nc \implies (a+(n-1)c)+c = (b + (n-1)c) + c$, and, since $c \leq a + (n-1)c$ and $c \leq b + (n-1)c$, this implies $a + (n-1)c = b + (n-1)c$. A simple induction then shows that $a = b$.

2 ⇒ 3 Since $c \prec d$, there is some $n \in \mathbb{N}$ such that $c \leq nd$, and hence $a + nd = b + nd$. There are also a' and b' such that $a = d + a'$ and $b = d + b'$, so $a' + (n+1)d = b' + (n+1)d$. Using 2, we cancel nd from this equation to get $a' + d = b' + d$, that is, $a = b$.

3 ⇒ 1 1 is just the special case of 3 in which $d = c$.

□

Proposition 8.10. *For a monoid M, the following are equivalent:*

1. *M is separative.*
2. $(\forall a, b \in M)(\forall m, n \in \mathbb{N})$ $(ma = mb$ and $na = nb \implies ka = kb$ where $k = \gcd(m, n))$
3. *There are $m, n \geq 2$ such that $\gcd(m, n) = 1$, and*
$$(\forall a, b \in M)\ (ma = mb \text{ and } na = nb \implies a = b).$$
4. $(\forall a, b, c \in M)$ $(a + c = b + c$ and $c \prec a$ and $c \prec b \implies a = b)$
5. *Each Archimedean component of M is cancellative.*

Further, any of these properties imply those of Proposition 8.9.

Proof. First we show that 1 implies property 1 of 8.9 ...

If $a + c = b + c$ with $c \leq a$ and $c \leq b$, then from 8.1.3, $2a = a + b = 2b$. Thus $a = b$.

1 ⇒ 2 Without loss of generality, we can assume $m \geq n$ and $m = hn + r$ for suitable $h, r \in \mathbb{Z}^+$ with $r < n$. By the Euclidean algorithm for calculating the gcd of m and n, it suffices to show that $na = nb$ and $ma = mb$ imply $ra = rb$...

If $r = 0$ there is nothing to prove, so we assume $r \geq 1$, then $ma = mb$ implies $hna + ra = hnb + rb$ and $hna + ra = hna + rb$. Now, using 2 of 8.9, this implies $a + ra = a + rb$, and adding $(r-1)a$ to each side we get $2ra = ra + rb$. By symmetry, $2rb = rb + ra$, and so using the hypothesis, $ra = rb$.

2 ⇒ 3 Trivial.

3 ⇒ 1 If $2a = a + b = 2b$, then by an easy induction, $ka = kb$ for all $k \geq 2$. In particular, $ma = mb$ and $na = nb$, so by 3, $a = b$.

We have now shown the equivalence of 1, 2, and 3, and that any of these imply the properties of Proposition 8.9.

2 ⇒ 4 Suppose $a + c = b + c$ and $n \in \mathbb{N}$ such that $c \leq na$ and $c \leq nb$. Then there are $x, y \in M$ such that $na = c + x$ and $nb = c + y$, and so $n(a + c) = n(b + c)$ implies $c + x + nc = c + y + nc$. Using 8.9 we get $c + x = c + y$, that is, $na = nb$. This same argument shows that $(n + 1)a = (n + 1)b$, and so, using the hypothesis, $a = b$.

4 ⇒ 5 Suppose a, b, c are in the same Archimedean component of M and $a + c = b + c$. Then in particular, $c \prec a$ and $c \prec b$, so by 4, $a = b$.

5 ⇒ 1 If $2a = a + b = 2b$ then $a \leq 2b$ and $b \leq 2a$, so a and b are in the same Archimedean component of M. Cancellation in this Archimedean component then gives $a = b$.

□

In general, the properties of 8.9 are weaker than separativity, but when the monoid has decomposition, they are equivalent:

Proposition 8.11. *If M is a decomposition monoid, then any of the properties of 8.9 is equivalent to separativity.*

Proof. From 8.10, we have already that separativity implies the properties of 8.9. For the converse, suppose we have $2a = a + b = 2b$ for some $a, b \in M$. Since $a \leq 2b$, there are $a_1, a_2 \in M$ such that $a_1, a_2 \leq b$ and $a = a_1 + a_2$. Now $a + a = a + b$ implies $a + a_1 + a_2 = b + a_1 + a_2$, and since $a_1, a_2 \leq a, b$ we can cancel a_1 and a_2 from this to get $a = b$. □

We now consider strong separativity. Here the situation is much simpler:

Proposition 8.12. *For a monoid M, the following are equivalent:*

1. *M is strongly separative*
2. $(\forall a, b, c \in M)\ (a + 2c = b + c \implies a + c = b)$
3. $(\forall a, b, c \in M)(\forall n \in \mathbb{N})\ (a + (n+1)c = b + nc \implies a + c = b)$
4. $(\forall a, b \in M)(\forall n \in \mathbb{N})\ ((n+1)a = na + b \implies a = b)$
5. $(\forall a, b, c \in M)\ (a + c = b + c \text{ and } c \leq a \implies a = b)$
6. $(\forall a, b, c \in M)\ (a + c = b + c \text{ and } c \prec a \implies a = b)$

Proof. The equivalence of 1, 2, 3, 4 and 5 is easy to prove, and 5 is just a special case of 6, so we prove here only that 4 implies 6 ...

Suppose $a + c = b + c$ with $c \leq na$ for some $n \leq \mathbb{N}$. Then $a + na = b + na$ and using 4, we get $a = b$. □

In a refinement monoid, the question of whether it is possible to cancel c from the relation $a + c = b + c$ often depends only on the properties of the subclass $\{\leq c\}$, rather than the properties of the whole monoid. As a prototype of this situation we have

Proposition 8.13. *Let a, b, c be elements of a refinement monoid such that $a + c = b + c$. If $\{\leq c\}$ is cancellative, then $a = b$.*

Proof. From the equation $a + c = b + c$ we get the refinement matrix

$$\begin{array}{c|cc} & b & c \\ \hline a & d' & a' \\ c & b' & c' \end{array}$$

a', b', c' are all in $\{\leq c\}$, so the equation $a' + c' = b' + c' = c$ cancels to give $a' = b'$. Hence $a = a' + d' = b' + d' = b$. □

Using Lemma 8.5 we can derive theorems about separativity similar to 8.13:

Proposition 8.14. *Let a, b and c be elements of a refinement monoid M such that $a + c = b + c$.*

1. *If $\{\leq c\}$ is strongly separative and $c \prec a$, then $a = b$.*
2. *If $\{\leq c\}$ is strongly separative and $c \prec a + b$, then $a = b$.*
3. *If $\{\leq c\}$ is separative, $c \prec a$ and $c \prec b$, then $a = b$.*

Proof.

1. Since $c \prec a$, there is some $n \in \mathbb{N}$ such that $c \leq na$. We will do the $n = 1$ case first...

$n = 1$ Since $c \leq a$ we can use 8.5.1 to get a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c_1 \leq a_1$. Since $a_1 + c_1 = b_1 + c_1$, this implies $2a_1 = a_1 + b_1$. The elements a_1 and b_1 are in the strongly separative subclass $\{\leq c\}$, and so $a_1 = b_1$. Thus $a = d_1 + a_1 = d_1 + b_1 = b$.

$n > 1$ Since $c \leq na$, we can write $c = \sum_{i=1}^{n} c_i$ where $c_i \leq a$ for all i. Thus

$$a + \sum_{i=1}^{n} c_i = b + \sum_{i=1}^{n} c_i.$$

The subclasses $\{\leq c_i\}$ are contained in $\{\leq c\}$ so they are strongly separative. Using the $n = 1$ case, the c_i can be canceled from the equation one by one to leave $a = b$.

2. We have $c \prec a + b$, so there is some $n \in \mathbb{N}$ such that $c \leq na + nb$, and also $c_1 \leq na$ and $c_2 \leq nb$ such that $c = c_1 + c_2$ and $a + c_1 + c_2 = b + c_1 + c_2$. The subclasses $\{\leq c_1\}$ and $\{\leq c_2\}$ are contained in $\{\leq c\}$ and so they are strongly separative. Since also $c_1 \prec a$ and $c_2 \prec b$, we can apply 1 to cancel c_1 and c_2 from the equation to give $a = b$.
3. Since $c \prec a, b$, there is some $n \in \mathbb{N}$ such that $c \leq na, nb$. We will do the $n = 1$ case first...

$n = 1$ Since $c \leq a, b$ we can use 8.5.2 to get a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c_1 \leq a_1, b_1$. Since $a_1 + c_1 = b_1 + c_1$, this implies $2a_1 = a_1 + b_1 = 2a_1$. The elements a_1 and b_1 are in the separative subclass $\{\leq c\}$, and so $a_1 = b_1$. Thus $a = d_1 + a_1 = d_1 + b_1 = b$.

$n > 1$ Since $c \leq na$, we can write $c = \sum_{i=1}^{n} c_i$ where $c_i \leq a$ for all i. For $i = 1, 2, \ldots, n$, we have $c_i \leq c \leq nb$, so we can write $c_i = \sum_{j=1}^{n} c_{ij}$ where now we have $c_{ij} \leq a, b, c$ for all i and j. Thus $c = \sum_{ij} c_{ij}$ and

$$a + \sum_{ij} c_{ij} = b + \sum_{ij} c_{ij}.$$

The subclasses $\{\leq c_{ij}\}$ are contained in $\{\leq c\}$ so they are separative. Using the $n = 1$ case, the c_{ij} can be canceled from the equation one by one to leave $a = b$.

$\square$

We next consider how the cancellation properties of a monoid effect those of quotient monoids, and vice versa.

Proposition 8.15. *Let I be a submonoid of monoid M. If M is cancellative (strongly separative, separative), then so are I and M/I.*

Proof. The claim that I has the same cancellation properties as M is clear, so we prove only the claims about M/I:

1. Suppose M is cancellative and $[a]_I + [c]_I = [b]_I + [c]_I$. Then there are $u, v \in I$ such that $a + c + u = b + c + v$. Cancellation in M gives $a + u = b + v$, so $[a]_I = [b]_I$.
2. Suppose M is strongly separative and $2[a]_I = [a]_I + [b]_I$. Then there are $u, v \in I$ such that $2a + u = a + b + v$. Thus $2(a + u) = (a + u) + (b + v)$, and using strong separativity, $a + u = b + v$. Hence $[a]_I = [b]_I$.
3. Suppose M is separative and $2[a]_I = [a]_I + [b]_I = 2[b]_I$. From the second equality, we get the inequality $[a]_I \leq 2[b]_I$, so there is $w \in I$ such that $a \leq 2b + w$. Thus $a \leq 2(b + w)$ and $a \prec b + w$.

   From the equality, $2[a]_I = [a]_I + [b]_I$, there are $u, v \in I$ such that $2a + u = a + b + v$. Thus $(a + u + w) + a = (b + v + w) + a$ with $a \prec a + u + w$ and $a \prec b + v + w$. Using separativity and 8.10.4, we get $a + u + w = b + v + w$, and since $u + w, v + w \in I$, $[a]_I = [b]_I$.

$\square$

The converse of this proposition is true in the strongly separative and separative cases if M has refinement and I is an order ideal:

Proposition 8.16. *Let I be an order ideal in a refinement monoid M. If I and M/I are (strongly separative) separative, then so is M.*

Proof. [2, Theorem 4.5]

1. Suppose I and M/I are strongly separative and $a, b, c \in M$ such that $a + c = b + c$ and $c \leq a$. We will show that $a = b$ ...

   From 8.5.1, there is a refinement matrix

$$\begin{array}{c|cc} & b & c \\ \hline a & d_1 & a_1 \\ c & b_1 & c_1 \end{array}$$

   with $c_1 \leq a_1$. In M/I we get $[a_1]_I + [c_1]_I = [b_1]_I + [c_1]_I$ with $[c_1]_I \leq [a_1]_I$, so using the strong separativity of M/I, $[a_1]_I = [b_1]_I$. Thus there are $u_1, v_1 \in I$ such that $a_1 + u_1 = b_1 + v_1$. From a refinement of this relationship,

$$\begin{array}{c|cc} & b_1 & v_1 \\ \hline a_1 & d_2 & v_2 \\ u_1 & u_2 & e_2 \end{array}$$

   we get the equation $a_1 + u_2 = b_1 + v_2$ and hence, $c + u_2 = c_1 + a_1 + u_2 = c_1 + b_1 + v_2 = c + v_2$. Since $c \leq a$, we have $a + u_2 = a + v_2$ and further

$$\begin{aligned} a + v_2 &= a + u_2 = d_1 + a_1 + u_2 = d_1 + b_1 + v_2 \\ &= b + v_2. \end{aligned}$$

   Note that $v_2 \leq v_1 \in I$ so $v_2 \in I$, and also $v_2 \leq a_1 \leq a$. Since $\{\leq v_2\} \subseteq I$ is strongly separative, we can use 8.14.1 to get $a = b$.

2. The proof in the separative case is similar to the strongly separative case...

   Suppose I and M/I are separative and $a, b, c \in M$ such that $a + c = b + c$ and $c \leq a, b$. We will show that $a = b$ ...

   From 8.5.2, there is a refinement matrix as above with $c_1 \leq a_1, b_1$. In M/I we get $[a_1]_I + [c_1]_I = [b_1]_I + [c_1]_I$ with $[c_1]_I \leq [a_1]_I, [b_1]_I$, so using the separativity of M/I, $[a_1]_I = [b_1]_I$. Thus there are $u_1, v_1 \in I$ such that $a_1 + u_1 = b_1 + v_1$. Exactly as above, we make a refinement of this equation and deduce that $a + v_2 = b + v_2$ with $\{\leq v_2\}$ separative. Since $v_2 \leq a_1 \leq c \leq a, b$ we can use 8.14.3 to cancel the v_2 and get $a = b$.

□

If I and M/I are cancellative then it is not necessarily true that M is cancellative, even if M has refinement. One example of this is 15.9.

Definition 8.17. *An element $a \in M$ of a monoid is* **free** *if for all $m, n \in \mathbb{N}$,*

$$ma \leq na \implies m \leq n.$$

A free element is not regular and regular elements are not free, but, in general there are elements which are neither. For example, in the monoid $M = \{0, 1, \infty\}$ such that $1+1 = \infty$, the element 1 is neither regular nor free.

In a separative monoid, however, every element is either regular or free:

Proposition 8.18. *Let a be an element of a monoid M.*

1. *If M is separative then a is free or regular.*
2. *If M is strongly separative then a is free or $a \leq 0$.*

Proof.

1. If M is separative and a is not free, then there are $m, n \in \mathbb{N}$ such that $m > n$ and $ma \leq na$. Since $n \geq 1$, we can use 8.9 and 8.10 to cancel a from this inequality until we get$(m - n + 1)a \leq a$. Since $m - n + 1 \geq 2$, this implies that a is regular.
2. If M is strongly separative, then it is separative and a is either free or regular. But if $2a \leq a$, then $2a + x = a$ for some $x \in M$ and using 8.12.2, we get $a + x = 0$. In particular, $a \leq 0$.

□

From this proposition, we have that if M is a strongly separative monoid then M is separative and has no proper regular elements. The converse of this statement is not true even for refinement monoids. For a counterexample, see 9.7.

Another consequence of this proposition is that every element of a finite separative monoid is regular: Given an element a of such a monoid, the elements $a, 2a, 3a, \ldots$ can not all be distinct, hence a is not free.

9 Weak Cancellation and Midseparativity

In the previous section, we defined cancellation, separativity, and strong separativity. In this section we will add to this list two other cancellation properties. Primely generated refinement monoids, which we will study in Section 11, are examples of monoids with these properties. The new cancellation properties are distinguished from the old ones because they contain existential quantifiers in their definitions.

Definition 9.1.

- *A semigroup M is* **weakly cancellative** *if $a+c=b+c$ for some $a,b,c \in M$ implies the existence of a refinement matrix*

$$\begin{array}{c c} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

  with $c \le c_1$. Note that this implies that $c \equiv c_1$ and $a_1, b_1 \ll c$.
- *A semigroup M is* **midseparative** *if*

  $(\forall a,b \in M)$ $(2a = a+b \implies \exists$ idempotent $e \in M$ such that $a = b+e)$.

  Note that $2e = e$ and $a = b+e$ imply that $a = a+e$.

A partially ordered weakly cancellative refinement monoid is called a strong refinement monoid by F. Wehrung [31].

Among the simple consequences of the definitions are: If M is a weakly cancellative (midseparative) monoid and $I \le M$ is an order ideal, then I is also weakly cancellative (midseparative). If $\{M_\alpha \mid \alpha \in \mathcal{I}\}$ is a family of monoids then $\prod_{\alpha\in\mathcal{I}} M_\alpha$ and $\bigoplus_{\alpha\in\mathcal{I}} M_\alpha$ are weakly cancellative (midseparative) if and only if M_α is weakly cancellative (midseparative) for each $\alpha \in \mathcal{I}$.

One can check that if M is a weakly cancellative (midseparative) semigroup then M^0 is a weakly cancellative (midseparative) monoid.

Similar to separativity and strong separativity, the midseparativity condition can be expressed in several equivalent ways. Note that in this proposition, unlike in 8.10 and 8.12, we assume that M has refinement.

Proposition 9.2. *Let M be a refinement monoid. Then the following are equivalent:*

1. *M is midseparative.*
2. $(\forall a,b,c \in M)$ $(a+c = b+c$ and $c \le a \implies \exists$ idempotent $e \le c$ such that $a = b+e)$
3. $(\forall a,b,c \in M)$ $(a+c = b+c$ and $c \prec a \implies \exists$ idempotent $e \le c$ such that $a = b+e)$

Proof.

$1 \Rightarrow 2$ If $a + c = b + c$ and $c \leq a$ then by 8.5.1, there is a refinement

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c_1 \leq a_1$. From $a_1 + c_1 = b_1 + c_1$, we get $2a_1 = a_1 + b_1$, so there is some idempotent $e \in M$ such that $a_1 = b_1 + e$. Thus $e \leq a_1 \leq c$ and $a = d_1 + a_1 = d_1 + b_1 + e = b + e$.

$2 \Rightarrow 3$ Since $c \prec a$, there is some $n \in \mathbb{N}$ such that $c \leq na$. We will proceed by induction on n. Since the $n = 1$ case is our hypothesis, we need show only the induction step...

Suppose the claim is true for some $n \in \mathbb{N}$, and $a + c = b + c$ with $c \leq (n+1)a$. We can write $c = c_1 + c_2$ with $c_1 \leq a$ and $c_2 \leq na$, and so $(a + c_1) + c_2 = (b + c_1) + c_2$ with $c_2 \leq n(a + c_1)$. By induction, there is $e_2 = 2e_2 \leq c_2$ such that $a + c_1 = (b + e_2) + c_1$. Since $c_1 \leq a$, the hypothesis provides $e_1 = 2e_1 \leq c_1$ such that $a = b + (e_1 + e_2)$. We have also $2(e_1 + e_2) = e_1 + e_2 \leq c_1 + c_2 = c$, so the induction is complete.

$3 \Rightarrow 1$ Set $c = a$ in 3.

□

Proposition 9.3. *For a monoid M we have the following implications:*

$$\begin{array}{ccccc} \textit{cancellative} & & \Longrightarrow & & \textit{weakly cancellative} \\ \Downarrow & & & & \Downarrow \\ \textit{strongly separative} & \Longrightarrow & \textit{midseparative} & \Longrightarrow & \textit{separative} \end{array}$$

Proof. All of the implications are easy except the following:

- Suppose M is weakly cancellative, and $a + c = b + c$ with $c \leq a, b$ for some $a, b, c \in M$. Then from the definition, there is a refinement

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $a_1 \ll c \leq a, b$. From 8.3.1, $a + a_1 = b + a_1$, and so using 8.2.1, we get $a = b$.
- Suppose M is midseparative, and $2a = a + b = 2b$ for some $a, b \in M$. From the definition, there are idempotents $e, f \in M$ such that $a = b + e$ and $b = a + f$. Thus $a = a + e$ and, since $a \leq b$, we have $b + e = b$. Thus $a = b + e = b$

□

Note that we need M to be a monoid, rather than just a semigroup, in this proposition so that strong separativity implies midseparativity, and cancellation implies weak cancellation.

Proposition 9.4. *Let M be a monoid such that*

$$(\forall a, b \in M)\ (a \leq b \text{ or } b \leq a).$$

Then the following are equivalent:

1. *M is separative.*
2. *M has refinement.*
3. *M is weakly cancellative.*

Proof.

$1 \Rightarrow 2$ Suppose $a_1 + a_2 = b_1 + b_2$ in M. Without loss of generality, we can assume that a_1 is minimal in the set $\{a_1, a_2, b_1, b_2\}$. Our hypothesis then implies that $a_1 \leq a_2, b_1, b_2$. In particular, $b_1 = a_1 + x_1$ for some $x_1 \in M$. Thus we have the equation $a_1 + a_2 = a_1 + (x_1 + b_2)$ with $a_1 \leq a_2, b_2 + x_1$. From 8.1.3, this implies that $2a_2 = a_2 + (b_2 + x_1) = 2(b_2 + x_1)$. Since M is separative, we have $a_2 = b_2 + x_1$ and then

$$\begin{array}{c c} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} a_1 & 0 \\ x_1 & b_2 \end{pmatrix} \end{array}$$

is a refinement of the original equation.

$2 \Rightarrow 3$ Suppose $a + c = b + c$ in M. We make a refinement of this equation:

$$\begin{array}{c c} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

Without loss of generality, we can assume $a_1 \leq b_1$. Thus there is some $x_1 \in M$ such that $b_1 = a_1 + x_1$. This gives a new refinement

$$\begin{array}{c c} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 + a_1 & 0 \\ x_1 & c_1 + a_1 \end{pmatrix} \end{array}$$

Thus M has weak cancellation.

$3 \Rightarrow 1$ This we have shown in 9.3.

□

There is another separativity property which serves as a "greatest common denominator" of weak cancellation and midseparativity, in the sense that it is implied by both, and it in turn implies separativity:

Proposition 9.5. *Let M be a refinement monoid. Then the following are equivalent:*

1. $(\forall a, b \in M)$ $(2a = a + b \implies b \leq a)$
2. $(\forall a, b, c \in M)$ $(a + c = b + c$ and $c \leq a \implies \exists x$ such that $c = c + x$ and $a = b + x)$
3. $(\forall a, b, c \in M)$ $(a + c = b + c$ and $c \leq a \implies b \leq a)$

If M is either weakly cancellative or midseparative then M has these properties. Further, any of these properties implies that M is separative.

Proof.

$1 \Rightarrow 2$ If $a + c = b + c$ and $c \leq a$, then by 8.5.1, there is a refinement

$$\begin{array}{c c} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c_1 \leq a_1$. Since $a_1 + c_1 = b_1 + c_1$, we have $2a_1 = a_1 + b_1$. By hypothesis this implies $b_1 \leq a_1$, thus $a_1 = b_1 + x$ for some $x \in M$. Hence $c + x = c_1 + b_1 + x = c_1 + a_1 = c$ and $b + x = d_1 + b_1 + x = d_1 + a_1 = a$.

$2 \Rightarrow 3$ Trivial.

$3 \Rightarrow 1$ Set $c = a$ in 3.

We have shown therefore the equivalence of 1, 2 and 3. We show next that weak cancellation implies 3…

Suppose M is weakly cancellative and $a+c = b+c$ with $c \le a$. Then there is a refinement

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $a_1 \ll c$. Since $c \le a$ we get $a_1 \ll a$ and, using 8.3.2, $a+a_1 = b+a_1$. Thus $b \le a+a_1 \le a$.

If M is midseparative, then 1 follows immediately.

Finally, we show that 2 implies separativity…

Suppose $a + c = b + c$ with $c \le a, b$. From 2, there is $x \in M$ such that $c = c + x$ and $a = b + x$. Since $c \le b$ we have $b = b + x$, and hence $a = b + x = b$. □

Unlike separative monoids, midseparative monoids are strongly separative if they have no proper regular elements:

Proposition 9.6. *A midseparative monoid is strongly separative if and only if it has no proper regular elements.*

Proof. Let M be a midseparative monoid with no proper regular elements, and $a, b \in M$ such that $2a = a + b$. Then there is some idempotent $e \in M$ such that $a = b + e = a + e$. Since e is regular, we have $e \le 0$, so by 8.2.1, $a = b$. Thus M is strongly separative.

The converse follows directly from 8.18.2. □

We will show by example that the implications in 9.3 are the strongest possible. In particular, not all weakly cancellative monoids are midseparative or vice versa – even for refinement monoids. For a midseparative refinement monoid which is not weakly cancellative, see 15.9.

To construct a weakly cancellative refinement monoid which is not midseparative, we proceed via the following example which has interesting properties of its own:

Example 9.7. *A separative refinement monoid with no proper regular elements which is not weakly cancellative, strongly separative or midseparative:*

Let M be the semigroup direct product $\{0,\infty\} \times \mathbb{R}^{++}$. Since $\{0,\infty\}$ and $\mathbb{R}^{++}$ are separative refinement semigroups, so is M, and since $\mathbb{R}^{++}$ has no regular elements, neither does M.

Let M^0 be the monoid obtained from M by adding a zero element. M^0 is a separative refinement monoid with no proper regular elements. It is easy to see that M^0 can be considered to be the monoid $(\{0,\infty\} \times \mathbb{R}^+) \setminus \{(\infty, 0)\}$ though it is clumsier to prove refinement in this form.

If we set $a = (\infty, 1)$ and $b = (0, 1)$, then $2a = a + b$, but $a \ne b$ and, even stronger, there is no x such that $a = b + x$, that is $b \not\le a$. Thus M^0 does not have the property of 9.5.1 and so, by that proposition, M^0 is not midseparative or weakly cancellative.

For future reference we will note the following properties of this monoid:

Let $I = \{\prec (0,1)\}$. This order ideal is isomorphic to $\mathbb{R}^+$ so is cancellative. The quotient M^0/I is easily seen to be isomorphic to the monoid $\{0,\infty\}$. So M^0/I has a proper regular element even though M^0 does not.

Also I and M^0/I are midseparative, but M^0 itself is not.

If $\mathbb{R}^{++}$ was weakly cancellative, then the example we have just constructed would be a weakly cancellative non-midseparative refinement monoid. Thus, to produce our desired example, we will first construct a semigroup which is like $\mathbb{R}^{++}$ but has, in addition, weak cancellation.

Example 9.8. *A weakly cancellative refinement semigroup with no regular elements:*
Let $N = (\mathbb{R}^{++})^2$ as a set, with addition defined by

$$(a_1, a_2) + (b_1, b_2) = \begin{cases} (a_1, a_2) & a_1 > b_1 \\ (b_1, b_2) & a_1 < b_1 \\ (a_1, a_2 + b_2) & a_1 = b_1 \end{cases}$$

for all $(a_1, a_2), (b_1, b_2) \in N$.

It is easily seen that N is a semigroup with no regular elements. In the following discussion the symbol $*$ as an entry in a refinement matrix denotes an element of N which adds nothing to either the column sum or row sum in which it appears. Such elements are always available since, for any elements $(a_1, a_2), (b_1, b_2) \in N$, there is another element (c_1, c_2) such that $(a_1, a_2) + (c_1, c_2) = (a_1, a_2)$ and $(b_1, b_2) + (c_1, c_2) = (b_1, b_2)$: For example, set $c_2 = 1$ and for c_1, pick any real number such that $0 < c_1 < a_1, b_1$.

- **Claim** *N has refinement.*
  Suppose $(a_1, a_2) + (b_1, b_2) = (c_1, c_2) + (d_1, d_2)$ in N. From the addition rule, we have $\max\{a_1, b_1\} = \max\{c_1, d_1\}$. Without loss of generality we will assume $a_1 = c_1$ is this maximum, that is, $b_1, d_1 \leq a_1 = c_1$. We consider the following cases:
  - *If $b_1 < a_1 = c_1$, then $(a_1, a_2) = (a_1, a_2) + (b_1, b_2) = (c_1, c_2) + (d_1, d_2)$ and $(c_1, c_2) + (b_1, b_2) = (c_1, c_2)$, so we can make the refinement*

$$\begin{array}{cc} & \begin{array}{cc} (c_1, c_2) & (d_1, d_2) \end{array} \\ \begin{array}{c} (a_1, a_2) \\ (b_1, b_2) \end{array} & \begin{pmatrix} (c_1, c_2) & (d_1, d_2) \\ (b_1, b_2) & * \end{pmatrix} \end{array}$$

  - *If $d_1 < c_1$, then we can construct a refinement by symmetry with the previous case.*
  - *If $a_1 = b_1 = c_1 = d_1$, then we have $a_2 + b_2 = c_2 + d_2$ in R^{++}. Since R^{++} has refinement, we can use a refinement of this equation to provide the second components of entries in a refinement of $(a_1, a_2) + (b_1, b_2) = (c_1, c_2) + (d_1, d_2)$.*
- **Claim** *N is weakly cancellative.*
  Suppose $(a_1, a_2) + (c_1, c_2) = (b_1, b_2) + (c_1, c_2)$ in N. We consider the following cases:
  - *If $a_1 < c_1$, then $(c_1, c_2) = (a_1, a_2) + (c_1, c_2) = (b_1, b_2) + (c_1, c_2)$ so we can make the refinement*

$$\begin{array}{cc} & \begin{array}{cc} (b_1, b_2) & (c_1, c_2) \end{array} \\ \begin{array}{c} (a_1, a_2) \\ (c_1, c_2) \end{array} & \begin{pmatrix} * & (a_1, a_2) \\ (b_1, b_2) & (c_1, c_2) \end{pmatrix} \end{array}$$

  - *If $a_1 > c_1$, then, since $\max\{a_1, c_1\} = \max\{b_1, c_1\}$, we must have $b_1 = a_1$. Thus $(a_1, a_2) = (a_1, a_2) + (c_1, c_2) = (b_1, b_2) + (c_1, c_2) = (b_1, b_2)$, and we can* *make the refinement*

$$\begin{array}{c|cc} & (b_1,b_2) & (c_1,c_2) \\ \hline (a_1,a_2) & (a_1,a_2) & * \\ (c_1,c_2) & * & (c_1,c_2) \end{array}$$

- *If $a_1 = c_1$, then since $\max\{a_1, c_1\} = \max\{b_1, c_1\}$, we have $b_1 \leq a_1$. In fact we must have $b_1 = a_1$, since if $b_1 < a_1 = c_1$, then $a_2 + c_2 = c_2$ which is not possible in $\mathbb{R}^{++}$. Thus $b_1 = a_1$ and $a_2 + c_2 = b_2 + c_2$. This implies $a_2 = b_2$ and hence $(a_1, a_2) = (b_1, b_2)$. We can make a refinement as in the previous case.*

In each of these refinement matrices, the lower right entry is (c_1, c_2), so we have shown that N is weakly cancellative.

From the semigroup N, we can now construct a weakly cancellative non-midseparative refinement monoid:

Example 9.9. *A weakly cancellative refinement monoid which is not midseparative:*

Let N be as in the previous example, and define the semigroup $M = \{0, \infty\} \times N$. Since $\{0, \infty\}$ and N are weakly cancellative refinement semigroups, so is M. Since N has no regular elements, neither does M.

Let M^0 be the monoid made by addition of a zero element to M. Then M^0 is a weakly cancellative monoid with no proper regular elements. But M is not strongly separative since, for example, if $a = (\infty, (1,1))$ and $b = (0, (1,1))$, then $2a = a + b = (\infty, (1,2))$, but $a \neq b$. In particular, by 9.6, M^0 is not midseparative.

Note that, by 9.5, given a and b as above, there should be some x such that $a = a + x = b + x$. It is easily checked that $x = (\infty, (1/2, 1))$, for example, has the required property.

As we saw in 8.13 and 8.14, for the equation $a + c = b + c$ in a refinement monoid, it is actually the cancellation properties of $\{\prec c\}$ that matter:

Proposition 9.10. *Let a, b, c be elements of a refinement monoid such that $a + c = b + c$.*

1. *If $\{\prec c\}$ is weakly cancellative, then there exists a refinement matrix*

$$\begin{array}{c} \\ a \\ c \end{array} \begin{array}{c} \begin{array}{cc} b & c \end{array} \\ \begin{pmatrix} * & * \\ * & c' \end{pmatrix} \end{array}$$

   with $c \leq c'$.

2. *If $\{\prec c\}$ is midseparative and $c \prec a$, then there exists an idempotent $e \leq c$ such that $a = b + e$.*

Proof.

1. From the equation $a + c = b + c$ we make a refinement

$$\begin{array}{c} \\ a \\ c \end{array} \begin{array}{c} \begin{array}{cc} b & c \end{array} \\ \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

This gives us the equation $a_1 + c_1 = b_1 + c_1$ in the weakly cancellative monoid $\{\prec c\}$. So there is a refinement

$$\begin{array}{cc} & \begin{array}{cc} b_1 & c_1 \end{array} \\ \begin{array}{c} a_1 \\ c_1 \end{array} & \begin{pmatrix} d_2 & a_2 \\ b_2 & c_2 \end{pmatrix} \end{array}$$

with $c_1 \leq c_2$. It is then easy to check that

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 + d_2 & a_2 \\ b_2 & a_1 + c_2 \end{pmatrix} \end{array}$$

is a refinement of the original equation. Further, setting $c' = a_1 + c_2$ we have $c = a_1 + c_1 \leq a_1 + c_2 = c'$.

2. Since $c \prec a$, there is some $n \in \mathbb{N}$ such that $c \leq na$. We do an induction on n...

   $n = 1$ Since $c \leq a$ we can use 8.5.1 to get a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

   with $c_1 \leq a_1$. Since $a_1, b_1, c_1 \in \{\prec c\}$, the equation $a_1 + c_1 = b_1 + c_1$ implies there is some idempotent $e \leq c_1$ such that $a_1 = b_1 + e$. Thus $a = d_1 + a_1 = d_1 + b_1 + e = b + e$.

   $n > 1$ Suppose the claim is true for some $n \in \mathbb{N}$, and $a + c = b + c$ with $c \leq (n+1)a$. We can write $c = c_1 + c_2$ with $c_1 \leq a$ and $c_2 \leq na$. Since $c_1, c_2 \leq c$, the monoids $\{\prec c_1\}$ and $\{\prec c_1\}$ are midseparative. We have $(a + c_1) + c_2 = (b + c_1) + c_2$ with $c_2 \leq n(a + c_1)$, so by induction, there is $2e_2 = e_2 \leq c_2$ such that $a + c_1 = (b + e_2) + c_1$. Since $c_1 \leq a$, the $n = 1$ case provides $2e_1 = e_1 \leq c_1$ such that $a = b + (e_1 + e_2)$. We have also $2(e_1 + e_2) = e_1 + e_2 \leq c_1 + c_2 = c$, so the induction is complete.

□

We next consider the extension question for weak cancellation and midseparativity: If I is an order ideal of M such that I and M/I are weakly cancellative (midseparative), is M weakly cancellative (midseparative)? Unlike separative and strongly separative monoids (8.16), we will see that weakly cancellative and midseparative monoids do not have this extension property. The refinement monoid of example 15.9 is not weakly cancellative, even though there is an order ideal $I \leq M$ such that I and M/I are both weakly cancellative. Example 9.7 is a counterexample in the midseparative case.

Later we will see that we get extension for midseparativity if I is not just midseparative, but either Artinian or semi-Artinian. See 13.6 and 15.5.

We consider next the converse question: If $I \leq M$, and M is weakly cancellative (midseparative), are I and M/I also weakly cancellative (midseparative)? Here weak cancellation and midseparativity behave like cancellation, separativity and strong separativity as seen in 8.15. Curiously, the proofs below that M/I is weakly cancellative (midseparative) do not require that I be an order ideal– it suffices that I be a submonoid. In contrast, to be certain that I has the same separativity properties as M, I must be an order ideal.

Proposition 9.11. *Let I be a order ideal in monoid M. If M is weakly cancellative (midseparative), then so are I and M/I.*

Proof. The claim that I has the same cancellation properties as M is clear, so we prove only the claims about M/I:

- Suppose M is weakly cancellative and $[a]_I+[c]_I=[b]_I+[c]_I$. Then there are $u,v\in I$ such that $a+c+u=b+c+v$. Since M is weakly cancellative, there is a refinement matrix
$$\begin{array}{c|cc} & b+v & c \\ \hline a+u & d_1 & a_1 \\ c & b_1 & c_1 \end{array}$$
with $c\le c_1$. In M/I, this refinement maps to
$$\begin{array}{c|cc} & [b]_I & [c]_I \\ \hline [a]_I & [d_1]_I & [a_1]_I \\ [c]_I & [b_1]_I & [c_1]_I \end{array}$$
with $[c]_I\le[c_1]_I$.
- Suppose M is midseparative and $2[a]_I=[a]_I+[b]_I$. Then there are $u,v\in I$ such that $2a+u=a+b+v$. Thus $2(a+u)=(a+u)+(b+v)$, and since M is midseparative, there is some idempotent $e\in M$ such that $a+u=b+v+e$. In M/I we then have $[a]_I=[b]_I+[e]_I$ and $2[e]_I=[e]_I$.

□

For the remainder of this section we will prove some important properties of weakly cancellative refinement monoids:

Lemma 9.12. *Let M be a weakly cancellative refinement monoid and $a,b,c_1,c_2\in M$ such that $a+c_1+c_2=b+c_1+c_2$. Then there is a refinement matrix*
$$\begin{array}{c|ccc} & b & c_1 & c_2 \\ \hline a & * & * & * \\ c_1 & * & c_1' & * \\ c_2 & * & * & c_2' \end{array}$$
such that $c_1\le c_1'$ and $c_2\le c_2'$.

Proof. Applying weak cancellation to the equation $(a+c_2)+c_1=(b+c_2)+c_1$ we get a refinement of the form
$$\begin{array}{c|cc} & b+c_2 & c_1 \\ \hline a+c_2 & d_1 & a_1 \\ c_1 & b_1 & c_3 \end{array}$$

with $c_1 \leq c_3$. Since $(a+b_1)+c_2 = d_1+a_1+b_1 = (b+a_1)+c_2$, there is a refinement

$$\begin{array}{c|cc} & b+a_1 & c_2 \\ a+b_1 & * & * \\ c_2 & * & c_2' \end{array}$$

with $c_2 \leq c_2'$. Refining further we get

$$\begin{array}{c} \\ a \\ b_1 \\ c_2 \end{array} \begin{array}{c} \begin{array}{ccc} b & a_1 & c_2 \end{array} \\ \begin{pmatrix} d_2 & a_2 & a_3 \\ b_2 & c_5 & b_4 \\ b_3 & a_4 & c_2' \end{pmatrix} \end{array}$$

It is then easy to check that the refinement matrix

$$\begin{array}{c} \\ a \\ c_1 \\ c_2 \end{array} \begin{array}{c} \begin{array}{ccc} b & c_1 & c_2 \end{array} \\ \begin{pmatrix} d_2 & a_2 & a_3 \\ b_2 & c_5+c_3 & b_4 \\ b_3 & a_4 & c_2' \end{pmatrix} \end{array}$$

has the required properties. □

Lemma 9.13. *Let a, b, c, c_1, c_2 be elements of a weakly cancellative refinement monoid M. Then*

1. $a+c \leq b+c \implies (\exists a_1 \ll c$ such that $a \leq b+a_1)$
2. $a \ll c_1+c_2 \implies (\exists a_1, a_2$ such that $a = a_1+a_2,\ a_1 \ll c_1$ and $a_2 \ll c_2)$
3. $\{\ll c_1+c_2\} = \{\ll c_1\} + \{\ll c_2\}$
4. $a \equiv c_1+c_2 \implies (\exists a_1, a_2$ such that $a = a_1+a_2,\ a_1 \equiv c_1$ and $a_2 \equiv c_2)$

Proof.

1. There is some $x \in M$ such that $a+x+c = b+c$, and hence a refinement matrix

$$\begin{array}{c} \\ a+x \\ c \end{array} \begin{array}{c} \begin{array}{cc} b & c \end{array} \\ \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

   with $c \leq c_1$. This implies $a \leq a+x = d_1+a_1 \leq b+a_1$ with $a_1 \ll c$.
2. We have $a+c_1+c_2 \leq c_1+c_2$. Using 1, there is some $a_1' \ll c_1$ such that $a+c_2 \leq a_1'+c_2$. Using 1 again, there is some $a_2' \ll c_2$ such that $a \leq a_1'+a_2'$. Decomposing this last inequality, there are $a_1 \leq a_1' \ll c_1$ and $a_2 \leq a_2' \ll c_2$ such that $a = a_1+a_2$.
3. This follows directly from 2 and 6.6.7.
4. We have $c_1+c_2 \leq a$ so there is some u such that $c_1+c_2+u = a \leq c_1+c_2$. Since $u \ll c_1+c_2$, by 1, there are u_1, u_2 with $u = u_1+u_2$, $u_1 \ll c_1$ and $u_2 \ll c_2$. Set $a_1 = c_1+u_1 \equiv c_1$ and $a_2 = c_2+u_2 \equiv c_2$, then $a = a_1+a_2$.

□

Proposition 9.14. *If M is a weakly cancellative refinement monoid then so is $\overline{M}$.*

Proof. We show first that $\overline{M}$ has refinement...

Suppose $[x_1] + [x_2] = [y_1] + [y_2]$ in $\overline{M}$, then $x_1 + x_2 \equiv y_1 + y_2$. From 9.13.4, there are x_1', x_2' such that, $x_1 \equiv x_1'$, $x_2 \equiv x_2'$ and $x_1' + x_2' = y_1 + y_2$. We make a refinement of this equation:

$$\begin{array}{cc} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1' \\ x_2' \end{array} & \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \end{array}$$

Since $[x_1] = [x_1']$ and $[x_2] = [x_2']$,

$$\begin{array}{cc} & \begin{array}{cc} [y_1] & [y_2] \end{array} \\ \begin{array}{c} [x_1] \\ [x_2] \end{array} & \begin{pmatrix} [z_{11}] & [z_{12}] \\ [z_{21}] & [z_{22}] \end{pmatrix} \end{array}$$

is a refinement of the original equation.

We show that $\overline{M}$ is weakly cancellative...

Suppose $[a] + [c] = [b] + [c]$ in $\overline{M}$, then $a + c \equiv b + c$. From 9.13.4, there are a', c' such that, $a \equiv a'$, $c \equiv c'$ and $a' + c' = b + c$. Let $u \ll c$ such that $c' = c + u$, then $a' + c + u = b + c$ and there is a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a' + u \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c \equiv c_1$. If we make a refinement of the equation $a' + u = d_1 + a_1$,

$$\begin{array}{cc} & \begin{array}{cc} d_1 & a_1 \end{array} \\ \begin{array}{c} a' \\ u \end{array} & \begin{pmatrix} d_1' & a_1' \\ u_1 & u_2 \end{pmatrix} \end{array}$$

we can rewrite the first refinement as

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a' \\ c' \end{array} & \begin{pmatrix} d_1' & a_1' \\ b_1 + u_1 & c_1 + u_2 \end{pmatrix} \end{array}$$

Since $u_2 \leq u \ll c$, we have $c \equiv c' \equiv c_1 + u_2$, and in $\overline{M}$,

$$\begin{array}{cc} & \begin{array}{cc} [b] & [c] \end{array} \\ \begin{array}{c} [a] \\ [c] \end{array} & \begin{pmatrix} [d_1'] & [a_1'] \\ [b_1 + u_1] & [c] \end{pmatrix} \end{array}$$

Thus $\overline{M}$ is weakly cancellative. □

Monoids with weak cancellation have a property called **Riesz interpolation** [10]: If there are elements $a_0, a_1, b_0, b_1 \in M$ with $a_0 \leq b_0$, $a_0 \leq b_1$, $a_1 \leq b_0$ and $a_1 \leq b_1$ (which we write as $a_0, a_1 \leq b_0, b_1$), then there is an element c that fits between, that is, $a_0, a_1 \leq c \leq b_0, b_1$. The proof which follows is a variation of the proof found in Tarski [30] which was applied to partially ordered monoids only.

Proposition 9.15. *Let M be a weakly cancellative refinement monoid. Then*

1. $(\forall a, b, c \in M)$ $(c \leq b \leq c + a \implies \exists d \leq a$ such that $b \equiv c + d)$
2. $(\forall a_0, a_1, b_0, b_1 \in M)$ $(a_0, a_1 \leq b_0, b_1 \implies \exists c$ such that $a_0, a_1 \leq c \leq b_0, b_1)$

Proof.

1. Let $x, y \in M$ be such that $b = c + x$ and $c + a = b + y$. Then $c + a = c + x + y$. Since M is weakly cancellative, there is a refinement matrix
$$\begin{array}{cc} & \begin{array}{cc} x+y & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$
with $b_1 \ll c$. Since $x \leq d_1 + b_1$ there are $d, x' \in M$ such that $x = d + x'$, $d \leq d_1 \leq a$ and $x' \leq b_1 \ll c$. In particular, $c \equiv c + x'$, and adding d to this equation we get $c + d \equiv c + d + x' = c + x = b$.
2. Let $d_0, d_1, f_0, f_1 \in M$ be such that $a_0 + d_0 = a_1 + d_1 = b_0$ and $a_0 + f_0 = a_1 + f_1 = b_1$. From the first of these equations we get a refinement matrix
$$\begin{array}{cc} & \begin{array}{cc} a_0 & d_0 \end{array} \\ \begin{array}{c} a_1 \\ d_1 \end{array} & \begin{pmatrix} r_1 & r_3 \\ r_2 & r_4 \end{pmatrix} \end{array}$$
Thus $r_1 \leq a_1 \leq b_1 = a_0 + f_0 = r_1 + (r_2 + f_0)$. Applying 1 to the inequality $r_1 \leq a_1 \leq r_1 + (r_2 + f_0)$, we get some s such that $a_1 \equiv r_1 + s$ and $s \leq r_2 + f_0$. Decompose this last inequality to get $s = s_0 + s_1$ with $s_0 \leq f_0$ and $s_1 \leq r_2$. Finally, we set $c = a_0 + s_0$ and check that c has the required properties:
$$c \geq a_0$$
$$c = a_0 + s_0 = r_1 + r_2 + s_0 \geq r_1 + s_0 + s_1 = r_1 + s \equiv a_1$$
$$c = a_0 + s_0 \leq a_0 + f_0 = b_1$$
$$c = a_0 + s_0 = r_1 + r_2 + s_0 \leq r_1 + r_2 + s \equiv a_1 + r_2 \leq a_1 + d_1 = b_0$$

□

10 Groups and Monoids

In Section 6 we investigated the order structure of monoids and found, among other things, that a monoid M such that $a \leq b$ for all elements, is an Abelian group. In this circumstance, $\overline{M}$ has one element. Thus, the information about M lost in going from M to $\overline{M}$ is contained in an Abelian group. This is a quite general phenomenon: For an arbitrary monoid M every element of $\overline{M}$ corresponds in a natural way to an Abelian group which represents algebraic information lost in mapping the elements of a $\equiv$-congruence class to their image in $\overline{M}$.

These groups are all trivial if and only if M is partially ordered, that is, $M = \overline{M}$. At the other extreme, $\overline{M}$ is trivial (contains one element) if and only if M is itself a group. So these groups and $\overline{M}$ contain complementary information about the monoid M. Nonetheless, knowing these groups and $\overline{M}$ is not, in general, sufficient to reconstruct M.

In our application to module categories, we will see in 19 that the ideal class group of a Dedekind domain is a group of the type to be discussed in this section.

For each element r of M, we will construct an Abelian group $G_r(M)$ whose elements are in bijection with the elements of the congruence class $\{\equiv r\} = \{x \in M \mid x \leq r \leq x\} \subseteq M$. Thus $\{\equiv r\}$ can itself be considered to be the group $G_r(M)$ if the addition is suitably redefined. In the discussion of this fact we will sometimes think of $\{\equiv r\}$ as a <u>subclass</u> of M, and sometimes as an <u>element</u> of $\overline{M}$. To distinguish these two situations, we will write $\{\equiv r\} \subseteq M$ or $\{\equiv r\} \in \overline{M}$ as appropriate.

We start by defining $G_0(M)$...

Definition 10.1. *For a monoid M, define $G_0(M) = \{\leq 0\} \subseteq M$.*

An element $a \in M$ is in $G_0(M)$ if and only if there is some $b \in M$ such that $a + b = 0$, that is, if and only if a has an inverse. Thus $G_0(M)$ is an Abelian group with the same operation and identity as in M. In fact, it is also the largest such group contained in M.

Now we can define $G_r(M)$ for any $r \in M$...

Definition 10.2. *Let r be an element of a monoid M.*

1. *Define a relation $\sim_r$ on M by*

$$a \sim_r b \iff r + a = r + b$$

   for $a, b \in M$. An easy calculation shows that $\sim_r$ is a congruence on M. We will write $[a]_r$ for the $\sim_r$-congruence class containing $a \in M$.

2. *Define*

$$G_r(M) = G_0(M/\sim_r).$$

   This definition is consistent with definition 10.1 since $\sim_0$ is the trivial congruence. We will often write G_0 and G_r instead of $G_0(M)$ and $G_r(M)$ if the monoid is clear from context.

3. *We will also find it useful to define the monoid*

$$H_r = \{\prec r\}/\sim_r .$$

We collect in the next proposition some simple facts about H_r and G_r:

Proposition 10.3. *Let M be a monoid and $r \in M$.*

1. *For all $a \in M$, $[a]_r \in G_r$ if and only if $a \ll r$. Thus, $G_r = \{\ll r\}/\sim_r$.*
2. $G_r = G_0(H_r)$
3. $G_0 = H_0$
4. *If $r \leq s$ in M, then the map $\phi_{sr}\colon M/\sim_r \to M/\sim_s$ given by $[a]_r \mapsto [a]_s$, is a monoid homomorphism. Further, when restricted to H_r, ϕ_{sr} is a monoid homomorphism into H_s, and when restricted to G_r, ϕ_{sr} is a group homomorphism into G_s.*
5. *If $r \leq s \leq t$ in M, then $\phi_{ts} \circ \phi_{sr} = \phi_{sr}$.*
6. *If $s \in \{\equiv r\} \subseteq M$, then $G_r = G_s$ and $H_r = H_s$.*

Proof. Easy. □

From this proposition we note that G_r and H_r depend only on the congruence class $\{\equiv r\} \in \overline{M}$, and so we could have more correctly used the notation $G_{\{\equiv r\}}$ and $H_{\{\equiv r\}}$.

We next want to show that there is a bijection from G_r to $\{\equiv r\} \subseteq M$, so that we can think of $\{\equiv r\}$ itself as a group.

Proposition 10.4. *Let M be a monoid and $r \in M$. Then the map $\rho : G_r \to \{\equiv r\}$ given by $\rho([a]_r) = r + a$ for $a \ll r$, is a bijection.*

Proof. For two elements $a_1, a_2 \ll r$ we have

$$[a_1]_r = [a_2]_r \iff r + a_1 = r + a_2,$$

so the map ρ is well defined and injective.

To show surjectivity, suppose $r_1 \in \{\equiv r\}$. Then there is some $a \ll r$ such that $r_1 = r + a$. Thus we get $[a]_r \in G_r$ and $r_1 = \rho([a]_r)$. □

Using the map ρ, we can construct a new operation $+_r$ on $\{\equiv r\} \subseteq M$ which makes it into a group: If $r_1 = r + a_1$ and $r_2 = r + a_2$ are in $\{\equiv r\}$ with $a_1, a_2 \ll r$, then we can define

$$r_1 +_r r_2 = \rho(\rho^{-1}(r_1) + \rho^{-1}(r_2)) = \rho([a_1]_r + [a_2]_r) = \rho([a_1 + a_2]_r) = r + a_1 + a_2.$$

Clearly, r is the identity element of the group $(\{\equiv r\}, +_r)$.

Notice that if $r_1 +_r r_2 = r_3$, then $r + r_3 = r_1 + r_2$. In a separative monoid, this suffices to define the operation $+_r$: If we had two elements $r_3, r_3' \in \{\equiv r\}$ such that $r + r_3 = r + r_3' = r_1 + r_2$, then, since $r \leq r_3, r_3'$, we can cancel r to get $r_3 = r_3'$.

Notice also that the map ρ depends on the element of $\{\equiv r\}$ used in its definition, so the group structure and operation $+_r$ on $\{\equiv r\}$ are not unique. Nonetheless, however ρ is constructed, the resulting group is always isomorphic to G_r.

We note the special case when r is regular…

Recall from 6.9 that if r is regular there is a unique idempotent e in $\{\equiv r\}$. It is simple task to show that with the existing monoid operation, $\{\equiv r\} \subseteq M$ is a group with identity e. We will confirm this fact by showing that in this situation the monoid operations $+$ and $+_e$ coincide on $\{\equiv r\} = \{\equiv e\} \subseteq M$.

Suppose $r_1, r_2 \in \{\equiv r\} = \{\equiv e\}$. Since $e \leq r_1, r_2$ and $2e = e$, we get $r_1 = e + r_1$ and $r_2 = e + r_2$. From the above definition of $+_e$ we get $r_1 +_e r_2 = e + r_1 + r_2 = r_1 + r_2$.

We have therefore shown the following:

Proposition 10.5. *If r is a regular element of a monoid, then*

$$G_r \cong (\{\equiv r\}, +)$$

as groups.

In general, we have from 10.3.6 that $G_r = G_s$ if $r \equiv s$. If M is cancellative, then it is easy to see that $G_r = G_s$ for any $r, s \in M$. If M is separative then we get an intermediate result, namely, all elements of an Archimedean component of M are associated to the same group:

Proposition 10.6. *Let r, s be elements of a separative monoid M such that $s \prec r \prec s$. Then the relations $\sim_r$ and $\sim_s$ coincide and, in particular, $G_r = G_s$, $H_r = H_s$ and $\{\ll r\} = \{\ll s\}$.*

Proof. Let $n \in \mathbb{N}$ such that $r \leq ns$. If $a \sim_r b$ for elements $a, b \in M$, then $a + r = b + r$, and so $a + ns = b + ns$. Separativity then implies that $a + s = b + s$, that is, $a \sim_s b$.

By symmetry, $a \sim_s b$ implies $a \sim_r b$ for all $a, b \in M$. From 10.2 we get $G_r = G_s$, $H_r = H_s$, and, by 10.3.1, $\{\ll r\} = \{\ll s\}$. $\square$

Thus for separative monoids, the groups G_r could be indexed by the Archimedean components of M, that is, by the elements of $\widetilde{M}$.

In separative monoids the significance of H_r becomes clearer:

Proposition 10.7. *Let r be an element of a separative monoid M. Then*

1. *The Archimedean component $\{\asymp r\}$ is embedded in H_r via the monoid homomorphism $a \mapsto [a]_r$.*
2. *H_r is cancellative.*

Proof.

1. We show that the homomorphism is injective when restricted to $\{\asymp r\}$...

   Suppose $a, b \in \{\asymp r\}$ such that $[a]_r = [b]_r$. Then $a + r = b + r$, so we can use the fact that Archimedean components are cancellative (8.10.5) to get $a = b$.
2. Suppose $a, b, c \in \{\prec r\}$ such that $[a]_r + [c]_r = [b]_r + [c]_r$. Then $a + c + r = b + c + r$. Since $c \prec r$, we can use 8.10.4 to cancel c from this equation to get $a + r = b + r$, that is, $[a]_r = [b]_r$.

$\square$

Every monoid has associated with it an Abelian group which is universal for homomorphisms from the monoid into groups.

Definition 10.8. *Let M be a monoid. Then the* **Grothendieck group** *of M,* [4, Section 1.3], [33, Appendix G]*, written $G(M)$, is constructed as a factor monoid of $M \times M$ as follows:*

1. *Define a relation $\approx$ on $M \times M$ by*
$$(a_1, a_2) \approx (b_1, b_2) \iff \exists x \in M \text{ such that } a_1 + b_2 + x = a_2 + b_1 + x,$$
for $(a_1, a_2), (b_1, b_2) \in M \times M$. Notice that $(a, a) \approx (0, 0)$ for any $a \in M$. It is not hard to check that $\approx$ is a congruence on $M \times M$.
2. *Define*
$$G(M) = (M \times M)/\approx .$$
Since $(a_1, a_2)+(a_2, a_1) \approx 0$ for all $a_1, a_2 \in M$, every element of $G(M)$ has an inverse, and so $G(M)$ is an Abelian group.
3. *We will write $\langle a\rangle_M$ for the image of $(a, 0) \in M \times M$ in $G(M)$. The image of $(0, a)$ is $-\langle a\rangle_M$, so every element of $G(M)$ can be written in the form $\langle a_1\rangle_M - \langle a_2\rangle_M$ for some $a_1, a_2 \in M$.*

Note also that for $a, b \in M$,
$$\langle a\rangle_M = \langle b\rangle_M \iff \exists x \in M \text{ such that } a + x = b + x.$$

The map $\langle\ \rangle_M \colon M \to G(M)$ is easily seen to be a monoid homomorphism. It is injective if and only if M is cancellative, and an isomorphism if and only if M is itself a group.

The main importance of $G(M)$ is its universal property:

Proposition 10.9. *Let M be a monoid, H an Abelian group, and $\psi\colon M \to H$, a monoid homomorphism. Then there is a unique group homomorphism $\hat{\psi} : G(M) \to H$ such that the following diagram commutes:*

$$\begin{tikzcd} M \arrow[r, "\langle\ \rangle_M"] \arrow[dr, "\psi"'] & G(M) \arrow[d, "\hat{\psi}"] \\ & H \end{tikzcd}$$

Proof. We define the map $\hat{\psi}$ by $\hat{\psi}(\langle a_1\rangle_M - \langle a_2\rangle_M) = \psi(a_1) - \psi(a_2)$ for $a_1, a_2 \in M$. It is easy to check that this map is well defined and makes the diagram commute. $\square$

Suppose A is a submonoid of M. Then the inclusion map of A in M followed by $\langle\ \rangle_M \colon M \to G(M)$ is a homomorphism of A into a group. So there is an induced group homomorphism from $G(A)$ to $G(M)$ such that $\langle a\rangle_A \mapsto \langle a\rangle_M$ for all $a \in A$. In general, this homomorphism is neither injective nor surjective.

Proposition 10.10. *Let A and B be order ideals of a refinement monoid. Then there is an exact sequence of Abelian groups*
$$G(A \cap B) \xrightarrow{\alpha} G(A) \times G(B) \xrightarrow{\beta} G(A + B) \longrightarrow 0$$
where α and β are group homomorphisms such that
$$\alpha(\langle x\rangle_{A\cap B}) = (\langle x\rangle_A, -\langle x\rangle_B)$$
$$\beta(\langle a\rangle_A, \langle b\rangle_B) = \langle a\rangle_{A+B} + \langle b\rangle_{A+B}$$
for all $x \in A \cap B$, $a \in A$ and $b \in B$.

Proof. Since $A \cap B \subseteq A, B \subseteq A + B$, the maps α and β are well defined group homomorphisms. That β is surjective and $\beta \circ \alpha = 0$ is easy to show, so we need to check only that $\ker \beta \subseteq \operatorname{im} \alpha \ldots$

Suppose $\beta(\langle a_1 \rangle_A - \langle a_2 \rangle_A, \langle b_1 \rangle_B - \langle b_2 \rangle_B) = 0$ for some $a_1, a_2 \in A$ and $b_1, b_2 \in B$, that is, $\langle a_1 + b_1 \rangle_{A+B} = \langle a_2 + b_2 \rangle_{A+B}$. Then there is some $a' + b' \in A + B$ with $a' \in A$ and $b' \in B$, such that $(a' + b') + (a_1 + b_1) = (a' + b') + (a_2 + b_2)$. We make a refinement of this equation:

$$\begin{array}{cc} & \begin{array}{cccc} a' & b' & a_1 & b_1 \end{array} \\ \begin{array}{c} a' \\ b' \\ a_2 \\ b_2 \end{array} & \begin{pmatrix} a_3 & x_3 & a_4 & x_4 \\ y_3 & b_3 & y_4 & b_4 \\ a_5 & x_5 & a_6 & x_6 \\ y_5 & b_5 & y_6 & b_6 \end{pmatrix} \end{array}$$

Since A and B are order ideals, we have $a_i \in A$, $b_i \in B$ and $x_i, y_i \in A \cap B$ for $i = 3, 4, 5, 6$.

From the equation $a' = a_3 + y_3 + a_5 + y_5 = a_3 + x_3 + a_4 + x_4$ in A, we get $\langle a_4 \rangle_A - \langle a_5 \rangle_A = \langle y_3 + y_5 \rangle_A - \langle x_3 + x_4 \rangle_A$. Thus

$$\begin{aligned} \langle a_1 \rangle_A - \langle a_2 \rangle_A &= \langle a_4 \rangle_A + \langle y_4 + y_6 \rangle_A - \langle a_5 \rangle_A - \langle x_5 + x_6 \rangle_A \\ &= \langle y_3 + y_4 + y_5 + y_6 \rangle_A - \langle x_3 + x_4 + x_5 + x_6 \rangle_A \\ &= \langle y \rangle_A - \langle x \rangle_A \end{aligned}$$

where $x = x_3 + x_4 + x_5 + x_6 \in A \cap B$ and $y = y_3 + y_4 + y_5 + y_6 \in A \cap B$.

Similarly, $\langle b_1 \rangle_B - \langle b_2 \rangle_B = \langle x \rangle_B - \langle y \rangle_B$. Hence

$$(\langle a_1 \rangle_A - \langle a_2 \rangle_A, \langle b_1 \rangle_B - \langle b_2 \rangle_B) = \alpha(\langle y \rangle_{A \cap B} - \langle x \rangle_{A \cap B}),$$

that is, $(\langle a_1 \rangle_A - \langle a_2 \rangle_A, \langle b_1 \rangle_B - \langle b_2 \rangle_B) \in \operatorname{im} \alpha$. □

We now consider the relationship between $G(M)$ and $G_r(M) \ldots$

From the above discussion we have

$$a \sim_r b \implies a + r = b + r \implies \langle a \rangle_M = \langle b \rangle_M$$

for any $a, b \in M$. Thus there is always a monoid homomorphism ψ_r from $M/\!\sim_r$ to $G(M)$ given by $[a]_r \mapsto \langle a \rangle_M$. This monoid homomorphism when restricted to G_r is a group homomorphism. One circumstance in which this homomorphism is injective is given by the following:

Proposition 10.11. *Let M be a separative monoid and $r \in M$ such that $M = \{\prec r\}$. Then for all $a, b \in M$*

$$a \sim_r b \iff \langle a \rangle_M = \langle b \rangle_M.$$

Thus the monoid homomorphism $\psi_r \colon H_r \to G(M)$ given by $[a]_r \mapsto \langle a \rangle_M$ is injective. In particular, G_r and H_r embed in $G(M)$. Further $G(H_r) \cong G(M)$.

Proof. We need to check that $\langle a \rangle_M = \langle b \rangle_M \implies a \sim_r b$ for all $a, b \in M \ldots$

If $\langle a \rangle_M = \langle b \rangle_M$, then there is some $x \in M$ such that $a + x = b + x$. Since $x \in \{\prec r\}$ there is some number n such that $x \leq nr$. Thus $a + nr = b + nr$. Separativity of M allows us to cancel $(n-1)r$ from this equation to get $a + r = b + r$, that is, $a \sim_r b$.

Thus we have shown that ψ_r is injective, and that G_r and H_r embed in $G(M)$.

We prove the last claim: Since ψ_r is a monoid homomorphism from H_r to a group $G(M)$, there is an induced group homomorphism $\hat{\psi}_r \colon G(H_r) \to G(M)$ given by $\hat{\psi}_r(\langle [a]_r \rangle_{H_r}) = \langle a \rangle_M$ for all $a \in M$. This homomorphism is easily seen to be surjective. To show injectivity we calculate $\ker \hat{\psi}_r$: Suppose $\hat{\psi}_r(\langle [a]_r \rangle_{H_r} - \langle [b]_r \rangle_{H_r}) = 0$. Then $\langle a \rangle_M = \langle b \rangle_M$, and since ψ_r is injective, $[a]_r = [b]_r$. Thus $\langle [a]_r \rangle_{H_r} - \langle [b]_r \rangle_{H_r} = 0$. $\square$

With this proposition and 10.7, we see that the Archimedean component $\{\asymp r\}$ is embedded in $G(M)$.

We again consider the special (and simpler) case when r is regular. Here we get similar results without needing separativity:

Proposition 10.12. *Let M be a monoid, $r \in M$ a regular element such that $\{\prec r\} = M$. Then $H_r = G_r \cong G(M)$, via the homomorphism ψ_r.*

Proof. Since r is regular, we have $\{\ll r\} = \{\leq r\} = \{\prec r\}$, so

$$H_r = \{\prec r\}/\sim_r = \{\ll r\}/\sim_r = G_r.$$

To show that ψ_r is injective, consider $a, b \in M$ such $\psi_r([a]_r) = \psi_r([a]_r)$. Then $\langle a \rangle_M = \langle b \rangle_M$, so there is some $x \in M$ with $a + x = b + x$. But $x \leq r$, so $a + r = b + r$, that is, $[a]_r = [b]_r$.

Any element of $G(M)$ is in the form $\langle a \rangle_M - \langle b \rangle_M$ for suitable $a, b \in M$. We have $a, b \ll r$, so $[a]_r$ and $[b]_r$ are in G_r, and $[b]_r$ has an inverse $-[b]_r$. Thus $\psi_r([a]_r - [b]_r) = \langle a \rangle_M - \langle b \rangle_M$, and ψ_r is surjective. $\square$

As well as the injection $\psi_r \colon G_r \to G(M)$, there is always a surjective group homomorphism $\nu_r \colon G(\{\ll r\}) \to G_r$ defined by $\langle a \rangle_{\{\ll r\}} \mapsto [a]_r$ for all $a \ll r$. These properties follow from the universal property of $G(\{\ll r\})$ and 10.3.1. Our final goal in this section is to use this group homomorphism to show a relationship between G_a, G_b and G_{a+b} for elements a and b of a weakly cancellative monoid. In this situation there are epimorphisms $\nu_a \colon G(\{\ll a\}) \to G_a$, $\nu_b \colon G(\{\ll b\}) \to G_b$, and $\nu_{a+b} \colon G(\{\ll a+b\}) \to G_{a+b}$. From 10.10, there is an exact sequence

$$G(\{\ll a\} \cap \{\ll b\}) \to G(\{\ll a\}) \times G(\{\ll b\}) \to G(\{\ll a\} + \{\ll b\}) \to 0.$$

If it happened that $\{\ll a\} + \{\ll b\} = \{\ll a + b\}$, then one might expect that, using the maps ν_a, ν_b and ν_{a+b}, it would be possible to construct a similar exact sequence relating the groups G_a, G_b and G_{a+b}. This is indeed the case for weakly cancellative monoids:

Proposition 10.13. *Let a, b be elements of a weakly cancellative refinement monoid M. Set $A = \{\ll a\}$ and $B = \{\ll b\}$. Then there is an exact sequence of Abelian groups*

$$G(A \cap B) \xrightarrow{\delta} G_a \times G_b \xrightarrow{\gamma} G_{a+b} \longrightarrow 0$$

where γ and δ are given by

$$\gamma([u]_a, [v]_b) = [u + v]_{a+b}$$

$$\delta(\langle x \rangle_{A \cap B}) = ([x]_a, -[x]_b)$$

for all $u \ll a$, $v \ll b$, and $x \in A \cap B$.

Proof. Since the monoid has weak cancellation, we have from 9.13 that

$$\{\ll a\} + \{\ll b\} = \{\ll a+b\},$$

that is, $A + B = \{\ll a+b\}$. Consider the diagram

$$\begin{tikzcd} G(A\cap B) \arrow[r, "\alpha"] \arrow[dr, "\delta"] & G(A)\times G(B) \arrow[r, "\beta"] \arrow[d, "\nu_a\times\nu_b"] & G(A+B) \arrow[r] \arrow[d, "\nu_{a+b}"] & 0 \\ & G_a\times G_b \arrow[r, "\gamma"] & G_{a+b} \arrow[r] & 0 \end{tikzcd}$$

The homomorphisms α and β are defined in 10.10 and make the top row exact. One easily confirms that the diagram commutes, so we get immediately that γ is surjective and $\gamma(\delta(G(A\cap B))) = 0$. Thus it remains to prove that $\ker\gamma \subseteq \operatorname{im}\delta$...

Suppose that $\gamma([u]_a, [v]_b) = [u+v]_{a+b} = 0$ for some $u \ll a$ and $v \ll b$. Then $a+b+u+v = a+b$ and so, using 9.12, there is a refinement matrix

$$\begin{array}{c c} & \begin{array}{cccc} a & b & u & v \end{array} \\ \begin{array}{c} a \\ b \end{array} & \begin{pmatrix} a' & x & u' & z \\ w & b' & y & v' \end{pmatrix} \end{array}$$

with $a \le a'$ and $b \le b'$, say $a' = a+c$ and $b' = b+d$. From the equation $a = (a+c)+x+u'+z$ we get $-[x+z]_a = [u'+c]_a$. Similarly, $-[w+y]_b = [v'+d]_b$. From the equation $a = a'+w = a+c+w$ we get $[c+w]_a = 0$. Similarly $[d+x]_b = 0$. Note also that $w, x, y, z \in A\cap B$, so that $w+y, x+z \in A\cap B$. Finally we calculate

$$\begin{aligned} \delta(\langle w+y\rangle_{A\cap B} - \langle x+z\rangle_{A\cap B}) &= ([w+y]_a - [x+z]_a, -[w+y]_b + [x+z]_b) \\ &= ([w+y+c+u']_a, [v'+d+x+z]_b) \\ &= ([c+w]_a + [u'+y]_a, [d+x]_b + [v'+z]_b) \\ &= ([u]_a, [v]_b). \end{aligned}$$

Thus $\ker\gamma \subseteq \operatorname{im}\delta$. $\square$

11 Primely Generated Refinement Monoids

As for many other algebraic structures, rings, modules, lattices, etc., much can be learned by focusing on those elements which are essential to any generating set for the structure. If one studies numbers then prime numbers are important because any number is a product of primes and because primes can not themselves be expressed as a product except in a trivial way. If one studies finitely generated modules over Artinian rings, then indecomposable modules are important because any such module is a direct sum of indecomposables.

Within a monoid, there are two properties that an element can have which are analogs of the properties of prime numbers and indecomposable modules:

Definition 11.1. *Let M be a monoid and $p \in M$. Then*

- *p is* **prime** *if for all $a_1, a_2 \in M$*

$$p \leq a_1 + a_2 \implies p \leq a_1 \text{ or } p \leq a_2,$$

- *p is* **indecomposable** *if for all $a_1, a_2 \in M$*

$$p \equiv a_1 + a_2 \implies p \equiv a_1 \text{ or } p \equiv a_2.$$

Note that 0 is a prime, indecomposable element in any monoid. A simple induction shows that these properties extend to arbitrary finite sums:

- If p is prime and $p \leq a_1 + a_2 + \ldots + a_n$ then there is some $i \in \{1, 2, \ldots, n\}$ such that $p \leq a_i$.
- If p is indecomposable and $p \equiv a_1 + a_2 + \ldots + a_n$ then there is some $i \in \{1, 2, \ldots, n\}$ such that $p \equiv a_i$.

Proposition 11.2. *Let M be a monoid and $p \in M$. Then*

$$p \text{ prime} \implies p \text{ indecomposable.}$$

Proof. If $p \equiv a_1 + a_2$, then $a_1 + a_2 \leq p$ and $p \leq a_1 + a_2$. The first inequality implies $a_1, a_2 \leq p$. The second and the primeness of p imply that either $p \leq a_1$ or $p \leq a_2$. Thus we have either $p \equiv a_1$ or $p \equiv a_2$. □

Proposition 11.3. *Let M be a monoid and $p \in M$. Then*

1. *p is prime in M $\iff$ $\{\equiv p\}$ is prime in $\overline{M}$ $\implies$ $\{\asymp p\}$ is prime in $\widetilde{M}$*
2. *p is indecomposable in M $\iff$ $\{\equiv p\}$ is indecomposable in $\overline{M}$*

Proof. The claims involving $\overline{M}$ are easy consequences of the fact that $\{\equiv a\} \leq \{\equiv b\}$ in $\overline{M}$ if and only if $a \leq b$ in M.

For the remaining claim, we have the rule $\{\asymp a\} \leq \{\asymp b\}$ in $\widetilde{M}$ if and only if $a \prec b$ in M. Thus, if $\{\asymp p\} \leq \{\asymp a_1\} + \{\asymp a_2\}$ then $p \prec a_1 + a_2$ and there is some $n \in \mathbb{N}$ such that

$p \leq n(a_1 + a_2)$. Since p is prime, we have either $p \leq a_1$ or $p \leq a_2$. So either $\{\asymp p\} \leq \{\asymp a_1\}$ or $\{\asymp p\} \leq \{\asymp a_2\}$. $\square$

Proposition 11.4. *Let p be an indecomposable element of a monoid M. Then for any $a \in M$,*

$$a \leq p \implies a \equiv p \text{ or } a \ll p.$$

Proof. Let $b \in M$ be such that $a + b = p$. Then either $a \equiv p$ or $b \equiv p$. In the second case, we get $a + p \equiv a + b = p$ so $a \ll p$. $\square$

Proposition 11.5. *Let I be an order ideal in M and $p \in M$. Then*

1. *p prime $\implies$ $[p]_I$ prime in M/I*
2. *If $p \in I$, then p is indecomposable in I if and only if p is indecomposable in M.*

Proof.

1. Suppose $[p]_I \leq [a_1]_I + [a_2]_I$ in M/I. Then there is some $u \in I$ such that $p \leq a_1 + a_2 + u$. Since p is prime, we have either $p \leq a_1$, $p \leq a_2$ or $p \leq u$. In the first two cases we get either $[p]_I \leq [a_1]_I$ or $[p]_I \leq [a_2]_I$. In the last case, $p \leq u \in I$ so $p \in I$ and $[p]_I = 0 \leq [a_1]_I, [a_2]_I$. Therefore $[p]_I$ is a prime in M/I.
2. Suppose p is indecomposable in I, and $p \equiv a_1 + a_2$ for some $a_1, a_2 \in M$. Since $a_1, a_2 \leq p \in I$, we get $a_1, a_2 \in I$, and so by the indecomposability of p in I, $p \equiv a_1$ or $p \equiv a_2$. This makes p indecomposable in M.

$\square$

In a decomposition monoid we have the simpler situation that primes and indecomposables are the same:

Proposition 11.6. *Let M be a decomposition monoid, $I \leq M$ an order ideal and $p \in M$. Then*

1. *p is prime $\iff$ p is indecomposable.*
2. *p is indecomposable $\implies$ $[p]_I$ is indecomposable in M/I.*
3. *If $p \in I$, then p is a prime element of I if and only if it is a prime element of M*

Proof.

1. We already know that p prime implies p indecomposable. For the converse, suppose p is indecomposable and $p \leq a + b$ for some $a_1, a_2 \in M$. Then there are p_1, p_2 such that $p_1 \leq a_1$, $p_2 \leq a_2$,and $p = p_1 + p_2$. But then $p \leq p_1 \leq a_1$ or $p \leq p_2 \leq a_2$. Thus p is prime.
2. This follows from 1 and 11.5.
3. This follows from 1 and 11.5.

$\square$

Since we will be working almost entirely with decomposition and refinement monoids, the distinction between primes and indecomposables will seldom appear.

Definition 11.7. *An element of a monoid is* **primely generated** *if it is a sum of prime elements. A monoid is* **primely generated** *if all its elements are primely generated.*

One of the main themes of this section is that, in refinement monoids, the existence of primely generated elements leads to cancellation properties of the types discussed in Section 9, namely, weak cancellation and midseparativity. We begin with weak cancellation:

Proposition 11.8. *Let M be a refinement monoid, $a, b, c \in M$ with c primely generated. If $a + c = b + c$, then there is a refinement matrix*

$$\begin{array}{c c} & \begin{matrix} b & c \end{matrix} \\ \begin{matrix} a \\ c \end{matrix} & \begin{pmatrix} * & * \\ * & c_1 \end{pmatrix} \end{array}$$

with $c \le c_1$.

Proof. Let $\mathcal{C}$ be the class of all elements $c \in M$ such that $a + c = b + c$ implies the existence of a refinement as above. Clearly $0 \in \mathcal{C}$. We will show that if $c \in \mathcal{C}$ and $p \in M$ is prime, then $c + p \in \mathcal{C}$...

Suppose we have $a + c + p = b + c + p$ for some $c \in \mathcal{C}$ and prime p. Since $c \in \mathcal{C}$, there is a refinement

$$\begin{array}{c c} & \begin{matrix} b+p & c \end{matrix} \\ \begin{matrix} a+p \\ c \end{matrix} & \begin{pmatrix} * & * \\ * & c_1 \end{pmatrix} \end{array}$$

with $c \le c_1$. Refining further we can get a new refinement matrix, still with c_1 as an entry:

$$\begin{array}{c c} & \begin{matrix} b & p & c \end{matrix} \\ \begin{matrix} a \\ p \\ c \end{matrix} & \begin{pmatrix} d_2 & a_2 & a_3 \\ b_2 & p_1 & p_2 \\ b_3 & p_3 & c_1 \end{pmatrix} \end{array}$$

Note that $p = b_2 + p_1 + p_2 = a_2 + p_1 + p_3$. Now we consider two cases:

- If $p \le p_1$ or $p \le p_2$ or $p \le p_3$, then $c + p \le c_1 + p_1 + p_2 + p_3$ and

$$\begin{array}{c c} & \begin{matrix} b & c+p \end{matrix} \\ \begin{matrix} a \\ c+p \end{matrix} & \begin{pmatrix} d_2 & a_2 + a_3 \\ b_2 + b_3 & c_1 + p_1 + p_2 + p_3 \end{pmatrix} \end{array}$$

  is a refinement of the form we seek.
- If $p \not\le p_1$ and $p \not\le p_2$ and $p \not\le p_3$, then since p is prime we must have $p \le a_2$ and $p \le b_2$. Hence $a_2 = p + a_4$ and $b_2 = p + b_4$ for some $a_4, b_4 \in M$, and

$$\begin{array}{c c} & \begin{matrix} b & c+p \end{matrix} \\ \begin{matrix} a \\ c+p \end{matrix} & \begin{pmatrix} d_2 + p & a_3 + a_4 \\ b_3 + b_4 & p + c_1 + p_1 + p_2 + p_3 \end{pmatrix} \end{array}$$

  is a refinement of the form we seek.

We have shown therefore that $c + p \in \mathcal{C}$, and by induction, that any primely generated element is in $\mathcal{C}$. $\square$

Corollary 11.9.

1. *Any primely generated refinement monoid is weakly cancellative.*
2. *If M is a primely generated refinement monoid, then so is $\overline{M}$.*

Proof.

1. This is immediate from 11.8 and the definition of weak cancellation, 9.1.
2. Since primes in M map to primes in $\overline{M}$ (11.3.1), it is easy to see that $\overline{M}$ is primely generated. Proposition 9.14 and 1 then show that $\overline{M}$ has refinement.

□

There are versions of the weak cancellation properties discussed in Section 9 that apply to refinement monoids containing primely generated elements:

Corollary 11.10. *Let M be a refinement monoid and $a, b, c, c_1, c_2 \in M$. Suppose c, c_1, c_2 are primely generated.*

1. *If $a + c_1 + c_2 = b + c_1 + c_2$, then there is a refinement matrix*

$$\begin{array}{c c} & \begin{array}{ccc} b & c_1 & c_2 \end{array} \\ \begin{array}{c} a \\ c_1 \\ c_2 \end{array} & \begin{pmatrix} * & * & * \\ * & c_1' & * \\ * & * & c_2' \end{pmatrix} \end{array}$$

such that $c_1 \leq c_1'$ and $c_2 \leq c_2'$.

2. $a + c \leq b + c \implies (\exists a_1 \ll c$ such that $a \leq b + a_1)$
3. $a \ll c_1 + c_2 \implies (\exists a_1, a_2$ such that $a = a_1 + a_2,\ a_1 \ll c_1$ and $a_2 \ll c_2)$
4. $\{\ll c_1 + c_2\} = \{\ll c_1\} + \{\ll c_2\}$
5. $a \equiv c_1 + c_2 \implies (\exists a_1, a_2$ such that $a = a_1 + a_2,\ a_1 \equiv c_1$ and $a_2 \equiv c_2)$

Proof. The proofs are exactly the same as for 9.12 and 9.13, except for the use of 11.8 in place of the weak cancellation hypothesis. □

Corollary 11.11. *Let a and b be elements of a refinement monoid such that $a \equiv b$. Then a is primely generated if and only if b is primely generated.*

Proof. Suppose $a = p_1 + p_2 + \ldots + p_K$ for some primes $p_1, p_2, \ldots, p_K \in M$. Then a simple induction from 11.10.5 shows that there are $p_1', p_2', \ldots, p_K'$ such that $b = p_1' + p_2' + \ldots + p_K'$ and $p_i \equiv p_i'$ for $i = 1, 2, \ldots, K$. This last condition says that p_i' is prime for all i and so b is primely generated. □

In 11.33, we will extend this result to elements such that $a \asymp b$.

Lemma 11.12. *Let M be a refinement monoid and $a, b, c \in M$. Suppose c is primely generated.*

1. *If $a + c = b + c$ and $c \leq a, b$, then $a = b$.*
2. *If $a + nc = b + nc$ for some $n \in \mathbb{N}$, then $a + c = b + c$.*
3. *If $a + c = b + c$ and $c \prec a, b$, then $a = b$.*

Proof.

1. From 11.8, there is a refinement matrix

$$\begin{array}{c c} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c \leq c_1$. In particular, $a_1 + c \leq a_1 + c_1 = c$ so that $a_1 \ll c$. From 8.3, we get $a + a_1 = b + a_1$. So, with $a_1 \ll c \leq a, b$, 8.2.1 implies that $a = b$.

2. Follows by induction from 1.
3. Suppose $c = p_1 + p_2 + \ldots + p_K$ for some primes $p_1, p_2, \ldots, p_K$. Since $c \prec a, b$ there is some $n \in \mathbb{N}$ such that $c \le na, nb$. For each $i \le K$, we have $p_i \le c \le na, nb$, so $p_i \le a, b$. Using 1, we can cancel the elements $p_1, p_2, \ldots, p_K$ one at a time from the equation $a + p_1 + p_2 + \ldots + p_K = b + p_1 + p_2 + \ldots + p_K$ to get $a = b$.

□

Proposition 11.13. *Let M be a refinement monoid and $a, b, c, d \in M$. Suppose d is primely generated.*

1. $(a + c = b + c$ and $c \prec d \prec a, b) \implies a = b$.
2. *The Archimedean component $\{\asymp d\}$ is cancellative.*
3. *If a, b or c is in $\{\asymp d\}$, then*
$$(a + c = b + c \text{ and } c \prec a, b) \implies a = b.$$
4. *If a or b is in $\{\asymp d\}$, then*
$$(2a = a + b = 2b) \implies a = b.$$
5. *If $c \in \{\asymp d\}$, then c is either free or regular.*
6. *If $a, b \in \{\asymp d\}$, then the congruences $\sim_a$ and $\sim_b$ coincide. In particular, $G_a = G_b$, $H_a = H_b$ and $\{\ll a\} = \{\ll b\}$.*

Proof.

1. Let $n \in \mathbb{N}$ be such that $c \le nd$. Then $a + nd = b + nd$, and $nd \prec d \prec a, b$. Since nd is primely generated, 11.12.3 implies that $a = b$.
2. If $a + c = b + c$ with $a, b, c \in \{\asymp d\}$, then $c \prec d \prec a, b$, and from 1, $a = b$.
3. We have two cases to prove:
    - Suppose $c \in \{\asymp d\}$. Then, in particular, we have $c \prec d \prec c \prec a, b$. So applying 1, we get $a = b$.
    - Suppose $a \in \{\asymp d\}$. Let $n \in \mathbb{N}$ be such that $c \le na, nb$. Since $c \le nb$, we get $a + nb = b + nb = (n+1)b$, so $a \prec b$. Since $c \le na$, we also get $a + na = b + na$. Using 1 with the relation $na \prec a \prec d \prec a \prec a, b$, we can cancel na from this equation to get $a = b$.
4. The equation $2a = 2b$ implies that a and b are in the same Archimedean component, namely, $\{\asymp d\}$. Using 2, we can cancel a from $a + a = a + b$, to get $a = b$.
5. If c is not free, then there are $m, n \in \mathbb{N}$ such that $m > n$ and $mc \le nc$. If $n = 1$ then c is regular and we are done.

   Otherwise we write the inequality as $(m - n + 1)c + x + (n - 1)c = c + (n - 1)c$ for some $x \in M$. We have $(n-1)c \prec c \prec d \prec c, (m - n + 1)c + x$, so we can use 1 to cancel $(n-1)c$ from this equation to get $(m - n + 1)c + x = c$. Since $m - n + 1 \ge 2$, this implies that c is regular.
6. Let $n \in \mathbb{N}$ be such that $a \le nb$. If $x \sim_a y$, then $x + a = y + a$, and so $x + nb = y + nb$. Since $(n-1)b \prec b \prec d \prec b \prec b + x, b + y$, we can use 1 to get $b + x = b + y$, that is, $x \sim_b y$. By symmetry, $x \sim_b y$ implies $x \sim_a y$. The rest of the claim then follows from 10.2 and 10.3.1.

□

Many of the statements of this proposition will seem more natural when we prove in 11.33 that if d is primely generated then all elements of $\{\asymp d\}$ are also primely generated.

Next we show that the existence of a primely generated element in a refinement monoid gives rise to a cancellation property similar to midseparativity:

Proposition 11.14. *Let M be a refinement monoid, $a, b, c \in M$ with c primely generated. If $a + c = b + c$ and $c \le a$, then there is a primely generated idempotent $e \le c$ such that $a = b + e$.*

Proof. Let $\mathcal{C}$ be the class of all elements $c \in M$ such that $a + c = b + c$ and $c \le a$ implies the existence of a primely generated idempotent $e \le c$ with $a = b + e$. Clearly $0 \in \mathcal{C}$. We will show that if $c \in \mathcal{C}$ and $p \in M$ is prime, then $c + p \in \mathcal{C}$...

Suppose we have $a + c + p = b + c + p$ with $c + p \le a$ for some $c \in \mathcal{C}$ and prime p. Since $c \in \mathcal{C}$ and $c \le a \le a + p$, there is a primely generated idempotent $e \le c$ with $a + p = b + p + e$. We have $p \le a$ so, using 8.5.1, we can make a refinement

$$\begin{array}{cc} & \begin{array}{cc} b+e & p \end{array} \\ \begin{array}{c} a \\ p \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & p_1 \end{pmatrix} \end{array}$$

with $p_1 \le a_1$.

Since p is prime we get two cases:

- If $p \not\le p_1$, then we must have $p \le a_1, b_1$ and hence $p \le a, b + e$. Using 11.12.1, we can then cancel p from the equation $a + p = b + p + e$ to give $a = b + e$. We also have $e \le c \le c + p$.
- If $p \le p_1$, then we have $2p \le 2p_1 \le a_1 + p_1 = p$, so $p \ll p$. Since we also have $p \le a$, we can use 8.2.2 to get $f \equiv p$ such that $2f = f$ and $a = b + e + f$. The element f is prime and so $e + f$ is a primely generated element such that $2(e + f) = (e + f)$. We also have $f \le p$, so $e + f \le c + p$.

We have shown therefore that $c + p \in \mathcal{C}$, and by induction, that any primely generated element is in $\mathcal{C}$. $\square$

Corollary 11.15.

1. *Any primely generated refinement monoid is midseparative.*
2. *A primely generated refinement monoid is strongly separative if and only if it has no proper regular elements.*

Proof. Directly from 11.14 and 9.6. $\square$

We now consider another kind of cancellation that has appeared only briefly in our discussion so far:

Definition 11.16. *A monoid M has* **$\le$-multiplicative cancellation** *if*

$$(\forall n \in \mathbb{N})(\forall a, b \in M)\ (na \le nb \implies a \le b).$$

Before proving that primely generated monoids have $\leq$-multiplicative cancellation, we will show that (ordinary) cancellation and $\leq$-multiplicative cancellation are distinct properties for refinement monoids...

The refinement monoid $M = \{0, \infty\}$ is not cancellative but has $\leq$-multiplicative cancellation simply because $na = a$ for all $a \in M$ and $n \in \mathbb{N}$. For an example of a refinement monoid which is cancellative but does not have $\leq$-multiplicative cancellation we have to work a bit harder:

Example 11.17. *Let $M = \mathbb{R}^{++} \times \mathbb{Z}_2$. Since $\mathbb{R}^{++}$ and $\mathbb{Z}_2$ are refinement semigroups (see 7.6), so is M. Let M^0 be the refinement monoid obtained by adjoining a zero element to M.*

That M^0 is cancellative can be checked directly or by recognizing that M^0 is isomorphic to the submonoid obtained by deleting the element $(0, 1)$ from the cancellative monoid $\mathbb{R}^+ \times \mathbb{Z}_2$.

Also easily checked is that for $(r_1, x_1), (r_2, x_2) \in M$ with the minimum preorder of M^0,

$$(r_1, x_1) \leq (r_2, x_2) \iff (r_1 = r_2 \text{ and } x_1 = x_2) \text{ or } r_1 < r_2.$$

Set $a = (1, 0)$ and $b = (1, 1)$. Then we have $2a = 2b$ but $a \not\leq b$, thus M^0 does not have $\leq$-multiplicative cancellation.

The above monoid has $(n = 3)$ $\leq$-multiplicative cancellation, that is

$$(\forall a, b \in M^0)\ (3a \leq 3b \implies a \leq b).$$

By replacing the group $\mathbb{Z}_2$ by other Abelian groups in the above construction, monoids which fail $\leq$-multiplicative cancellation in other ways can be constructed. Another monoid of this type can be constructed as a submonoid of $(\mathbb{C}, \cdot)$, the set of complex numbers with multiplication as operation. Set

$$M = \{z \in \mathbb{C} \mid |z| > 1\} \cup \{1\} \subseteq \mathbb{C}.$$

Then M is a cancellative refinement monoid which fails $\leq$-multiplicative cancellation for all $n > 1$.

We next establish a $\leq$-multiplicative cancellation property for primely generated elements of refinement monoids:

Proposition 11.18. *Let a be a primely generated element of a refinement monoid M. Then for all $b \in M$ and $n \in \mathbb{N}$,*

$$na \leq nb \implies a \leq b.$$

Proof. Let $\mathcal{A}$ be the class of all primely generated elements $a \in M$ such that for all $b \in M$ and $n \in \mathbb{N}$, $na \leq nb$ implies $a \leq b$. Clearly $0 \in \mathcal{A}$. We will show that if $a \in \mathcal{A}$ and $p \in M$ is prime, then $a + p \in \mathcal{A}$...

Suppose we have $n(a + p) \leq nb$ for some $a \in \mathcal{A}$ and prime p. Since $na \leq nb$ we have $a \leq b$, so $a + x = b$ for some $x \in M$. Thus $na + np \leq na + nx$ and there is some $u \in M$ such that $na + np + u = na + nx$. The element na is primely generated, so using 11.8, we can make a refinement of this equation,

$$\begin{array}{cc} & \begin{array}{cc} nx & na \end{array} \\ \begin{array}{c} np + u \\ na \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

such that $na \leq c_1$. From the inequality $p \leq np + u = d_1 + a_1$ we get two cases:

- If $p \leq d_1$, then $p \leq nx$ and thus $p \leq x$. Consequently, $a + p \leq a + x = b$.
- If $p \leq a_1$, then $na + p \leq c_1 + a_1 = na$. From 11.12.2, this implies $a + p \leq a \leq b$.

We have shown therefore that $a + p \in \mathcal{A}$, and by induction, that any primely generated element is in $\mathcal{A}$. $\square$

As an immediate corollary we get

Corollary 11.19. *Any primely generated refinement monoid has $\leq$-multiplicative cancellation.*

If a is a primely generated element of a monoid, there will in general be many ways of expressing a as a sum of primes. Thus we consider the question of whether there is a canonical form for such elements. Among all such expressions for a there are certainly expressions which have the least possible number of terms. These we will call **minimum prime expressions** for a.

Proposition 11.20. *Let a be a primely generated element of a monoid, and $a = \sum_{i=1}^{K} p_i$ a minimum prime expression for a. Then for any $i, j \leq K$,*

$$p_i \leq p_j \implies p_i \equiv p_j.$$

Any two such expressions contain the same primes up to $\equiv$, that is, if

$$a = \sum_{i=1}^{K} p_i = \sum_{i=1}^{K} q_i$$

are minimum prime expressions for a, then for any $i \leq K$ there is some $j \leq K$ such that $p_i \equiv q_j$.

Proof. Suppose $a = \sum_{i=1}^{K} p_i$ has minimum length and $p_i \leq p_j$ with $i \neq j$. By 11.2 and 11.4, $p_i \equiv p_j$ or $p_i \ll p_j$. In the second case, $p_i + p_j \equiv p_j$, so $p_i + p_j$ would be a prime and we could shorten the expression for a. This contradicts the minimality of the expression, so we must have $p_i \equiv p_j$.

Further, if there are two minimum prime expressions for a as above, and $i \leq K$, then, since $p_i \leq \sum_{j=1}^{K} q_j$, there is some $j \leq K$ such that $p_i \leq q_j$. Similarly, there is some $k \leq K$ such that $q_j \leq p_k$. Thus $p_i \leq q_j \leq p_k$, and $p_i \equiv q_j \equiv p_k$. $\square$

This proposition does not rule out the possibility that an element could have two minimal prime expressions in which the multiplicities of the primes which appear are different. The following gives just such a monoid:

Example 11.21. *Let M be the monoid generated by two elements p and q such that $3p = 2p$, $3q = 2q$ and $2p + q = 2q + p$. Then M has seven elements: $\{0, p, 2p, q, 2q, p + q, 2p + q\}$ and is partially ordered. 0, p and q are the prime elements, and 0, $2p$, $2q$ and $2p + q$ are the regular elements of M. Thus $2p + q$ and $2q + p$ are two minimum prime expressions for the same element of M.*

The complication shown by this example disappears in refinement monoids. To prove this, we will define functions n_p, one for each prime p, which will enable us to get a handle on the primes which appear in expressions for primely generated elements.

Definition 11.22. *For any element p of a monoid M, define a function $n_p\colon M \to (\mathbb{Z}^+)^\infty$ by*

$$n_p(a) = \sup\{n \in \mathbb{Z}^+ \mid np \leq a\}$$

for all $a \in M$.

As easy consequences of the definition we get

Proposition 11.23. *Let a, a', p, p' be elements of a monoid.*

1. *$n_p(a) = 0$ if and only if $p \not\leq a$*
2. *If $p \leq 0$ then $n_p(a) = \infty$*
3. *If $p \ll p$ then $n_p(a) \in \{0, \infty\}$*
4. *If $p' \equiv p$ and $a' \equiv a$ then $n_p(a) = n_{p'}(a')$*

Proposition 11.24. *If M is a refinement monoid and $p \in M$ is prime such that $p \not\leq 0$, then n_p is a monoid homomorphism.*

Proof. Since $p \not\leq 0$ we have $n_p(0) = 0$, so we need to show only that $n_p(a_1 + a_2) = n_p(a_1) + n_p(a_2)$ for all $a_1, a_2 \in M$...

Suppose $n_1 p \leq a_1$ and $n_2 p \leq a_2$ for some $n_1, n_2 \in \mathbb{Z}^+$. Then $(n_1 + n_2)p \leq a_1 + a_2$, and hence $n_1 + n_2 \leq n_p(a_1 + a_2)$. Taking the supremum over all such n_1 and n_2 gives the inequality $n_p(a_1) + n_p(a_2) \leq n_p(a_1 + a_2)$.

To show the opposite inequality, suppose $np \leq a_1 + a_2$ for some $n \in \mathbb{Z}^+$. Then we get the refinement matrix

$$\begin{array}{cc} & \begin{array}{cccc} p & p & \dots & p \end{array} \\ \begin{array}{c} a_1 \geq \\ a_2 \geq \end{array} & \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{pmatrix} \end{array}$$

Since p is prime, for each i, we have either $p \leq x_i$ or $p \leq y_i$. So there are some $n_1, n_2 \in \mathbb{Z}^+$ such that $n = n_1 + n_2$, $n_1 p \leq a_1$ and $n_2 p \leq a_2$. Thus $n \leq n_p(a_1) + n_p(a_2)$, and taking the supremum over all such n we get $n_p(a_1 + a_2) \leq n_p(a_1) + n_p(a_2)$. □

Proposition 11.25. *If p and q are primes in a refinement monoid, then $n_p(q) \in \{0, 1, \infty\}$ and*

1. $n_p(q) = 0 \iff p \not\leq q$
2. $n_p(q) = 1 \iff$ ($p \equiv q$ and p is free).
3. $n_p(q) = \infty \iff p \ll q$.

Proof. If $p \leq 0$ then $n_p(q) = \infty$ and the claim is trivially true. So it remains to consider the case where $p \not\leq 0$ and hence n_p is a homomorphism:

1. This we have already noted in 11.23.
2. If $n_p(q) = 1$ then $p \leq q$, so by 11.3, either $p \ll q$ or $p \equiv q$. The first case implies $n_p(q) = \infty$, so we must be in the second case. To show freeness, suppose $mp \leq np$ for some $m, n \in \mathbb{N}$, then $mq \leq np \leq nq$, and so $m = n_p(mq) \leq n_p(nq) = n$.

   Conversely, if $q \equiv p$ and p is free, then for any $n \in \mathbb{N}$, $np \leq q \iff np \leq p \iff n \leq 1$. Thus $n_p(q) = 1$.

3. Notice that if $n_p(q) \geq 2$ then $2p \leq q$ and, by 11.3, $p \equiv q$ or $p \ll q$. But in the first case, $q + p \equiv 2p \leq q$. So in either case, $p \ll q$. Conversely, if $p \ll q$, then a simple induction shows that $np + q \leq q$ for all $n \in \mathbb{N}$, so $n_p(q) = \infty$.

Finally we note that the argument of 3 shows that $n_p(q) \in \{0, 1, \infty\}$. □

Corollary 11.26. *If p is a prime in a refinement monoid, then $n_p(p) \in \{1, \infty\}$, and*

1. $n_p(p) = 1 \iff p$ *is free.*
2. $n_p(p) = \infty \iff p$ *is regular.*

With the homomorphisms n_p in hand, we return to the question of uniqueness of minimal prime expressions–this time for refinement monoids...

Proposition 11.27. *Let a be a primely generated element in a refinement monoid. Then any minimum prime expression for a is unique up to $\equiv$-equivalence. Specifically, if*

$$a = \sum_{i=1}^{K} p_i = \sum_{i=1}^{K} q_i$$

are minimum prime expressions for a, then after suitable renumbering, we get $p_i \equiv q_i$ for all $i \leq K$.

Proof. Fix a p_I. From 11.2, there is some q_J such that $q_J \equiv p_I$. Now we consider the following two cases:

- p_I free: For any $i \leq K$, we have either $p_I \not\leq p_i$ and $n_{p_I}(p_i) = 0$, or $p_I \leq p_i$, which implies $p_i \equiv p_I$ and $n_{p_I}(p_i) = n_{p_I}(p_I) = 1$. Thus $n_{p_I}(a) = \sum_i n_{p_I}(p_i)$ is just equal to the number of terms in $\sum_i p_i$ which are $\equiv$-equivalent to p_I. Similarly, $n_{q_J}(a)$ is the number of terms in $\sum_j q_j$ which are $\equiv$-equivalent to q_J. But $p_I \equiv q_J$, so $n_{p_I}(a) = n_{q_J}(a)$ and $\sum_i p_i$ and $\sum_j q_j$ have the same number of terms $\equiv$-equivalent to p_I. Thus free primes, and their multiplicities, match up in the two minimum prime expressions for a.
- p_I regular: In this case, there is no other p_i such that $p_i \equiv p_I$, or q_j such that $q_j \equiv q_J$, since otherwise, we could shorten the minimum prime expressions. Thus the sums $\sum_i p_i$ and $\sum_j q_j$ each have exactly one term $\equiv$-equivalent to p_I.

□

One might hope that it would be possible to collect $\equiv$-equivalent primes together so that any primely generated element could be written in the form $a = n_1p_1 + n_2p_2 + \ldots + n_Kp_K$ where $n_i \in \mathbb{N}$ and $p_1, p_2, \ldots, p_K$ are primes such that $p_i \leq p_j$ implies $p_i = p_j$. This is not always possible: Let $M = \mathbb{Z} \setminus \{0\}$ with multiplication as its monoid operation. Then one easily checks that 2 and -2 are $\equiv$-equivalent primes and so -4 is primely generated (as is any element), but there is no element $p \in M$ such that $-4 = p^2$.

To state the best possible result of this type we will call a set X of prime elements of a monoid **incomparable** if for all $p, p' \in X$, $p \leq p'$ implies $p = p'$.

Proposition 11.28. *Let a be a primely generated element in a refinement monoid. Then*

$$a \equiv n_1p_1 + n_2p_2 + \ldots + n_Kp_K + q_1 + q_2 + \ldots + q_L$$

where $\{p_1, p_2, \ldots, p_K, q_1, q_2, \ldots, q_L\}$ is a set of incomparable primes with p_i free, q_i regular and $n_i = n_{p_i}(a) \in \mathbb{N}$ for all i. In any expression for a of this form, the numbers

$K, L, n_1, n_2, \ldots, n_K$ are uniquely determined by a and the primes are determined up to $\equiv$-equivalence.

Proof. From 11.27 we have that there is a minimum prime expression for a in the form $a = \sum_i p_i$. From this sum we will construct the required expression...

We first relabel the regular primes as $q_1, q_2, \ldots, q_L$. Note that $q_i \equiv q_j$ implies $i = j$, since otherwise $q_i + q_j$ is prime and we could shorten the minimum prime expression for a.

Among the remaining primes we pick a representative of each $\equiv$-equivalence class and label these primes $p_1, p_2, \ldots, p_K$. Thus any prime appearing in the original minimum prime expression for a is $\equiv$-equivalent to a unique element of $X = \{p_1, p_2, \ldots, p_K, q_1, q_2, \ldots, q_L\}$.

It is easy to see that X is a set of incomparable primes. As in 11.27, for each free prime p_i, $n_i = n_{p_i}(a)$ is the number of primes $\equiv$-equivalent to p_i in the minimal prime expression. Thus

$$a \equiv n_1 p_1 + n_2 p_2 + \ldots + n_K p_K + q_1 + q_2 + \ldots + q_L.$$

We now consider the uniqueness of such expressions...

Suppose we have, in addition, the expression

$$a \equiv n'_1 p'_1 + n'_2 p'_2 + \ldots + n'_{K'} p'_{K'} + q'_1 + q'_2 + \ldots + q'_{L'}$$

where $X' = \{p'_1, p'_2, \ldots, p'_{K'}, q'_1, q'_2, \ldots, q'_{L'}\}$ is a set of incomparable primes with p'_i free, q'_i regular and $n_i \in \mathbb{N}$ for all i. The argument used in 11.20 shows that there is a bijection between the primes of X and the primes of X', and so, after suitable relabeling, we can assume that $K' = K$, $L' = L$, $p_i \equiv p'_i$ and $q_i \equiv q'_i$ for all i. A simple calculation then shows that

$$n_i = n_{p_i}(a) = n_{p'_i}(a) = n'_i.$$

□

An expression of the form given in this proposition will be called a **canonical form** for a. We will see in 11.37 that the number of free primes, that is, the number K, in a canonical form for a plays a crucial role in determining the structure of $\{\asymp a\}$. Regular primes in the canonical form will take the sidelines. In the extreme case, the Archimedean component of a primely generated regular element is trivial in the sense of the following proposition:

Proposition 11.29. *Let a be a primely generated regular element in a refinement monoid. Then*

$$\{\asymp a\} = \{\equiv a\} \cong G_a = H_a \cong G(\{\prec a\}).$$

Proof. We will show first that $\{\asymp a\} \subseteq \{\equiv a\}$. The opposite inclusion is trivially true.

Suppose $b \prec a \prec b$. Then $b \leq na$ for some $n \in \mathbb{N}$. Since a is regular, we have $na \leq a$ and so $b \leq a$.

Also there is some $m \in \mathbb{N}$ such that $a \leq mb$. This gives $ma \leq mb$, and using 11.18 we get $a \leq b$.

The remaining claims of this proposition follow directly from 10.5 and 10.12. □

The structure of primely generated regular (6.7) refinement monoids has been studied by H. Dobbertin [8].

We remark that if $a = q_1 + q_2 + \ldots + q_L$ for a set $\{q_1, q_2, \ldots, q_L\}$ of incomparable regular primes, then $\{\asymp a\} = \{\equiv a\}$, even without refinement. Without the refinement hypothesis, however, not every primely generated regular element is a sum of regular primes. Consider, for example, the monoid $M = \{0, 1, \infty\}$ with $1 + 1 = 1 + \infty = \infty + \infty = \infty$. The element

1 is a prime in M but is neither free nor regular. The element ∞ is a primely generated regular element of M which is not a sum of regular primes. Further, $\{\asymp \infty\} = \{1, \infty\}$ and $\{\equiv \infty\} = \{\infty\}$.

We turn next to the structure of $\{\asymp a\}$ when a is a <u>free</u> primely generated element of a refinement monoid M. From 11.28, we have

$$a \equiv n_1 p_1 + n_2 p_2 + \ldots + n_K p_K + q_1 + q_2 + \ldots + q_L$$

where $\{p_1, p_2, \ldots, p_K, q_1, q_2, \ldots, q_L\}$ is a set of incomparable primes with p_i free, q_i regular and $n_i \in \mathbb{N}$ for all i. Since a is free we must have $K \geq 1$. We will write $q = q_1 + q_2 + \ldots + q_L$ and $f = n_1 p_1 + n_2 p_2 + \ldots + n_K p_K$, so $a \equiv f + q$ and $a + q \leq a$.

For each prime p_i there is a monoid homomorphism $n_{p_i}\colon M \to (\mathbb{Z}^+)^\infty$ defined in 11.22. We note again that $n_{p_i}(p_j) = \delta_{ij}$, $n_{p_i}(q_j) = 0$ and $n_{p_i}(a) = n_i$ for all i, j. We can combine these to make a homomorphism $N_a\colon M \to ((\mathbb{Z}^+)^\infty)^K$ by defining

$$N_a(x) = (n_{p_1}(x), n_{p_2}(x), \ldots, n_{p_K}(x))$$

for all $x \in M$.

We will henceforth restrict our attention to the order ideal $\{\prec a\}\ldots$

If $x \prec a$, then $x \leq ma$ for some $m \in \mathbb{N}$, so $n_{p_i}(x) \leq n_{p_i}(ma) = mn_i < \infty$. Thus the homomorphism N_a maps $\{\prec a\}$ into $(\mathbb{Z}^+)^K$. In particular, $N_a(\{\prec a\})$ is cancellative.

It will be useful to define a right inverse for N_a, $\nu\colon (\mathbb{Z}^+)^K \to \{\prec a\}$ by

$$\nu(\mathbf{m}) = m_1 p_1 + m_2 p_2 + \ldots + m_K p_K,$$

where $\mathbf{m} = (m_1, m_2, \ldots, m_K) \in (\mathbb{Z}^+)^K$. One readily confirms that $N_a \circ \nu$ is the identity on $(\mathbb{Z}^+)^K$, so that $N_a(\{\prec a\}) = (\mathbb{Z}^+)^K$, and also $\nu(N_a(a)) = f$. Notice also that ν depends on the particular primes used in its definition, whereas N_a depends only on a and the order of the primes $p_1, p_2, \ldots, p_K$.

Proposition 11.30. *Let a be a free primely generated element in a refinement monoid M. Let $a \equiv n_1 p_1 + n_2 p_2 + \ldots + n_K p_K + q_1 + q_2 + \ldots + q_L$, q, f, N_a and ν be as above. Then for all $x \prec a$,*

1. $\nu(N_a(x)) \leq x$
2. $N_a(x) = 0 \iff x \ll a$
3. $x = \nu(N_a(x)) + u$ *for some* $u \ll a$.

Proof.

1. We want to show that $n_{p_1}(x)p_1 + n_{p_2}(x)p_2 + \ldots + n_{p_K}(x)p_K \leq x$.

   By the definition of n_{p_1}, we have $n_{p_1}(x)p_1 \leq x$, so there is some $x_1 \in M$ such that $n_{p_1}(x)p_1 + x_1 = x$. Applying the homomorphism n_{p_2} to this equation gives $n_{p_2}(x_1) = n_{p_2}(x)$. Thus there is some $x_2 \in M$ such that $x_1 = n_{p_2}(x)p_2 + x_2$, that is, $n_{p_1}(x)p_1 + n_{p_2}(x)p_2 + x_2 = x$. Repeating this process in the obvious way gives the required inequality.
2. If $x \ll a$, then $x + a \leq a$. Applying N_a we get $N_a(a) + N_a(x) \leq N_a(a)$. Thus $N_a(x) = 0$.

   Conversely, suppose $N_a(x) = 0$. Since $x \prec a$, there are $y \prec a$ and $m \in \mathbb{N}$ such that $x + y = ma$. Applying the homomorphism N_a to this equation gives $N_a(y) = mN_a(a)$, so with 1, we have $y \geq \nu(N_a(y)) = m\nu(N_a(a)) = mf$. Thus $x + mf \leq x + y = ma$. Adding mq to this yields $x + ma \leq x + mf + mq \leq ma + mq \leq ma$. Finally, using 11.12.2, we can cancel $(n-1)a$ from this inequality to get $x + a \leq a$.

3. From 1, we know that $x = \nu(N_a(x)) + u$ for some $u \in M$. We need to check only that $u \ll a$...

   Applying N_a to this equation yields $N_a(x) = N_a(x) + N_a(u)$. So $N_a(u) = 0$, and from 2, $u \ll a$.

□

We specialize 11.30 further to the case of $x \in \{\asymp a\}$...

Proposition 11.31. *Let a be a free primely generated element in a refinement monoid M. Let $a \equiv n_1p_1 + n_2p_2 + \ldots + n_Kp_K + q_1 + q_2 + \ldots + q_L$, q, N_a and ν be as above. Then for all $x \in \{\asymp a\}$,*

1. $N_a(x) \in \mathbb{N}^K$
2. $x \equiv \nu(N_a(x)) + q$

Proof.

1. Since x is in the Archimedean component containing a, $N_a(x)$ is in the Archimedean component containing $N_a(a) = (n_1, n_2, \ldots, n_K)$, which is easily seen to be $\mathbb{N}^K$.
2. Let $n \in \mathbb{N}$ be such that $a \leq nx$, then $nq \leq q \leq a \leq nx$, so, from 11.18, we get $q \leq x$. Thus $x = y + q$ for some $y \in M$. Since $N_a(q) = 0$, we have $N_a(y) = N_a(x)$. From 11.30.3, $y = \nu(N_a(x)) + u$ for some $u \ll a$, and so $x = \nu(N_a(x)) + q + u$.

   From 1, we have $N_a(x) \geq (1, 1, \ldots, 1)$, so $\nu(N_a(x)) + q \geq p_1 + p_2 + \ldots + p_K + q$. This implies that $a \prec \nu(N_a(x)) + q \prec a$. From 11.13.6 we get $u \ll \nu(N_a(x)) + q$, and so $x = \nu(N_a(x)) + q + u \equiv \nu(N_a(x)) + q$

□

It is worth recording the special case of these last two propositions when a is itself a prime:

Proposition 11.32. *Let a be a free prime element in a refinement monoid M. Then for all $x \prec a$,*

1. $n_a(x) < \infty$
2. $n_a(x) = 0 \iff x \ll a$
3. $(\forall n \in \mathbb{N})\ (n_a(x) = n \in \mathbb{N} \iff x \equiv na)$

Proof. Since $x \prec a$, there is some $n \in \mathbb{N}$ such that $x \leq na$. Using 11.26, this implies that $n_a(x) \leq n_a(na) = n < \infty$.

Since a is prime, N_a and n_a coincide, and so 2 is immediate from 11.30. If $n_a(x) = n \in \mathbb{N}$, then $na \leq x$. Since a is prime, this implies $a \leq x$ and so $x \asymp a$. Thus, from 11.31.2, $x \equiv \nu(N_a(x)) = na$. The converse is immediate since n_a is a homomorphism. □

Proposition 11.33. *Let a be a primely generated element in a refinement monoid M. Then all elements of $\{\asymp a\}$ are primely generated.*

Proof. Let $x \asymp a$. By 11.13.5, we consider two cases:

- If a is regular, then from 11.29, $x \equiv a$, and so by 11.11, x is primely generated.

- If a is free, then from 11.31.2, $x \equiv \nu(N_a(x)) + q$. Since $\nu(N_a(x)) + q$ is primely generated, the claim follows from 11.11.

$\square$

Monoids with the property that $nx \leq a$ for all $n \in \mathbb{N}$ implies $x \ll a$, are called **Archimedean** by F. Wehrung [31], [32]. Example 14.2 is a cancellative refinement monoid which does not have this property, so it is quite independent of any of the other cancellation properties we have been studying.

Proposition 11.34. *Let a be a primely generated element in a refinement monoid M. Then for all $x \in M$,*

$$(x \ll a) \iff (\forall n \in \mathbb{N})\ (nx \leq a).$$

Proof. If $x \ll a$, then it true, in general, that $nx \leq a$ for any $n \in \mathbb{N}$.

For the converse, by 11.13.5, we consider two cases:

- Suppose a is regular. Then $x \leq a$ suffices to imply that $x \ll a$.
- Suppose a is free. Then $x \leq a$ implies $nx \prec a$, and we can apply the homomorphism N_a to the inequality $nx \leq a$ to get $nN_a(x) \leq N_a(a) \in (\mathbb{Z}^+)^K$. Since this is true for all $n \in \mathbb{N}$ we get $N_a(x) = 0$, and so, by 11.30.2, $x \ll a$.

$\square$

Before going further with the structure of Archimedean components, we need to establish analogs of 10.7 and 10.11 that apply to order ideals generated by primely generated elements.

Proposition 11.35. *Let r be a primely generated element of a refinement monoid M.*

1. *The Archimedean component $\{\asymp r\}$ is embedded in H_r by the monoid homomorphism $a \mapsto [a]_r$.*
2. *H_r is cancellative.*

Proof. The proofs are essentially the same as for 10.7:

1. We show that the homomorphism is injective when restricted to $\{\asymp r\}$...
   Suppose $a, b \in \{\asymp r\}$ such that $[a]_r = [b]_r$. Then $a + r = b + r$, so we can use the fact the Archimedean component containing r is cancellative (11.13.2) to get $a = b$.
2. Suppose $a, b, c \in \{\prec r\}$ such that $[a]_r + [c]_r = [b]_r + [c]_r$. Then $a + c + r = b + c + r$. Since $c \prec r$, there is some $n \in \mathbb{N}$ such that $c \leq nr$, and so $a + r + nr = b + r + nr$. Canceling using 11.12.2, we get $a + r = b + r$, that is, $[a]_r = [b]_r$.

$\square$

Proposition 11.36. *Let r be a primely generated element of a refinement monoid M such that $M = \{\prec r\}$. Then for all $a, b \in M$*

$$a \sim_r b \iff \langle a \rangle_M = \langle b \rangle_M.$$

Thus the monoid homomorphism $\psi_r \colon H_r \to G(M)$ given by $[a]_r \mapsto \langle a \rangle_M$ is injective. In particular, G_r and H_r embed in $G(M)$. Further $G(H_r) \cong G(M)$.

Proof. The proof of this proposition is the same as for 10.11 except for the obvious replacement of the separativity hypothesis by 11.12.2 $\square$

Now we can prove the main structure theorem for $\{\asymp a\}$:

Theorem 11.37. *Let a be a free primely generated element in a refinement monoid M, and K the number of free primes in a canonical form for a. Then*

1. $H_a \cong (\mathbb{Z}^+)^K \times G_a$ *(monoid isomorphism)*
2. $\{\asymp a\} \cong \mathbb{N}^K \times G_a$ *(semigroup isomorphism)*
3. $G(\{\prec a\}) \cong \mathbb{Z}^K \times G_a$ *(group isomorphism)*

Proof. We will continue to use the notation established for Proposition 11.30.

1. Let $\mu\colon (\mathbb{Z}^+)^K \times G_a \to H_a$ be the homomorphism defined by
$$\mu(\mathbf{m}, [u]_a) = [\nu(\mathbf{m})]_a + [u]_a,$$
for all $\mathbf{m} \in (\mathbb{Z}^+)^K$ and $u \ll a$.

   To show that μ is injective, we suppose that $\mu(\mathbf{m}, [u]_a) = \mu(\mathbf{m}', [u']_a)$ for some $\mathbf{m}, \mathbf{m}' \in (\mathbb{Z}^+)^K$ and $u, u' \ll a$. Then $[\nu(\mathbf{m}) + u]_a = [\nu(\mathbf{m}') + u']_a$, that is,
$$\nu(\mathbf{m}) + u + a = \nu(\mathbf{m}') + u' + a.$$
Since $N_a(\nu(\mathbf{m})) = \mathbf{m}$, $N_a(\nu(\mathbf{m}')) = \mathbf{m}'$ and $N_a(u) = N_a(u') = 0$, we have from this equation that
$$\mathbf{m} + N_a(a) = \mathbf{m}' + N_a(a).$$
Thus $\mathbf{m} = \mathbf{m}'$, $\nu(\mathbf{m}) = \nu(\mathbf{m}')$ and $\nu(\mathbf{m}) + u + a = \nu(\mathbf{m}) + u' + a$.

   Now $\nu(\mathbf{m}) \prec a$ so there is some $n \in M$ such that $\nu(\mathbf{m}) \leq na$. From the equation $\nu(\mathbf{m}) + u + a = \nu(\mathbf{m}) + u' + a$ we then get $na + u + a = na + u' + a$. Using 11.12.2, we can cancel na from this to get $a + u = a + u'$, that is $[u]_a = [u']_a$. We have shown therefore that μ is injective.

   To show that μ is surjective, we suppose that $[x]_a \in H_a$ with $x \prec a$. From 11.30.3, $x = \nu(N_a(x)) + u$ for some $u \ll a$. Hence $[u]_a \in G_a$ and
$$[x]_a = [\nu(N_a(x))]_a + [u]_a = \mu(N_a(x), [u]_a).$$
2. We have from 11.35.1 that $\{\asymp a\}$ is embedded in H_a, so by 1, it is embedded in $(\mathbb{Z}^+)^K \times G_a$. Using 11.31, it is then easy to show that $\{\asymp a\} \cong \mathbb{N}^K \times G_a$ via this embedding.
3. From 11.36,
$$G(\{\prec a\}) \cong G(H_a) \cong G((\mathbb{Z}^+)^K \times G_a) \cong \mathbb{Z}^K \times G_a.$$

$\square$

12 Artinian Decomposition Monoids

We have seen in 8.4 that the relation $a_0 + c_0 = b_0 + c_0$ in a refinement monoid gives rise to a descending sequence of such relationships, $a_n + c_n = b_n + c_n$, for $a_0 \geq a_1 \geq a_2 \geq \ldots$, $b_0 \geq b_1 \geq b_2 \geq \ldots$ and $c_0 \geq c_1 \geq c_2 \geq \ldots$. It is then natural to expect that a refinement monoid with a descending chain condition would have some strong cancellation properties.

In Section 11, we saw that certain cancellation properties occur in refinement monoids containing primely generated elements.

The goal of this section is to show that these two sources of cancellation are closely related: Every Artinian decomposition monoid is primely generated, and conversely, every primely generated decomposition monoid such that the primes form an Artinian class, is Artinian (12.13).

We begin by applying the definitions from Section 2 to monoids: Every monoid is a preordered class when given its minimum order, so lower classes, exact maps and strictly increasing maps are defined and we get

Proposition 12.1. *Let $\sigma\colon M \to N$ be a monoid homomorphism, and $I \subseteq M$ a subclass.*

1. *I is an order ideal if and only if I is both a submonoid and a lower class of M.*
2. *An exact homomorphism maps order ideals to order ideals.*
3. *M has decomposition if and only if the addition homomorphism $+\colon M \times M \to M$ is exact.*
4. *If M has decomposition and I is an order ideal, then the quotient homomorphism $\sigma_I\colon M \to M/I$ is exact.*
5. *The homomorphism σ is strictly increasing if and only if*

$$(\forall a, b \in M)\ (\sigma(a) \ll \sigma(b) \implies a \ll b).$$

Proof.

1. This is just a restatement of 6.11.2 and the definition of order ideals.
2. Use 1 and the facts that a homomorphism maps submonoids to submonoids, and an exact function maps lower classes to lower classes.
3. We temporarily write $b_1 + b_2 = +(b_1, b_2)$ so that $+$ appears as a function from $M \times M$ to M.

   Suppose M has decomposition. If $a \leq +(b_1, b_2) = b_1 + b_2$ in M, then there are a_1, a_2 such that $a = +(a_1, a_2) = a_1 + a_2$, $a_1 \leq b_1$ and $a_2 \leq b_2$. In $M \times M$ this implies $(a_1, a_2) \leq (b_1, b_2)$. Thus $+\colon M \times M \to M$ is exact. The converse is just as easy.
4. We need to show that $\{\leq \sigma_I(b)\} \subseteq \sigma_I(\{\leq b\})$ for all $b \in M$.

   Suppose then that $[a]_I \leq \sigma_I(b) = [b]_I$. Then there is some $u \in I$ such that $a \leq b + u$. By decomposition, there are a', u' with $a' \leq b$, $u' \leq u$ and $a = a' + u'$. Since I is an order ideal, $u' \in I$ and $[a]_I = [a']_I \in \sigma_I(\{\leq b\})$.

5. Suppose σ is strictly increasing and there are $a, b \in M$ such that $\sigma(a) \ll \sigma(b)$. Then $\sigma(a+b) = \sigma(a)+\sigma(b) \leq \sigma(b)$ and $a+b \geq b$. Since σ is strictly increasing, this implies $a + b \leq b$, that is, $a \ll b$.

   For the converse, we know already that σ is increasing, so it remains to show that
$$(a \leq b \text{ and } \sigma(a) \geq \sigma(b)) \implies a \geq b$$
   for all $a, b \in M$.

   If $a \leq b$ then $b = a + c$ for some $c \in M$. Thus $\sigma(a) + \sigma(c) = \sigma(b) \leq \sigma(a)$, that is, $\sigma(c) \ll \sigma(a)$. By hypothesis, this implies $c \ll a$, and so, $a \geq a + c = b$.

□

We will also need the following monoid version of 2.12:

Proposition 12.2. *Consider the following commutative diagram of monoids and monoid homomorphisms:*

$$\begin{tikzcd} K \arrow[r, "\sigma"] \arrow[dr, "\psi"'] & L \arrow[d, "\tau"] \\ & M \end{tikzcd}$$

1. *If σ is surjective and ψ is exact, then τ is exact.*
2. *If τ is injective and ψ is exact, then σ is exact.*

Proof. This follows from 2.12 and the fact that any monoid homomorphism is an increasing function on the underlying preordered class. □

As a simple example of the use of these propositions, consider the following situation: Suppose $A \leq B$ are order ideals in a decomposition monoid M. Let σ and ψ be quotient homomorphisms as in the diagram:

$$\begin{tikzcd} M \arrow[r, "\sigma"] \arrow[dr, "\psi"'] & M/A \arrow[d, "\tau"] \\ & M/B \end{tikzcd}$$

Since $A \leq B$, there is an induced homomorphism τ from M/A to M/B making the diagram commute. The homomorphism σ is surjective and, from 12.1.4, ψ is exact. So using 12.2.1, we get that τ is exact.

We now introduce the the main subject of this section:

Definition 12.3. *A monoid M is an* **Artinian monoid** *if M is an Artinian preordered class. The* **Artinian radical** *of a monoid M is* $\operatorname{Arad} M$ *as defined for preordered classes. See 2.14 and 2.16.*

Thus, by definition, M is Artinian if and only if $\overline{M}$ is Artinian. For example, if M_1 and M_2 are monoids, then $M_1 \times M_2$ is Artinian if and only if M_1 and M_2 are Artinian.

Continuing the pattern established in Section 2 for preordered classes, if X is a subclass of a monoid M, then a **minimal element** of X is an element $a \in X$ such that for any

$b \in X$, $b \le a$ implies $a \le b$. Hence, a monoid M is Artinian if and only if either of the following are true:

1. Every nonempty subclass of M has a minimal element.
2. For every decreasing sequence $x_1 \ge x_2 \ge x_3 \ge \dots$ there is an $N \in \mathbb{N}$ such that $x_n \ge x_N$ for all $n \ge N$.

If I is a submonoid of M then, as we have seen before, the order of I as a subclass of M may be different than its minimum preorder as an independent monoid. Thus a submonoid of an Artinian monoid may not be Artinian with its own minimum preorder. For an example of this see 14.2. Conversely, a submonoid I which is Artinian with its minimum preorder may not be an Artinian subclass of M.

Example 12.4. *The minimum order on $M = (\mathbb{R}^+, +)$ is the same as the usual order on real numbers, so $\mathbb{R}^+$ is not Artinian. Let $I = \{0\} \cup [1, \infty) \subseteq \mathbb{R}^+$. Then I is a submonoid which is not an Artinian subclass of $\mathbb{R}^+$. Nonetheless, with its own minimum order I is an Artinian monoid.*

The complications described above do not occur for order ideals, and so we can use the expression "Artinian order ideal" without ambiguity.

Note that $\{\le 0\}$ is an Artinian lower class of M, so for any monoid, $\{\le 0\} \subseteq \operatorname{Arad} M$. Also, any order ideal of M is a lower class, so every Artinian order ideal is contained in $\operatorname{Arad} M$. On the other hand, $\operatorname{Arad} M$ may not be a submonoid or order ideal of M. Of course, M is Artinian if and only if $M = \operatorname{Arad} M$.

From 2.17 and 2.18 we get

Proposition 12.5. *Let $\sigma\colon M \to N$ be a monoid homomorphism, and $I \subseteq M$, $J \subseteq N$.*

1. *If σ is strictly increasing and J is an Artinian order ideal, then $\sigma^{-1}(J)$ is an Artinian order ideal in M, so $\sigma^{-1}(J) \subseteq \operatorname{Arad} M$.*
2. *If σ is exact and I is an Artinian order ideal, then $\sigma(I)$ is an Artinian order ideal in N, so $\sigma(I) \subseteq \operatorname{Arad} N$.*
3. *If σ is strictly increasing then $\sigma^{-1}(\operatorname{Arad} N) \subseteq \operatorname{Arad} M$.*
4. *If σ is exact then $\sigma(\operatorname{Arad} M) \subseteq \operatorname{Arad} N$.*
5. *If σ is exact, strictly increasing and injective, then $\operatorname{Arad} M = \psi^{-1}(\operatorname{Arad} N)$.*
6. *If σ is exact, strictly increasing and surjective, then $\operatorname{Arad} N = \psi(\operatorname{Arad} M)$.*
7. *If I is an order ideal in M, then $\operatorname{Arad} I = \operatorname{Arad} M \cap I$.*

Proof. These results follow directly from the above-mentioned propositions, 12.1.2, 6.13.3 and the fact that any monoid homomorphism is an increasing function. □

As we have already seen, if M is a decomposition monoid then the addition homomorphism $+\colon M \times M \to M$, and division by an order ideal are exact homomorphisms, so applying the previous proposition we get

Proposition 12.6. *Let $A, B \leq M$ be two order ideals of a decomposition monoid and $\sigma_A \colon M \to M/A$ the quotient homomorphism.*

1. $\operatorname{Arad} M$ *is an order ideal.*
2. $\operatorname{Arad}(A + B) = \operatorname{Arad} A + \operatorname{Arad} B$
3. $\operatorname{Arad}(A \cap B) = \operatorname{Arad} A \cap \operatorname{Arad} B$
4. *A and B are Artinian if and only if $A + B$ is Artinian.*
5. *For all $a \in M$, $\{\leq a\}$ is Artinian if and only if $\{\prec a\}$ is Artinian.*
6. *$\sigma_A(\operatorname{Arad} B) \subseteq \operatorname{Arad}(M/A)$. In particular, if B is Artinian then $\sigma_A(B) = (A+B)/A$ is an Artinian order ideal in M/A, and if M is Artinian then so is M/A.*
7. *If $A = \{\leq 0\}$, then $\sigma_A(\operatorname{Arad} M) = \operatorname{Arad}(M/A)$ and $\sigma_A^{-1}(\operatorname{Arad}(M/A)) = \operatorname{Arad} M$. In particular, M is Artinian if and only if M/A is Artinian.*

Proof.

1. By definition, $\operatorname{Arad} M$ is a lower class, so we need to show that it is a submonoid...
   We have already noted that $0 \in \operatorname{Arad} M$ so it remains to show that $\operatorname{Arad} M$ is closed under addition: If $x_1, x_2 \in \operatorname{Arad} M$ then the lower classes $\{\leq x_1\}$ and $\{\leq x_2\}$ in M are Artinian, as is $\{\leq x_1\} \times \{\leq x_2\} \subseteq M \times M$. The addition map $+\colon \{\leq x_1\} \times \{\leq x_2\} \to \{\leq (x_1 + x_2)\}$ is exact and surjective, so from 12.5.2, the lower class $\{\leq (x_1 + x_2)\}$ is Artinian. That is, $x_1 + x_2 \in \operatorname{Arad} M$.
2. $\operatorname{Arad}(A+B) = (A+B) \cap \operatorname{Arad} M = (A \cap \operatorname{Arad} M) + (B \cap \operatorname{Arad} M) = \operatorname{Arad} A + \operatorname{Arad} B$. Here we have used the distributivity of $\mathcal{L}(M)$ (7.10).
3. Easy using 12.5.7, and, in fact, does not require that M have decomposition.
4. This follows easily from 2.
5. If $\{\leq a\}$ is Artinian, then $\{\leq a\} \subseteq \operatorname{Arad} M$. But $\operatorname{Arad} M$ is an order ideal, so it also contains the order ideal generated by a. Thus $\{\prec a\}$ is Artinian.
6. Since σ_A is exact, this follows from 12.5.2 .
7. This follows easily from the fact that $\overline{M} \cong \overline{M/A}$.

□

Clearly any finite monoid is Artinian, as is any monoid M such that $\overline{M}$ is finite. Any Abelian group, for example, is trivially an Artinian monoid. Any free monoid is Artinian since for any element x, $\{\leq x\}$ is a finite set (Reminder: Any element of a free monoid is a <u>finite</u> sum of the generators. See 5.13).

A finitely generated free monoid is isomorphic to $(\mathbb{Z}^+)^n$ for some $n \in \mathbb{N}$, and the minimal order of this monoid coincides with the ordering of this same set as a direct product of posets. Thus we get the following important result:

Proposition 12.7. *Any finitely generated submonoid of a monoid is an Artinian subclass.*

Proof. Let I be a finitely generated submonoid of M. Then I is the image of the monoid $(\mathbb{Z}^+)^n$ for some $n \in \mathbb{N}$ under a monoid homomorphism σ. Since σ is increasing, we can apply 2.26, to get that I is an Artinian subclass of M. □

This proposition has the consequence that any finitely generated monoid is Artinian. But this fact alone does not suffice to prove the proposition, since, as mentioned before, a submonoid may be Artinian when given its minimum preorder, but nonetheless fail to be an Artinian subclass of the monoid that it is embedded in.

In the remainder of this section we investigate the relationship between the Artinian condition and the existence of primes, indecomposables and regular elements in monoids.

Proposition 12.8. *Every element in an Artinian partially ordered monoid is a sum of indecomposables.*

Proof. Let M be an Artinian monoid and let $\mathcal{B}$ be the subclass of elements of M which are not sums of indecomposables. If $\mathcal{B}$ is not empty it has a minimal element p. We will show that p is indecomposable, contradicting $p \in \mathcal{B}$.

Suppose $p = a + b$ for some $a, b \in M$, then either a or b must be in $\mathcal{B}$ otherwise p would be a sum of indecomposables. But the minimality of p in $\mathcal{B}$ then implies that either $a = p$ or $b = p$, thus p is indecomposable. □

In an Artinian monoid M that is not partially ordered it is possible that elements are not sums of indecomposables, even though every element of $\overline{M}$ is a sum of indecomposables of $\overline{M}$. The complication is that two elements $a, b \in M$ may satisfy $a \equiv b$ with a, but not b, a sum of indecomposables. This happens in the following example:

Example 12.9. *Let M be the monoid generated by two elements a and b such that $3a = a$ and $2a = 2b$. M has six elements, $\{0, a, 2a, b, a + b, 2a + b\}$, and these form the three $\equiv$-equivalence classes: $\{0\}, \{b\}, \{a, 2a, a + b, 2a + b\}$. Thus $\overline{M} \cong \{0, 1, \infty\}$ with $1 + 1 = \infty$. Since 0 and 1 are all the indecomposables in $\{0, 1, \infty\}$, the elements 0 and b are all the indecomposables of M. Both M and $\{0, 1, \infty\}$ are, of course, Artinian. Every element of $\{0, 1, \infty\}$ is a sum of indecomposables, but only 0, b, $2a$, and $2a + b$ are sums of indecomposables in M.*

Note that the elements of M which are not sums of indecomposables, namely, a and $a+b$, are regular.

Proposition 12.10. *Every element in an Artinian monoid is a sum of indecomposables and regular elements.*

Proof. Let M be an Artinian monoid and let $\mathcal{B}$ be the subclass of elements of M which are not sums of indecomposables and regular elements. If $\mathcal{B}$ is not empty it has a minimal element p. We will show that p is indecomposable or regular, contradicting $p \in \mathcal{B}$.

First we note that if p is not regular, then any $c \ll p$ is not in $\mathcal{B}$, since otherwise the minimality of p would give $p \leq c$ and then $2p \leq p + c \leq p$.

Suppose then that p is not regular and $p \equiv a + b$ for some $a, b \in M$, then $p = a + b + c$ for some $c \ll p$. Since c is not in $\mathcal{B}$, either a or b must be in $\mathcal{B}$, otherwise p would be a sum of indecomposables and regular elements. But the minimality of p in $\mathcal{B}$ then implies that either $p \leq a$ or $p \leq b$, thus p is indecomposable. □

If the monoid M has decomposition we get a much stronger result:

Proposition 12.11. *Any Artinian decomposition monoid is primely generated.*

Proof. Let M be an Artinian decomposition monoid and let $\mathcal{B}$ be the subclass of elements of M which are not sums of primes. If $\mathcal{B}$ is not empty it has a minimal element p. We will show that p is prime, contradicting $p \in \mathcal{B}$.

Suppose $p \leq a_1 + a_2$ for some $a_1, a_2 \in M$, then there are $p_1, p_2 \in M$ such that $p_1 \leq a_1$, $p_2 \leq a_2$ and $p = p_1 + p_2$. Either p_1 or p_2 must be in $\mathcal{B}$ otherwise p would be a sum of primes. But the minimality of p in $\mathcal{B}$ then implies that either $p \leq p_1 \leq a_1$ or $p \leq p_2 \leq a_2$, thus p is prime. $\square$

Next we consider a condition on the primes of a monoid that ensures that the monoid is Artinian:

Proposition 12.12. *Let M be a primely generated monoid such that the prime elements of M form an Artinian subclass of M. Then M is an Artinian monoid.*

Proof. Note that if M satisfies the hypothesis, then the same is true of $\overline{M}$, and also, by definition, M is Artinian if and only if $\overline{M}$ is Artinian. Thus without loss of generality, we can assume that M is partially ordered.

Let $\mathcal{L}$ be the class of prime elements of M with order induced from M, then by hypothesis, $\mathcal{L}$ is an Artinian poclass.

Let $a_1 \geq a_2 \geq \ldots$ be a decreasing sequence in M. For each a_n let $\mathcal{A}_n \subseteq \mathcal{L}$ be the primes which appear in a minimal prime expression for a_n. By 11.20, $\mathcal{A}_n$ is a finite antichain in $\mathcal{L}$. Since $a_n \geq a_{n+1}$, every prime in $\mathcal{A}_{n+1}$ is less than or equal to some prime in $\mathcal{A}_n$. Thus this sequence of antichains satisfies the hypothesis of 2.28, and by that proposition, the set $\mathcal{A} = \bigcup_{n \in \mathbb{N}} \mathcal{A}_n$ is finite.

The sequence $a_1 \geq a_2 \geq \ldots$ is then contained in the submonoid generated by $\mathcal{A}$. By 12.7, this submonoid is an Artinian subclass of M. Thus there is some $N \in \mathbb{N}$ such that $a_n \geq a_N$ for all $n \geq N$. $\square$

Finally, putting the last two propositions together we get:

Theorem 12.13. *A decomposition monoid M is Artinian if and only if*

- *M is primely generated, and*
- *The prime elements of M form an Artinian subclass of M.*

Proof. Combine the previous proposition, 12.12 and the fact that, if M is Artinian, then so is the subclass of prime elements of M. $\square$

One might wonder at this point whether an Artinian monoid whose elements are all primely generated is necessarily a decomposition monoid. This is not true in general, and one counterexample is the monoid M of Example 11.21 which is generated by two elements p and q subject to the relations $3p = 2p$, $3q = 2q$ and $2p + q = 2q + p$. p and q are primes, so every element in the monoid is primely generated, but the inequality $(2p) \leq (2q) + (p)$ can not be decomposed, that is, $2p$ can not be written as $2p = p_1 + p_2$ with $p_1 \leq 2q$ and $p_2 \leq p$. So M is not a decomposition monoid.

13 Artinian Refinement Monoids

In this section we will consider Artinian refinement monoids and their properties. From Theorem 12.13, Artinian refinement monoids are primely generated, so the results of Section 11 apply to such monoids:

Proposition 13.1. *Let M be an Artinian refinement monoid. Then*

1. *$\overline{M}$ is an Artinian refinement monoid.*
2. *M and $\overline{M}$ are weakly cancellative, midseparative and separative.*
3. *M has no proper regular elements $\iff$ $\overline{M}$ has no proper regular elements $\iff$ M is strongly separative $\iff$ $\overline{M}$ is strongly separative.*
4. *M and $\overline{M}$ have $\leq$-multiplicative cancellation.*

Proof. Since Artinian refinement monoids are primely generated, these results follow from 11.9, 11.15 and 11.19. □

One interesting consequence of weak cancellation for Artinian monoids is

Proposition 13.2. *Let M be a partially ordered Artinian refinement monoid. Then M is a join-semilattice (6.19).*

Proof. We need to show that $a \vee b$ exists for any $a, b \in M \ldots$

Suppose $a, b \in M$ and set

$$\mathcal{A} = \{c \in M \mid a \leq c \text{ and } b \leq c\}.$$

The element $a+b$ is in $\mathcal{A}$, so this subclass is not empty and must have some minimal element. Suppose c_1 and c_2 are two minimal elements of $\mathcal{A}$.

The monoid M is weakly cancellative, and we have $a, b \leq c_1, c_2$, so from the Riesz interpolation property (9.15), there is some $c \in M$ such that $a, b \leq c \leq c_1, c_2$. c is in $\mathcal{A}$, so the minimality of c_1 and c_2 then implies $c_1 = c = c_2$.

Thus $\mathcal{A}$ has a unique minimum which is then the supremum of a and b. □

Unlike the situation for $\widetilde{M}$ (6.21), in this proposition, the operations $+$ and $\vee$ do not coincide.

Combining this result with 13.1, we have that for any Artinian refinement monoid M, $\overline{M}$ is a semilattice.

Even though we know already that Artinian refinement monoids are weakly cancellative, it is worthwhile giving a direct proof since it serves as a prototype for the more complicated proofs to come:

Proposition 13.3. *Let a, b, c be elements of a refinement monoid such that $a + c = b + c$. If $\{\prec c\}$ is Artinian, then there exists a refinement matrix*

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} * & * \\ * & c_1 \end{pmatrix} \end{array}$$

with $c \leq c_1$.

Proof. Define

$$\mathcal{C} = \{c' \leq c \mid a + c' = b + c'\}.$$

Since $\mathcal{C} \subseteq \{\prec c\}$ and $c \in \mathcal{C}$, the subclass $\mathcal{C}$ is Artinian and nonempty. Let c_0 be a minimal element of $\mathcal{C}$. Then $a + c_0 = b + c_0$ and we can make a refinement of this equation:

$$\begin{array}{cc} & \begin{array}{cc} b & c_0 \end{array} \\ \begin{array}{c} a \\ c_0 \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1' \end{pmatrix} \end{array}$$

Now $c_1' \leq c_0 \leq c$ and $a + c_1' = d_1 + c_0 = b + c_1'$, so $c_1' \in \mathcal{C}$. By the minimality of c_0 in $\mathcal{C}$, we get $c_0 \leq c_1'$. Since $c_0 \leq c$ there is some x_0 such that $c = c_0 + x_0$. Set $c_1 = c_1' + x_0$. Then $c \leq c_1$ and we have $c = c_0 + x_0 = c_1' + a_1 + x_0 = c_1 + a_1$ and, similarly, $c = c_1 + b_1$. $\square$

A similar direct proof that Artinian refinement monoids are midseparative could be given, but, in fact, we have a stronger result:

Proposition 13.4. *Let a and b be elements of a refinement monoid M such that $2a = a+b$. If $\{\prec b\}$ is Artinian, then there exists an idempotent e such that $a = b + e$.*

Proof. Define

$$\mathcal{B} = \{b' \in M \mid \exists a', d' \text{ such that } a = d' + a', b = d' + b', 2a' = a' + b'\}.$$

Since $\mathcal{B} \subseteq \{\prec b\}$ and $b \in \mathcal{B}$, the subclass $\mathcal{B}$ is Artinian and nonempty. Let b_0 be a minimal element of $\mathcal{B}$ and a_0, d_0 such that $a = d_0 + a_0$, $b = d_0 + b_0$, and $2a_0 = a_0 + b_0$.

From 8.6 (with $n = 2,\ c_0 = a_0$), there is a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b_0 & a_1 \end{array} \\ \begin{array}{c} a_0 \\ a_1 \end{array} & \begin{pmatrix} d_2 & a_2 \\ b_2 & c_2 \end{pmatrix} \end{array}$$

with $c_2 \leq a_2$ and $2b_2 \leq a_0$. We also have $a = (d_0 + d_2) + a_2$, $b = (d_0 + d_2) + b_2$, and, by 8.1.2, $2a_2 = a_2 + b_2$, so $b_2 \in \mathcal{B}$. Since $b_2 \leq b_0$, the minimality of b_0 implies $b_0 \leq b_2$. In particular, $2b_0 \leq 2b_2 \leq a_0$.

Since $2b_0 \leq a_0$, there is some x_0 such that $a_0 = 2b_0 + x_0$. From $2a_0 = a_0 + b_0$ we then get $4b_0 + 2x_0 = 3b_0 + x_0$.

By 13.1, $\{\prec b\}$ is separative. Since $b_0 \in \{\prec b\}$, we can use 8.14, to cancel $2b_0$ from the above equation to get $2b_0 + 2x_0 = b_0 + x_0$. Set $e = b_0 + x_0$. Then e is an idempotent such that $b + e = (d_0 + b_0) + (b_0 + x_0) = d_0 + a_0 = a$. $\square$

Note that we have actually shown that $0 \in \mathcal{B}$, since if we set $d' = b$, $a' = e$ and $b' = 0$, then $a = d' + a'$, $b = d' + b'$ and $2a' = a' + b'$.

The significance of this proposition is that it is not true with the weaker hypothesis that $\{\prec b\}$ is midseparative (instead of Artinian). A counterexample is the monoid of Example

9.7. Putting $a = (\infty, 1)$ and $b = (0, 1)$ in this monoid, we have $2a = a + b$. Notice that $\{\prec b\} \cong \mathbb{R}^+$ so is cancellative and midseparative, but since the only idempotent is 0, there is no idempotent e such that $a = b + e$. Thus 13.4 implies, but is not implied by, the fact that Artinian refinement monoids have midseparativity.

This situation is to be contrasted with proposition 13.3, which remains true (by 9.10.1) if the hypothesis that $\{\prec c\}$ is Artinian is replaced by $\{\prec c\}$ is weakly cancellative.

This difference will allow us to prove an extension theorem (13.6) for midseparativity, which is not true for weak cancellation, and will eventually be used to show in Section 15 that semi-Artinian refinement monoids are midseparative but not weakly cancellative.

To prove the extension theorem we make use of the following lemma:

Lemma 13.5. *Let I be an Artinian order ideal in a refinement monoid M.*

1. *If $a, b, c \in M$ with $a + c = b + c$, $c \leq a$ and $b \in I$, then there is an idempotent $e \leq c$ such that $a = b + e$.*
2. *If $2[e]_I = [e]_I$ for some $e \in M$, then there is an idempotent $e' \leq e$ such that $[e']_I = [e]_I$.*
3. *If $2[e]_I = [e]_I \leq [a]_I$ for some $e, a \in M$, then there is an idempotent $e' \leq a$ such that $[e']_I = [e]_I$.*

Proof.

1. If $a + c = b + c$ and $c \leq a$ then, by 8.5.1, there is a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

   with $c_1 \leq a_1$. From 8.1.2 we get $2a_1 = a_1 + b_1$. We have $b_1 \leq b \in I$ so $\{\prec b_1\} \subseteq I$ and $\{\prec b_1\}$ is Artinian. Using 13.4, there is some idempotent $e \in M$ such that $a_1 = b_1 + e$. Thus $e \leq a_1 \leq c$ and $a = d_1 + a_1 = d_1 + b_1 + e = b + e$.
2. Since $2[e]_I = [e]_I$, there are elements $u, b \in I$ such that $2e + u = e + b$. Hence $(e + u) + e = b + e$ with $e \leq e + u$ and $\{\prec b\}$ Artinian. From 1, there is some idempotent $e' \leq e$ such that $e + u = b + e'$. Since $u, b \in I$, we have $[e']_I = [e]_I$.
3. From $[e]_I \leq [a]_I$, there is some $u \in I$ such that $e \leq a + u$. Decomposing this we can write $e = e'' + u''$ with $e'' \leq a$ and $u'' \leq u$. Since $u'' \in I$, we have $[e'']_I = [e]_I$ and so, $2[e'']_I = [e'']_I$. Using 2, there is some idempotent e' such that $e' \leq e'' \leq a$ and $[e']_I = [e'']_I = [e]_I$.

□

Proposition 13.6. *Let I be an Artinian order ideal in a refinement monoid M. Then if M/I is midseparative, so is M.*

The proof of this proposition is a variation of the proof of 8.16:

Proof. Suppose there are $a, b \in M$ such that $2a = a + b$. From 8.5.1, there is a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b & a \end{array} \\ \begin{array}{c} a \\ a \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

with $c_1 \leq a_1$. In M/I we get $2[a_1]_I = [a_1]_I + [b_1]_I$, so using the midseparativity of M/I, there is some idempotent $[e_1]_I$ such that $[a_1]_I = [b_1]_I + [e_1]_I$. We have $[e_1]_I \leq [a_1]_I$, so using 13.5.3, we can assume that $2e_1 = e_1 \leq a$. In particular, $a + e_1 = a$.

Since $[a_1]_I = [b_1]_I + [e_1]_I$, there are $u_1, v_1 \in I$ such that $a_1 + u_1 = b_1 + e_1 + v_1$. From a refinement of this relationship,

$$\begin{array}{c|cc} & b_1 + e_1 & v_1 \\ \hline a_1 & d_2 & v_2 \\ u_1 & u_2 & e_2 \end{array}$$

we get the equation $a_1 + u_2 = b_1 + e_1 + v_2$, and hence,

$$\begin{aligned} a + v_2 &= a + e_1 + v_2 = c_1 + b_1 + e_1 + v_2 = c_1 + a_1 + u_2 \\ &= a + u_2 = d_1 + a_1 + u_2 = d_1 + b_1 + e_1 + v_2 \\ &= (b + e_1) + v_2. \end{aligned}$$

Note that $v_2 \leq v_1$ so $v_2 \in I$, and also $v_2 \leq a_1 \leq a$. Since $\{\prec v_2\}$ is contained in I, it is Artinian, and, from 13.1, midseparative. Using 9.10.2, there exists some idempotent $e_2 \leq v_2$ such that $a = (b + e_1) + e_2$. Set $e = e_1 + e_2$, then $2e = e$ and $a = b + e$. □

Artinian refinement monoids are primely generated so, by 11.34, they have the property that $na \leq b$ for all $n \in \mathbb{N}$ implies $a \ll b$. We give next a direct proof of this fact:

Proposition 13.7. *Let M be an Artinian refinement monoid. Then*

$$(\forall a, b \in M)\ (a \ll b \iff (\forall n \in \mathbb{N})\ (na \leq b)).$$

Proof. The implication $\Rightarrow$ is easy to prove and true in any monoid. We prove here the converse…

Since $na \leq b$ for all $n \in \mathbb{N}$, there are $c_1, c_2, \ldots$ such that $na + c_n = b$. Let $N \in \mathbb{N}$ be such that $a + c_N$ is minimal in the set $\{a + c_n \mid n \in \mathbb{N}\}$.

Now $Na + c_N = b = (N+1)a + c_{N+1}$, so, since M is separative, $a + c_N = 2a + c_{N+1}$. In particular, $a + c_N \geq a + c_{N+1}$. The minimality of $a + c_N$, then implies $a + c_N \leq a + c_{N+1}$, and so $a + b = a + Na + c_N \leq a + Na + c_{N+1} = b$, that is, $a \ll b$. □

14 Semi-Artinian Decomposition Monoids

We will see in Section 19 that the monoid $M(R\text{-}\mathbf{Noeth})$, for a commutative Noetherian (or FBN) ring R, is an Artinian refinement monoid. So for these rings, $M(R\text{-}\mathbf{Noeth})$ has all the cancellation properties described in the previous two sections. For other rings, however, $M(R\text{-}\mathbf{Noeth})$ may not be Artinian. See Example 17.1.

Nonetheless, for any ring, the monoid $M(R\text{-}\mathbf{Noeth})$ is semi-Artinian. In the next section we will see that semi-Artinian refinement monoids have midseparativity but not weak cancellation or $\leq$-multiplicative cancellation. In this section we consider the semi-Artinian chain condition in decomposition monoids.

Definition 14.1. *A monoid is* **semi-Artinian** *if there is a well ordered chain of order ideals $I_0 \leq I_1 \leq \cdots \leq I_\alpha \leq \cdots \leq M$ with indices $\alpha \in \mathbf{Ord}$, such that*

1. $I_0 = \{\leq 0\}$
2. $I_{\alpha+1}/I_\alpha$ *is an Artinian monoid for all* $\alpha \in \mathbf{Ord}$.
3. $I_\alpha = \bigcup_{\beta<\alpha} I_\beta$ *for all limit ordinals* $\alpha \in \mathbf{Ord}$.
4. $M = \bigcup_{\alpha\in\mathbf{Ord}} I_\alpha$.

A chain of order ideals as above which satisfies 1, 2 *and* 3, *we will call a* **semi-Artinian series**.

Obviously, any Artinian monoid is semi-Artinian. The following example is probably the simplest semi-Artinian/non-Artinian monoid:

Example 14.2. *A cancellative semi-Artinian refinement monoid which is not Artinian:*

We will construct this example as a submonoid of the cancellative partially-ordered Artinian refinement monoid $\mathbb{Z}^+ \times \mathbb{Z}$...

Let $L = \{(0,n) \mid n \in \mathbb{N}\} = (0,\mathbb{N}) \subseteq \mathbb{Z}^+ \times \mathbb{Z}$ and $M = (\mathbb{Z}^+ \times \mathbb{Z}) \setminus L$. It is easily checked that M is a submonoid of $\mathbb{Z}^+ \times \mathbb{Z}$. M is a subset of a cancellative partially ordered monoid so is itself cancellative and partially ordered.

We prove that M has refinement:

Suppose we have $(w_1,w_2),(x_1,x_2),(y_1,y_2),(z_1,z_2) \in M$ such that $(w_1,w_2)+(x_1,x_2) = (y_1,y_2)+(z_1,z_2)$. Since these elements are in the refinement monoid $\mathbb{Z}^+ \times \mathbb{Z}$, there is a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} (w_1,w_2) & (x_1,x_2) \end{array} \\ \begin{array}{c} (y_1,y_2) \\ (z_1,z_2) \end{array} & \begin{pmatrix} (s_1,s_2) & (t_1,t_2) \\ (u_1,u_2) & (v_1,v_2) \end{pmatrix} \end{array}$$

with entries in $\mathbb{Z}^+ \times \mathbb{Z}$. We will show that this refinement matrix can be modified so that all entries are in M

If all the entries happen to be in M then we are done. So suppose (without loss of generality) that $(s_1,s_2) \in L$, that is, $(s_1,s_2) = (0,n)$ for some $n \in \mathbb{N}$. We can rewrite the *refinement matrix as*

$$\begin{array}{cc} & \begin{array}{cc} (w_1,w_2) & (x_1,x_2) \end{array} \\ \begin{array}{c} (y_1,y_2) \\ (z_1,z_2) \end{array} & \begin{pmatrix} (0,0) & (t_1,t_2+n) \\ (u_1,u_2+n) & (v_1,v_2-n) \end{pmatrix} \end{array}$$

We have $(0,0) \in M$*,* $(t_1,t_2+n) = (y_1,y_2) \in M$ *and* $(u_1,u_2+n) = (w_1,w_2) \in M$ *so the only possible remaining problem is if* $(v_1,v_2-n) \in L$*, that is if* $(v_1,v_2-n) = (0,m)$ *for some* $m \in \mathbb{N}$*. If this happens then we can rewrite the refinement matrix as*

$$\begin{array}{cc} & \begin{array}{cc} (w_1,w_2) & (x_1,x_2) \end{array} \\ \begin{array}{c} (y_1,y_2) \\ (z_1,z_2) \end{array} & \begin{pmatrix} (0,-m) & (t_1,t_2+n+m) \\ (u_1,u_2+n+m) & (0,0) \end{pmatrix} \end{array}$$

Now $(0,0),(0,-m) \in M$*,* $(t_1,t_2+n+m) = (x_1,x_2) \in M$ *and* $(u_1,u_2+n+m) = (z_1,z_2) \in M$ *so we are finished.*

Thus we have shown that M *has refinement.*

M *is not finitely generated. A convenient generating subset is* $\{c, a_0, a_1, a_2, \ldots\}$ *where* $c = (0,-1)$ *and* $a_n = (1,n)$ *for* $n \in \mathbb{Z}^+$*. These generators satisfy the relations* $a_n = a_{n+1} + c$ *for all* $n \in \mathbb{Z}^+$*.*

The order in M *is not the same as in* $\mathbb{Z}^+ \times \mathbb{Z}$*. In particular,* $a_n \leq a_m \iff n \geq m$*, so the sequence* $a_0 \geq a_1 \geq a_2 \geq \ldots$ *is strictly decreasing. Thus* M *is not Artinian and no* a_n *is in* $\operatorname{Arad} M$*. On the other hand, the order ideal* $I = \{\prec c\} = \{(0,-n) \mid n \in \mathbb{Z}^+\}$ *is isomorphic to* $\mathbb{Z}^+$ *so is Artinian. Thus we have shown that* $\operatorname{Arad} M = I$*. In the quotient monoid* M/I*, we have* $[a_n]_I = [a_m]_I$ *for any* $m,n \in \mathbb{Z}^+$ *so* M/I *has one generator and is isomorphic to* $\mathbb{Z}^+$*. In particular,* M/I *is Artinian and hence* M *is semi-Artinian.*

Notice that M *is not Artinian even though it is a submonoid of an Artinian monoid, and that since* 0 *and* c *are the only primes,* M *is not primely generated.*

Notice also that $nc + a_n = a_0$ *for all* $n \in \mathbb{N}$*. So* $nc \leq a_0$ *for all* $n \in \mathbb{N}$ *does not imply that* $c \ll a_0$*, even though* M *has the strongest possible cancellation property. See 13.7 and 11.34.*

The above example happens to be cancellative and so weakly cancellative, but example 15.9 shows that, in general, semi-Artinian refinement monoids do not have to be either.

Recall that in a decomposition monoid, the Artinian radical is the largest Artinian order ideal. This suggests the following definition:

Definition 14.3. *Let* M *be a decomposition monoid. Define inductively an increasing sequence of order ideals,* $\operatorname{Arad}_0 M \leq \operatorname{Arad}_1 M \leq \cdots \leq \operatorname{Arad}_\alpha M \leq \ldots$ *for* $\alpha \in \mathbf{Ord}$*, as follows:*

1. $\operatorname{Arad}_0 M = \{\leq 0\}$
2. $\operatorname{Arad}_{\alpha+1} M = \sigma_\alpha^{-1}(\operatorname{Arad}(M/\operatorname{Arad}_\alpha M))$ *where* σ_α *is the quotient homomorphism from* M *to* $M/\operatorname{Arad}_\alpha M$*.*
3. *If* α *is a limit ordinal, define* $\operatorname{Arad}_\alpha M = \bigcup_{\beta<\alpha} \operatorname{Arad}_\beta M$

In addition, we define the **semi-Artinian radical** *of* M *by*

$$\operatorname{srad} M = \bigcup_{\alpha \in \mathbf{Ord}} \operatorname{Arad}_\alpha M.$$

Notice that, by 5.10, we have

$$(\operatorname{Arad}_{\alpha+1} M)/(\operatorname{Arad}_\alpha M) = \operatorname{Arad}(M/\operatorname{Arad}_\alpha M).$$

So $(\operatorname{Arad}_{\alpha+1} M)/(\operatorname{Arad}_\alpha M)$ is Artinian for all α and $\operatorname{Arad}_0 M \leq \operatorname{Arad}_1 M \leq \ldots$ is a semi-Artinian series in M.

From 12.6.7, we get directly that $\operatorname{Arad} M = \operatorname{Arad}_1 M$. A simple induction argument shows that if there is some $\alpha \in \mathbf{Ord}$ such that $\operatorname{Arad}_{\alpha+1} M = \operatorname{Arad}_\alpha M$, then $\operatorname{Arad}_\beta M = \operatorname{Arad}_\alpha M$ for all $\beta \geq \alpha$, so that $\operatorname{srad} M = \operatorname{Arad}_\alpha M$. In particular, if $\operatorname{Arad}_1 M = \operatorname{Arad}_0 M$, that is, if $\operatorname{Arad} M = \{\leq 0\}$, then $\operatorname{srad} M = \{\leq 0\}$.

Proposition 14.4. *Let M be a decomposition monoid, and $I_0 \leq I_1 \leq \cdots \leq M$ a semi-Artinian series. Then $I_\alpha \subseteq \operatorname{Arad}_\alpha M$ for all $\alpha \in \mathbf{Ord}$.*

Proof. Proof by induction on α...

- $\underline{\alpha = 0}$ Trivial.
- $\underline{\alpha \text{ is a successor ordinal}}$ Suppose $\alpha = \beta + 1$ and $I_\beta \subseteq \operatorname{Arad}_\beta M$. Let τ and σ_β be quotient homomorphisms as in the diagram.

$$\begin{tikzcd} M \arrow[r, "\tau"] \arrow[dr, "\sigma_\beta"'] & M/I_\beta \arrow[d, "\rho"] \\ & M/\operatorname{Arad}_\beta M \end{tikzcd}$$

  Then since $I_\beta \subseteq \operatorname{Arad}_\beta M$, there is a homomorphism ρ making the diagram commute. σ_β is exact and τ is surjective, so from 12.2, ρ is exact. $\tau(I_\alpha) = I_\alpha/I_\beta$ is Artinian so, using 12.5.2, $\rho(\tau(I_\alpha)) = \sigma_\beta(I_\alpha) \subseteq \operatorname{Arad}(M/\operatorname{Arad}_\beta M)$. Thus

$$I_\alpha \subseteq \sigma_\beta^{-1}(\sigma_\beta(I_\alpha)) \subseteq \sigma_\beta^{-1}(\operatorname{Arad}(M/\operatorname{Arad}_\beta M)) = \operatorname{Arad}_\alpha M,$$

  and the claim is true for successor ordinals.
- $\underline{\alpha \text{ is a limit ordinal}}$ We have $I_\beta \subseteq \operatorname{Arad}_\beta M$ for all $\beta < \alpha$. Thus

$$I_\alpha = \bigcup_{\beta<\alpha} I_\beta \subseteq \bigcup_{\beta<\alpha} \operatorname{Arad}_\beta M = \operatorname{Arad}_\alpha M,$$

  so the claim is true for limit ordinals.

□

Corollary 14.5. *A decomposition monoid M is semi-Artinian if and only if $M = \operatorname{srad} M$.*

Proof. If $M = \operatorname{srad} M$ then the chain of order ideals $\operatorname{Arad}_0 M \leq \operatorname{Arad}_1 M \leq \ldots$ satisfies the requirements of the definition of semi-Artinian monoids.

Conversely, if $I_0 \leq I_1 \leq \cdots \leq M$ is a chain of order ideals as required by the definition, then from the previous proposition,

$$M = \bigcup_{\alpha\in\mathbf{Ord}} I_\alpha \subseteq \bigcup_{\alpha\in\mathbf{Ord}} \operatorname{Arad}_\alpha M = \operatorname{srad} M \subseteq M.$$

□

Proposition 14.6. *If I is an order ideal in a decomposition monoid M, then for all $\alpha \in \mathbf{Ord}$*

$$\operatorname{Arad}_\alpha I = (\operatorname{Arad}_\alpha M) \cap I$$

and

$$\operatorname{srad} I = (\operatorname{srad} M) \cap I.$$

Proof. We prove the first claim by induction on α...

- $\underline{\alpha = 0}$ Since $\{\leq 0\} \subseteq I$, we get $\operatorname{Arad}_0 I = \{\leq 0\} = \operatorname{Arad}_0 M = (\operatorname{Arad}_0 M) \cap I$.
- $\underline{\alpha \text{ is a successor ordinal}}$ Suppose $\alpha = \beta + 1$ and $\operatorname{Arad}_\beta I = (\operatorname{Arad}_\beta M) \cap I$. Let σ_I and σ_M be quotient homomorphisms and ι the inclusion homomorphism as in the diagram:

$$\begin{tikzcd} I \arrow[r, "\sigma_I"] \arrow[d, "\iota"] & I/\operatorname{Arad}_\beta I \arrow[d, "\psi"] \\ M \arrow[r, "\sigma_M"] & M/\operatorname{Arad}_\beta M \end{tikzcd}$$

Since $\operatorname{Arad}_\beta I \subseteq \operatorname{Arad}_\beta M$, there is a unique homomorphism ψ which makes the diagram commute. Since $\operatorname{Arad}_\beta I = (\operatorname{Arad}_\beta M) \cap I$, from 7.14 we have that ψ induces a monoid isomorphism from $\overline{I/\operatorname{Arad}_\beta I}$ to $\overline{(I + \operatorname{Arad}_\beta M)/\operatorname{Arad}_\beta M}$. In particular,

$$\operatorname{Arad}(I/\operatorname{Arad}_\beta I) = \psi^{-1}(\operatorname{Arad}((I + \operatorname{Arad}_\beta M)/\operatorname{Arad}_\beta M)).$$

Using 12.5.7, 5.10 and 7.11.2 we calculate

$$\begin{aligned} \operatorname{Arad}((I + \operatorname{Arad}_\beta M)/\operatorname{Arad}_\beta M) &= \operatorname{Arad}(M/\operatorname{Arad}_\beta M) \cap (I + \operatorname{Arad}_\beta M)/\operatorname{Arad}_\beta M \\ &= \sigma_M(\operatorname{Arad}_\alpha M) \cap \sigma_M(I) \\ &= \sigma_M((\operatorname{Arad}_\alpha M) \cap I) \end{aligned}$$

Finally, using 7.11 and the distributivity of $\mathcal{L}(M)$ 7.10, we get

$$\begin{aligned} \operatorname{Arad}_\alpha I &= \sigma_I^{-1}(\operatorname{Arad}(I/\operatorname{Arad}_\beta I)) \\ &= \sigma_I^{-1}(\psi^{-1}(\operatorname{Arad}((I + \operatorname{Arad}_\beta M)/\operatorname{Arad}_\beta M))) \\ &= \sigma_I^{-1}(\psi^{-1}(\sigma_M((\operatorname{Arad}_\alpha M) \cap I))) \\ &= \iota^{-1}(\sigma_M^{-1}(\sigma_M((\operatorname{Arad}_\alpha M) \cap I))) \\ &= \iota^{-1}((\operatorname{Arad}_\alpha M) \cap I + \operatorname{Arad}_\beta M) \\ &= I \cap ((\operatorname{Arad}_\alpha M) \cap I + \operatorname{Arad}_\beta M) \\ &= (\operatorname{Arad}_\alpha M) \cap I + (\operatorname{Arad}_\beta M) \cap I \\ &= (\operatorname{Arad}_\alpha M) \cap I \end{aligned}$$

- $\underline{\alpha \text{ is a limit ordinal}}$ We have $\operatorname{Arad}_\beta I = (\operatorname{Arad}_\beta M) \cap I$ for all $\beta < \alpha$. Thus

$$\operatorname{Arad}_\alpha I = \bigcup_{\beta<\alpha} \operatorname{Arad}_\beta I = \bigcup_{\beta<\alpha} ((\operatorname{Arad}_\beta M) \cap I) = (\bigcup_{\beta<\alpha} \operatorname{Arad}_\beta M) \cap I = (\operatorname{Arad}_\alpha M) \cap I.$$

The second claim is immediate from the first claim. $\square$

Corollary 14.7. *Let I be an order ideal in a decomposition monoid M and $x \in M$.*

1. *I is semi-Artinian if and only if $I \subseteq \operatorname{srad} M$.*
2. *$\{\prec x\}$ is semi-Artinian if and only if $x \in \operatorname{srad} M$*
3. *M is semi-Artinian if and only if $\{\prec x\}$ is semi-Artinian for all $x \in M$*
4. *$\operatorname{srad} M$ is the union of all semi-Artinian order ideals of M*

Proof. Immediate from 14.6. □

Corollary 14.8. *Let A, B be two order ideals in a decomposition monoid M. Then for all $\alpha \in \mathbf{Ord}$*

1. $\operatorname{Arad}_\alpha(A + B) = \operatorname{Arad}_\alpha A + \operatorname{Arad}_\alpha B$
2. $\operatorname{Arad}_\alpha(A \cap B) = (\operatorname{Arad}_\alpha A) \cap (\operatorname{Arad}_\alpha B)$
3. $\operatorname{srad}(A + B) = \operatorname{srad} A + \operatorname{srad} B$
4. $\operatorname{srad}(A \cap B) = (\operatorname{srad} A) \cap (\operatorname{srad} B)$
5. *$A + B$ is semi-Artinian if and only if A and B are semi-Artinian.*

Proof. Similar to proof of 12.6. □

Lemma 14.9. *Let $I_0 \leq I_1 \leq \cdots \leq M$ be a semi-Artinian series in M and $\sigma \colon M \to N$ an exact monoid homomorphism with N a decomposition monoid. Set $J_\alpha = \sigma(I_\alpha)$ for each $\alpha \in \mathbf{Ord}$. Then $J_0 \leq J_1 \leq \cdots \leq N$ is a semi-Artinian series in N.*

Proof. Since σ is exact, each J_α is an order ideal. We have three requirements to check...

1. Since $I_0 = \{\leq 0\}$, we have $J_0 = \sigma(I_0) \subseteq \{\leq 0\}$. Conversely, $J_0 = \sigma(I_0)$ is an order ideal which contains $\sigma(0) = 0$, so $J_0 \supseteq \{\prec 0\} = \{\leq 0\}$.
2. Let $\alpha \in \mathbf{Ord}$, and τ, ρ the quotient homomorphisms as in the diagram.

$$\begin{tikzcd} I_{\alpha+1} \arrow[r, "\sigma"] \arrow[d, "\tau"] & J_{\alpha+1} \arrow[d, "\rho"] \\ I_{\alpha+1}/I_\alpha \arrow[r, "\psi"] & J_{\alpha+1}/J_\alpha \end{tikzcd}$$

   Since $I_\alpha \subseteq \ker(\rho \circ \sigma)$, there is a homomorphism ψ making the diagram commute.
   Since N has decomposition, ρ and, hence $\rho\circ\sigma$ are exact. τ is surjective so, by 12.2, ψ is exact. $I_{\alpha+1}/I_\alpha$ is Artinian, so by 12.5.2, $\psi(I_{\alpha+1}/I_\alpha) = J_{\alpha+1}/J_\alpha$ is Artinian.
3. If $\alpha \in \mathbf{Ord}$ is a limit ordinal, then

$$J_\alpha = \sigma(I_\alpha) = \sigma(\bigcup_{\beta<\alpha} I_\beta) = \bigcup_{\beta<\alpha} \sigma(I_\beta) = \bigcup_{\beta<\alpha} J_\beta.$$

□

Proposition 14.10. *Let $\sigma \colon M \to N$ be an exact monoid homomorphism with N and M decomposition monoids. Then for all $\alpha \in \mathbf{Ord}$*

$$\sigma(\operatorname{Arad}_\alpha M) \subseteq \operatorname{Arad}_\alpha N,$$

and

$$\sigma(\operatorname{srad} M) \subseteq \operatorname{srad} N.$$

In particular, if M is semi-Artinian, then so is $\sigma(M)$.

Proof. The sequence $\operatorname{Arad}_0 M \leq \operatorname{Arad}_1 M \leq \ldots \leq M$ is a semi-Artinian series in M, and σ is exact, so by the previous proposition, $\sigma(\operatorname{Arad}_0 M) \leq \sigma(\operatorname{Arad}_1 M) \leq \ldots \leq N$ is a semi-Artinian series in N. From 14.4 we therefore get $\sigma(\operatorname{Arad}_\alpha M) \subseteq \operatorname{Arad}_\alpha N$. The rest of the claim then follows directly. □

If I is an order ideal in a decomposition monoid M then from 12.1.4, the quotient homomorphism is exact, so we get immediately

Corollary 14.11. *Let M be a decomposition monoid and $I \leq M$ an order ideal, with $\sigma\colon M \to M/I$ the quotient homomorphism. Then for all $\alpha \in$* **Ord**

$$\sigma(\operatorname{Arad}_\alpha M) \subseteq \operatorname{Arad}_\alpha(M/I),$$

and

$$\sigma(\operatorname{srad} M) \subseteq \operatorname{srad}(M/I).$$

In particular, if M is semi-Artinian, then so is M/I.

So far we have been constructing a semi-Artinian series using the Artinian radical. Since the Artinian radical is the largest Artinian order ideal (for decomposition monoids), the resulting series takes the largest possible jumps between successive order ideals.

We consider next the opposite extreme, namely, taking at each step a minimal non-trivial Artinian order ideal. If such an order ideal exists it will be generated by an atom:

Definition 14.12. *Let M be a monoid. An element $a \in M$ is an* **atom** *of M if it is minimal in the subclass $M \setminus \{\leq 0\}$.*

The following are easily checked consequences of the definition:

Proposition 14.13. *Let M be a monoid, $I \leq M$ an order ideal, and $a \in I$. Then the following are equivalent*

1. *a is an atom of M*
2. *$a \not\leq 0$ and for all $b \leq a$, either $b \leq 0$ or $b \equiv a$.*
3. *a is an atom of I.*
4. *$[a]$ is an atom of $\overline{M}$.*
5. *$[a] \neq [0]$ and $\overline{\{\leq a\}} = \{[0], [a]\} \subseteq \overline{M}$.*

In addition, atoms are indecomposable.

Also worth noting is that in a conical monoid, an element a is an atom if and only if $a \neq 0$ and $(b \leq a \implies b = 0 \text{ or } b = a)$.

Proposition 14.14. *If a is an atom of a decomposition monoid M, then $\{\prec a\}$ is Artinian.*

Proof. Since $\overline{\{\leq a\}} = \{[0], [a]\}$ is a finite set, $\{\leq a\}$ is Artinian. From 12.6.5, this implies that $\{\prec a\}$ is Artinian. □

Corollary 14.15. *For a decomposition monoid M,* $\operatorname{Arad} M \supset \{\leq 0\}$ *if and only if M has an atom.*

Proof. If M has an atom, then $\{\prec a\}$ is Artinian, so $\operatorname{Arad} M$ contains $a \notin \{\leq 0\}$. Conversely, if $\operatorname{Arad} M \supset \{\leq 0\}$ then the Artinian subclass $\operatorname{Arad} M \setminus \{\leq 0\}$ has a minimal element, which will be an atom of $\operatorname{Arad} M$ and hence an atom of M. □

Since $\operatorname{Arad} M = \{\leq 0\}$ if and only if $\operatorname{srad} M = \{\leq 0\}$ we get immediately

Corollary 14.16. *For a decomposition monoid M, $\operatorname{srad} M \supset \{\leq 0\}$ if and only if M has an atom.*

In particular, if M is a semi-Artinian decomposition monoid such that $M \supset \{\leq 0\}$ then M has an atom. More generally

Corollary 14.17. *If $I \subset J$ are order ideals in a semi-Artinian decomposition monoid M, then J/I has an atom.*

Proof. From 14.11 and 14.7, J/I is semi-Artinian. By hypothesis $J/I \supset \{0\} = \{\leq 0\}$, and so, from the previous proposition, J/I has an atom. □

The significance of this proposition is that, except for set theoretical considerations, the existence of atoms in subquotients characterizes semi-Artinian decomposition monoids. To show this, we will need to consider the difference between sets and proper classes:

Proposition 14.18. *If M is a decomposition monoid whose elements form a set, then there is some $\alpha \in$ **Ord** such that $\operatorname{srad} M = \operatorname{Arad}_\alpha M$.*

Proof. We prove the contrapositive...

Suppose $\operatorname{srad} M \neq \operatorname{Arad}_\alpha M$ for any $\alpha \in$ **Ord**. Consider the increasing map from **Ord** to $\mathcal{L}(M)$ given by $\alpha \mapsto \operatorname{Arad}_\alpha M$. We claim that this map is injective. Indeed if $\operatorname{Arad}_{\alpha+1} M = \operatorname{Arad}_\alpha M$ for some α, then, as we have already noted, $\operatorname{srad} M = \operatorname{Arad}_\alpha M$. This contradicts our assumption.

We have then an injective map from the proper class **Ord** into $\mathcal{L}(M)$, hence $\mathcal{L}(M)$ is a proper class. M itself must therefore be a proper class, since otherwise, if M were a set, then $\mathcal{L}(M)$ would be also. □

Lemma 14.19. *For an element x in a decomposition monoid, $\{\leq x\}$ is a set if and only if $\{\prec x\}$ is a set.*

Proof. Suppose $\{\leq x\}$ is a set. If $z \prec x$ then there is some $n \in \mathbb{N}$ such that $z \leq nx$. By the decomposition property there are $z_1, z_2, \ldots, z_n$ such that $z = z_1 + z_2 + \ldots + z_n$ and $z_i \leq x$ for all $i \leq n$. Thus every element of $\{\prec x\}$ is a finite sum of elements of $\{\leq x\}$. Since the class of finite subsets of $\{\leq x\}$ is a set, so is $\{\prec x\}$. The converse is trivial. □

Proposition 14.20. *Let M be a decomposition monoid.*

1. *If M is a set and M/I has an atom for all order ideals $I \subset M$, then M is semi-Artinian.*
2. *If $\{\leq x\}$ is a set for all $x \in M$ and for all order ideals $I \subset J$ of M, J/I has an atom, then M is semi-Artinian.*

Proof.

1. Since M is a set there is some $\alpha \in$ **Ord** such that $\operatorname{srad} M = \operatorname{Arad}_\alpha M$. In particular, $\operatorname{Arad}_{\alpha+1} M = \operatorname{Arad}_\alpha M$ so $\operatorname{Arad}(M/\operatorname{Arad}_\alpha M) = \{\leq 0\}$. From 14.15, $M/\operatorname{Arad}_\alpha M$ has no atoms, so by hypothesis, $\operatorname{Arad}_\alpha M = M$, and M is then semi-Artinian.
2. Fix an $x \in M$. From the lemma, $\{\prec x\}$ is a set and, by hypothesis, $\{\prec x\}/I$ has an atom for all $I \subset \{\prec x\}$. From 1, $\{\prec x\}$ is semi-Artinian.

   Using 14.7.3, M is then semi-Artinian.

□

A semi-Artinian series for a monoid M can be constructed inductively using atoms as follows:

1. Set $I_0 = \{\leq 0\}$
2. For any ordinal α, pick an atom a_α in M/I_α, or, if M/I_α has no atoms, pick $a_\alpha = 0$. Set $I_{\alpha+1} = \sigma_\alpha^{-1}(\{\prec a_\alpha\})$ where σ_α is the quotient homomorphism from M to M/I_α.
3. For any limit ordinal α set $I_\alpha = \bigcup_{\beta<\alpha} I_\beta$

A monoid may have many atoms, so it may have many semi-Artinian series of this type. In this construction, there is no obvious way of choosing a particular atom at each stage. Thus we are led to consider the order ideal generated by all the atoms. This will give us a canonical way of producing a semi-Artinian series, the socle series, in any monoid.

Definition 14.21. *Let M be a monoid. The* **socle** *of M (written $\operatorname{soc} M$) is the order ideal generated by all atoms of M.*

The following are easily checked consequences of the definition:

Proposition 14.22. *Let M be a monoid, and $I \leq M$ an order ideal. Then*

1. $\operatorname{soc} \overline{M} = \overline{\operatorname{soc} M}$
2. $\operatorname{soc} I \subseteq I \cap \operatorname{soc} M$

Proposition 14.23. *Let M be a decomposition monoid, $I \leq M$ an order ideal. Then*

1. $\operatorname{soc} M = \{\text{all finite sums of atoms of } M\} \bigcup \{\leq 0\}$
2. $\operatorname{soc} I = I \cap \operatorname{soc} M$

Proof.

1. If $x \in \operatorname{soc} M$ there are atoms $a_1, a_2, \ldots, a_n$ such that $x \leq a_1 + a_2 + \cdots + a_n$. From this relationship we get the decomposition matrix
$$\begin{array}{cc} & \begin{array}{cccc} \leq a_1 & \leq a_2 & \ldots & \leq a_n \end{array} \\ x & \begin{pmatrix} x_1 & x_2 & \ldots & x_n \end{pmatrix}. \end{array}$$
Since the a_i are atoms, the x_i are either atoms or in $\{\leq 0\}$.
If $x_i \in \{\leq 0\}$ for all i, then $x \in \{\leq 0\}$. If, on the other hand, one of the x_i is an atom then adding any elements of $\{\leq 0\}$ to this element leaves it an atom. Thus x can be written as a sum of atoms.
2. From 14.22, we have already that $\operatorname{soc} I \subseteq I \cap \operatorname{soc} M$.
To show the opposite inclusion, suppose $x \in I \cap \operatorname{soc} M$. Then from 1, either $x \in \{\leq 0\} \subseteq \operatorname{soc} I$, or there are atoms $a_1, a_2, \ldots, a_n$ such that $x = a_1 + a_2 + \cdots + a_n$. Since $a_i \leq x \in I$, each a_i is in I. Thus, using 14.13, $x \in \operatorname{soc} I$.

□

Proposition 14.24. *If M is a decomposition monoid then $\operatorname{soc} M$ is an Artinian order ideal.*

Proof. From 14.22.1, we can assume M is partially ordered.

Let $0 \neq a \in \operatorname{soc} M$. Then there are atoms $a_1, a_2, \ldots, a_n$ such that $a = a_1 + a_2 + \ldots + a_n$. Suppose $b \leq a$, then using decomposition, b can be written $b = b_1 + b_2 + \ldots + b_n$ with $b_i \leq a_i$ for all $i = 1, 2, \ldots, n$. For each i, a_i is an atom, so either $b_i = a_i$ or $b_i = 0$. Thus b is a sum of a subset of the atoms $a_1, a_2, \dots, a_n$, and $\{\leq a\}$ is a finite set. In particular, $\{\leq a\}$ is Artinian, and $a \in \operatorname{Arad} M$. Since this is true for all $a \in \operatorname{soc} M$, we have $\operatorname{soc} M \subseteq \operatorname{Arad} M$. □

This proposition allows us to construct a semi-Artinian series in any decomposition monoid based on the socle:

Definition 14.25. *Let M be a decomposition monoid. Define inductively an increasing sequence of order ideals, $\operatorname{soc}_0 M \leq \operatorname{soc}_1 M \leq \dots \leq \operatorname{soc}_\alpha M \leq \dots$ for $\alpha \in \mathbf{Ord}$, as follows:*

1. $\operatorname{soc}_0 M = \{\leq 0\}$
2. $\operatorname{soc}_{\alpha+1} M = \sigma_\alpha^{-1}(\operatorname{soc}(M/\operatorname{soc}_\alpha M))$ *where σ_α is the quotient homomorphism from M to $M/\operatorname{soc}_\alpha M$.*
3. *If α is a limit ordinal, define* $\operatorname{soc}_\alpha M = \bigcup_{\beta<\alpha} \operatorname{soc}_\beta M$

In addition, we define the **Loewy radical** *of M by*

$$\operatorname{Lrad} M = \bigcup_{\alpha \in \mathbf{Ord}} \operatorname{soc}_\alpha M.$$

As in 14.3, we have

$$(\operatorname{soc}_{\alpha+1} M)/\operatorname{soc}_\alpha M = \operatorname{soc}(M/\operatorname{soc}_\alpha M)$$

and so using 14.24, $\operatorname{soc}_0 M \leq \operatorname{soc}_1 M \leq \dots$ is a semi-Artinian series in M. Consequently, $\operatorname{Lrad} M \subseteq \operatorname{srad} M$.

Proposition 14.26. *If I is an order ideal in a decomposition monoid M, then for all $\alpha \in \mathbf{Ord}$*

$$\operatorname{soc}_\alpha I = (\operatorname{soc}_\alpha M) \cap I$$

and

$$\operatorname{Lrad} I = (\operatorname{Lrad} M) \cap I.$$

Proof. The proof is the same as the proof of 14.6, except for the use of 14.23.2 in place of 12.5.7. □

Given reasonable set theoretic conditions we get $\operatorname{Lrad} M = \operatorname{srad} M$:

Proposition 14.27. *Let M be a decomposition monoid.*

1. *If M is a set then there is some $\alpha \in \mathbf{Ord}$ such that* $\operatorname{Lrad} M = \operatorname{soc}_\alpha M$.
2. *If M is a set then* $\operatorname{Lrad} M = \operatorname{srad} M$.
3. *If $\{\leq x\}$ is a set for all $x \in M$ then* $\operatorname{Lrad} M = \operatorname{srad} M$.

Proof.

1. The proof is almost the same as for 14.18, since if $\operatorname{soc}_\alpha M = \operatorname{soc}_{\alpha+1} M$ for some $\alpha \in \mathbf{Ord}$ then $\operatorname{Lrad} M = \operatorname{soc}_\alpha M$.
2. From 1, there is some $\alpha \in \mathbf{Ord}$ such that $\operatorname{Lrad} M = \operatorname{soc}_\alpha M \subseteq \operatorname{srad} M$. So, in particular, $M/\operatorname{Lrad} M$ and hence, $(\operatorname{srad} M)/\operatorname{Lrad} M$, have no atoms. Since $\operatorname{srad} M$ is semi-Artinian, by 14.17, this is only possible if $\operatorname{Lrad} M = \operatorname{srad} M$.
3. Let $x \in \operatorname{srad} M$. Using 14.19, $\{\prec x\}$ is a set, so using 2 and 14.26,
$$x \in \{\prec x\} = \operatorname{srad}(\{\prec x\}) = \operatorname{Lrad}\{\prec x\} \subseteq \operatorname{Lrad} M.$$
Thus $\operatorname{srad} M \subseteq \operatorname{Lrad} M$. The opposite inclusion is always true, as we have already noted.

□

Since atoms are defined in terms of the order of the monoid, it is no surprise that they behave well with respect to exact and strictly increasing functions:

Proposition 14.28. *Let $\sigma\colon M \to N$ be a monoid homomorphism.*

1. *If σ is strictly increasing and $a \in N$ an atom, then all elements of $\sigma^{-1}(a)$ are atoms.*
2. *If σ is exact and $a \in M$ an atom, then $\sigma(a)$ is an atom or $\sigma(a) \leq 0$.*
3. *If σ is exact then $\sigma(\operatorname{soc} M) \subseteq \operatorname{soc} N$.*

Proof.

1. Let $a' \in \sigma^{-1}(a)$. If $a' \leq 0$ we would have $a = \sigma(a') \leq 0$, contrary to a being an atom. So we must have $a' \not\leq 0$.
   Now suppose $b \leq a'$. Then $\sigma(b) \leq \sigma(a') = a$, so either $\sigma(b) \leq 0$ or $\sigma(b) \geq a$. In the first case, we have $0 \leq b$ and $\sigma(0) \geq \sigma(b)$, so, since σ is strictly increasing, $b \leq 0$. In the second case we have $b \leq a'$ and $\sigma(b) \geq \sigma(a')$, so $b \geq a'$. Therefore a' is an atom of M.
2. Let $a' = \sigma(a)$ and suppose $a' \not\leq 0$. Then to show a' is an atom it remains to check that $b' \leq a'$ implies either $b' \leq 0$ or $b' \geq a'$...
   Suppose $b' \leq a' = \sigma(a)$. Then, using exactness, $b' \in \{\leq \sigma(a)\} \subseteq \sigma(\{\leq a\})$. So there is some $b \leq a$ such that $b' = \sigma(b)$. Since a is an atom, either $b \leq 0$ or $b \geq a$, and hence, either $b' \leq 0$ or $b' \geq a'$. Therefore a' is an atom of N.
3. If $a \in \operatorname{soc} M$ then there are atoms $a_1, a_2, \ldots, a_n$ such that $a \leq a_1 + a_2 + \ldots + a_n$. Thus in N, $\sigma(a) \leq \sigma(a_1) + \sigma(a_2) + \ldots + \sigma(a_n)$. From 2, for each $i \leq n$, $\sigma(a_i)$ is either an atom or in $\{\leq 0\}$. Thus $\sigma(a) \in \operatorname{soc}(N)$.

□

Proposition 14.29. *Let $\sigma\colon M \to N$ be an exact monoid homomorphism with N a decomposition monoid. Then for all $\alpha \in$ **Ord***

$$\sigma(\operatorname{soc}_\alpha M) \subseteq \operatorname{soc}_\alpha N.$$

Proof. By induction...

- $\underline{\alpha = 0}$ $\sigma(\operatorname{soc}_0 M) = \sigma(\{\leq 0\}) \subseteq \{\leq 0\} = \operatorname{soc}_0 N$
- $\underline{\alpha \text{ is a successor ordinal}}$ Suppose $\alpha = \beta + 1$ and $\sigma(\operatorname{soc}_\beta(M)) \subseteq \operatorname{soc}_\beta(N)$. Let τ, ρ be the quotient homomorphisms as in the diagram.

$$\begin{tikzcd} M \arrow[r, "\sigma"] \arrow[d, "\tau"] & N \arrow[d, "\rho"] \\ M/\operatorname{soc}_\beta M \arrow[r, "\psi"] & N/\operatorname{soc}_\beta N \end{tikzcd}$$

  Since $\sigma(\operatorname{soc}_\beta(M)) \subseteq \operatorname{soc}_\beta N$, there is a homomorphism ψ making the diagram commute.

Since N has decomposition, ρ and, hence $\rho \circ \sigma$, are exact. τ is surjective so, by 12.2, ψ is exact. By 14.28.3,

$$\psi(\operatorname{soc}(M/\operatorname{soc}_\beta M)) \subseteq \operatorname{soc}(N/\operatorname{soc}_\beta N)$$

Now suppose we have $a \in \operatorname{soc}_\alpha M$. Then $\tau(a) \in \operatorname{soc}(M/\operatorname{soc}_\beta M)$, and therefore $\rho(\sigma(a)) = \psi(\tau(a)) \in \operatorname{soc}(N/\operatorname{soc}_\beta N)$. This means that $\sigma(a) \in \operatorname{soc}_\alpha N$.

- $\underline{\alpha \text{ is a limit ordinal}}$ We assume $\sigma(\operatorname{soc}_\beta(M)) \subseteq \operatorname{soc}_\beta N$ for all $\beta < \alpha$. Thus

$$\begin{aligned}\sigma(\operatorname{soc}_\alpha M) &= \sigma(\bigcup_{\beta<\alpha} \operatorname{soc}_\beta M) = \bigcup_{\beta<\alpha} \sigma(\operatorname{soc}_\beta M)\\ &\subseteq \bigcup_{\beta<\alpha} \operatorname{soc}_\beta N = \operatorname{soc}_\alpha(N).\end{aligned}$$

□

For the remainder of this section we will turn our attention to strongly separative monoids. The reason for this is that the monoids $M(R\text{-}\mathbf{Noeth})$ and $M(R\text{-}\mathbf{Art})$, to be defined in Section 16, are strongly separative (as well as semi-Artinian).

Proposition 14.30.

1. *If M is a strongly separative decomposition monoid, then for all $a, b \in M$*

$$a \ll b \in \operatorname{soc} M \implies a \leq 0.$$

2. *If M is an Artinian monoid, then for all $b \in M$,*

$$((\forall a \in M)\ (a \ll b \implies a \leq 0)) \implies b \in \operatorname{soc} M.$$

Proof.

1. If $b \leq 0$ then $a \leq 0$ and we are done. Thus by 14.23, we can assume $b = b_1 + b_2 + \ldots + b_n$ for some atoms $b_1, b_2, \ldots, b_n$.

   We will assume first that a is an atom, and show by induction on n that this leads to a contradiction...

   $n = 1$ We have $a \leq b_1$. Since $a \not\leq 0$ and b_1 is an atom, this implies $a \equiv b_1$, and so $a \ll a$. But a strongly separative monoid has no proper regular elements, so this contradicts our hypothesis.

   $n > 1$ Since M has decomposition, a is prime. So from $a \leq b_1 + b_2 + \ldots + b_n$, there is some index I such that $a \leq b_I$. As above, this implies $a \equiv b_I$. Let $b' = \sum_{i \neq I} b_i$. Then $b \equiv a + b'$ and from $a \ll b$, we get $2a + b' \leq a + b'$. Since M is strongly separative, we can cancel a from this to get $a + b' \leq b'$, that is $a \ll b'$. Since b' is a sum of $n - 1$ atoms, we have completed the induction step.

   Finally, we consider the general situation. We have $a \in \operatorname{soc} M$, so either $a \leq 0$ or $a = a_1 + a_2 + \ldots + a_m$ for some atoms $a_1, a_2, \ldots, a_m$. In the second case we would have $a_1 \leq a \ll b$, so $a_1 \ll b$. Since a_1 is an atom, this leads to the contradiction discussed above. Thus we must have $a \leq 0$.

2. We prove the contrapositive...
   Suppose $b \notin \operatorname{soc} M$. Define

$$\mathcal{B} = \{b' \mid b' \leq b \text{ and } b' \notin \operatorname{soc} M\}.$$

   Since $b \in \mathcal{B}$, $\mathcal{B}$ is not empty and has a minimal element, b_0. $b_0 \notin \operatorname{soc} M$ so, in particular, b_0 is not an atom. Thus there is some $b_1 \leq b_0$ such that $b_1 \not\leq 0$ and $b_1 \not\geq b_0$. Let b_2 be such that $b_0 = b_1 + b_2$. We must have $b_2 \not\leq 0$, since otherwise we would get $b_0 \leq b_1$.
   $\operatorname{soc} M$ is an order ideal, so at least one of b_1 and b_2 must not be in $\operatorname{soc} M$. Since in the following we need only that $b_1, b_2 \not\leq 0$, we can, without loss of generality, assume that $b_1 \notin \operatorname{soc} M$.
   We have $b_1 \in \mathcal{B}$ and $b_1 \leq b_0$, so, by the minimality of b_0, $b_0 \leq b_1$. Thus, setting $a = b_2$, we get $b_0 + a \leq b_1 + a = b_0$, that is, $a \ll b_0$. Since $b_0 \leq b$, we get finally $a \ll b$ with $a \not\leq 0$.

□

Combining the two parts of this proposition we get

Corollary 14.31. *If M is a strongly separative Artinian decomposition monoid, then for all $b \in M$,*

$$((\forall a \in M)\ (a \ll b \implies a \leq 0)) \iff b \in \operatorname{soc} M.$$

Proposition 14.32. *Let N be a strongly separative decomposition monoid, and $\sigma\colon M \to N$ a monoid homomorphism such that $\sigma^{-1}(\{\leq 0\}) = \{\leq 0\}$. Then*

$$\sigma^{-1}(\operatorname{soc} N) \subseteq \operatorname{soc} M.$$

Proof. First we note the following fact:

Suppose $a, b \in \sigma^{-1}(\operatorname{soc} N)$ such that $\sigma(a) \ll \sigma(b)$. We have $\sigma(b) \in \operatorname{soc} N$ with N a strongly separative decomposition monoid, so from 14.30.1, $\sigma(a) \leq 0$. By the hypothesis, this implies $a \leq 0$. In particular, $a \ll b$.

From 12.1.5, this implies that σ is strictly increasing on $\sigma^{-1}(\operatorname{soc} N)$. From 14.24, $\operatorname{soc} N$ is Artinian, and so, from 12.5.1, $\sigma^{-1}(\operatorname{soc} N)$ is Artinian.

Let $b \in \sigma^{-1}(\operatorname{soc} N)$. Then from the above discussion we have that $a \ll b$ implies $a \leq 0$. Since $\sigma^{-1}(\operatorname{soc} N)$ is Artinian, 14.30.2 shows that $b \in \operatorname{soc}(\sigma^{-1}(\operatorname{soc} N))$, and so, by 14.22.2, $b \in \operatorname{soc} M$.

□

Proposition 14.33. *Let M and N be decomposition monoids with N strongly separative, and $\sigma\colon M \to N$ a monoid homomorphism such that $\sigma^{-1}(\{\leq 0\}) = \{\leq 0\}$. Then for all $\alpha \in$* **Ord**

$$\sigma^{-1}(\operatorname{soc}_\alpha N) \subseteq \operatorname{soc}_\alpha M.$$

Proof. By induction...

- <u>$\alpha = 0$</u> By hypothesis, we have $\sigma^{-1}(\operatorname{soc}_0 N) = \sigma^{-1}(\{\leq 0\}) = \{\leq 0\} = \operatorname{soc}_0 M$.
- <u>α is a successor ordinal</u> Suppose $\alpha = \beta + 1$ and $\sigma^{-1}(\operatorname{soc}_\beta N) \subseteq \operatorname{soc}_\beta M$.

Let μ, τ and ρ be quotient homomorphisms as in the diagram:

$$\begin{tikzcd}
 & M \arrow[dl, "\mu"'] \arrow[d, "\tau"] \arrow[r, "\sigma"] & N \arrow[d, "\rho"] \\
M/\operatorname{soc}_\beta M & M/\sigma^{-1}(\operatorname{soc}_\beta N) \arrow[l, "\nu"'] \arrow[r, "\psi"] & N/\operatorname{soc}_\beta N
\end{tikzcd}$$

Since $\sigma(\sigma^{-1}(\operatorname{soc}_\beta N)) \subseteq \operatorname{soc}_\beta N$ and $\sigma^{-1}(\operatorname{soc}_\beta N) \subseteq \operatorname{soc}_\beta M$, we can fill in the diagram with homomorphisms ν and ψ making it commute.

N is strongly separative, so by 8.15, $N/\operatorname{soc}_\beta N$ is also strongly separative. Also, by construction, $\psi^{-1}(\{\leq 0\}) = \{\leq 0\}$, so, using 14.32, we get

$$\psi^{-1}(\operatorname{soc}(N/\operatorname{soc}_\beta N)) \subseteq \operatorname{soc}(M/\sigma^{-1}(\operatorname{soc}_\beta N)).$$

Since M has decomposition, the homomorphism ν is exact (12.2). From 14.28.3, we then get

$$\nu(\operatorname{soc}(M/\sigma^{-1}(\operatorname{soc}_\beta N))) \subseteq \operatorname{soc}(M/\operatorname{soc}_\beta M).$$

Now suppose $a \in \sigma^{-1}(\operatorname{soc}_\alpha(N))$. We will show that $a \in \operatorname{soc}_\alpha M$...

We have $\sigma(a) \in \operatorname{soc}_\alpha N$, so $\psi(\tau(a)) = \rho(\sigma(a)) \in \operatorname{soc}(N/\operatorname{soc}_\beta N)$. This implies that $\tau(a) \in \psi^{-1}(\operatorname{soc}(N/\operatorname{soc}_\beta N))$. From above, $\tau(a) \in \operatorname{soc}(M/\sigma^{-1}(\operatorname{soc}_\beta N)$. Then $\mu(a) = \nu(\tau(a)) \in \nu(\operatorname{soc}(M/\sigma^{-1}(\operatorname{soc}_\beta N)))$, and from above, $\mu(a) \in \operatorname{soc}(M/\operatorname{soc}_\beta M)$. Hence $a \in \operatorname{soc}_\alpha M$.

- <u>α is a limit ordinal</u> We assume $\sigma^{-1}(\operatorname{soc}_\beta N) \subseteq \operatorname{soc}_\beta M$ for all $\beta < \alpha$. Thus

$$\begin{aligned}\sigma^{-1}(\operatorname{soc}_\alpha(N)) = \sigma^{-1}(\bigcup_{\beta<\alpha} \operatorname{soc}_\beta N) &= \bigcup_{\beta<\alpha} \sigma^{-1}(\operatorname{soc}_\beta N) \\ &\subseteq \bigcup_{\beta<\alpha} \operatorname{soc}_\beta M) = \operatorname{soc}_\alpha M\end{aligned}$$

□

15 Semi-Artinian Refinement Monoids

In this section we consider the properties of semi-Artinian refinement monoids. Unlike Artinian monoids, semi-Artinian monoids are not, in general, primely generated, so we do not expect to get all the properties that Artinian monoids have. The main result of this section is that semi-Artinian refinement monoids are midseparative (and hence separative). We will also show by example that semi-Artinian monoids do not have weak cancellation or $\leq$-multiplicative cancellation.

The key to showing that semi-Artinian monoids are separative is the extension property of separative monoids given in 8.16: If I is a separative order ideal in a refinement monoid M such that M/I is separative, then M is separative. In particular, if I and M/I are Artinian, then M is separative.

Proposition 15.1. *Any semi-Artinian refinement monoid is separative.*

Proof. Let $I_0 \leq I_1 \leq \cdots \leq M$ be a semi-Artinian series such that $M = \bigcup_\alpha I_\alpha$. We do the proof by induction…

First we note that $I_0 = \{\leq 0\}$ is a group so I_0 is both cancellative and separative. Now suppose $\alpha \in \mathbf{Ord}$ such that I_β is separative for all $\beta < \alpha$. We have two cases:

- If α is a successor ordinal, $\alpha = \beta + 1$, then I_β is separative, and I_α/I_β is an Artinian refinement monoid, hence is also separative. From 8.16, this implies I_α is separative.
- If α is a limit ordinal, then $I_\alpha = \bigcup_{\beta<\alpha} I_\beta$. This is easily seen to imply that I_α is separative.

Finally, since I_α is separative for all α, M is also separative. $\square$

The extension property of separative monoids used in this proof is not shared by weakly cancellative or midseparative monoids. See Examples 9.7 and 15.9. Nonetheless for the midseparative case we have the extension property given in 13.6: If I is an Artinian order ideal in a refinement monoid M such that M/I is midseparative, then M is midseparative. In particular, if I and M/I are Artinian, then M is midseparative. The existence of this property suggests that semi-Artinian refinement monoids are midseparative. To prove this we will first prove a semi-Artinian monoid version of Proposition 13.4.

Theorem 15.2. *Let a and b be elements of a refinement monoid M such that $2a = a + b$. If $\{\prec b\}$ is semi-Artinian, then there exists an idempotent e such that $a = b + e$.*

Proof. Let $I_0 \leq I_1 \leq \ldots \leq \{\prec b\}$ be a semi-Artinian series such that $\{\prec b\} = \bigcup_\alpha I_\alpha$. As in 13.4, we define

$$\mathcal{B} = \{b' \in M \mid \exists a', d' \text{ such that } a = d' + a', b = d' + b', 2a' = a' + b'\}.$$

Now $b \in \mathcal{B}$, and $b \in I_\gamma$ for some $\gamma \in \mathbf{Ord}$, so $\mathcal{B} \cap I_\gamma \neq \emptyset$.

Let $\alpha \in \mathbf{Ord}$ be the least ordinal such that $\mathcal{B} \cap I_\alpha \neq \emptyset$. We will prove that $\alpha = 0$ by showing that if α is either a limit or successor ordinal then we get a contradiction…

- $\underline{\alpha \text{ is a limit ordinal}}$ Let $b_0 \in \mathcal{B} \cap I_\alpha$. Since $I_\alpha = \bigcup_{\beta<\alpha} I_\beta$, there must be some $\beta < \alpha$ with $b_0 \in I_\beta$ and hence $b_0 \in \mathcal{B} \cap I_\beta$. This contradicts the minimality of α.
- $\underline{\alpha \text{ is a successor ordinal}}$ Suppose $\alpha = \beta + 1$. Let $\sigma_\beta \colon \{\prec b\} \to \{\prec b\}/I_\beta$ be the quotient homomorphism. Then $\sigma_\beta(I_\alpha) = I_\alpha/I_\beta$ is Artinian and $\mathcal{B}' = \sigma_\beta(\mathcal{B} \cap I_\alpha)$ is a nonempty subclass of an Artinian monoid.

  Let $b_0 \in \mathcal{B} \cap I_\alpha$ be chosen so that $\sigma_\beta(b_0)$ is minimal in $\mathcal{B}'$. Since $b_0 \in \mathcal{B}$, there are a_0 and d_0 such that $a = d_0 + a_0$, $b = d_0 + b_0$, and $2a_0 = a_0 + b_0$.

  Exactly as in 13.4, by 8.6, there is a refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} b_0 & a_1 \end{array} \\ \begin{array}{c} a_0 \\ a_1 \end{array} & \begin{pmatrix} d_2 & a_2 \\ b_2 & c_2 \end{pmatrix} \end{array}$$

  with $c_2 \leq a_2$ and $2b_2 \leq a_0$. We also have $a = (d_0 + d_2) + a_2$, $b = (d_0 + d_2) + b_2$, and, by 8.1.2, $2a_2 = a_2 + b_2$, so $b_2 \in \mathcal{B}$. Since $b_2 \leq b_0$, the minimality of $\sigma_\beta(b_0)$ implies $\sigma_\beta(b_0) \leq \sigma_\beta(b_2)$. In particular, $\sigma_\beta(2b_0) \leq \sigma_\beta(2b_2) \leq \sigma_\beta(a_0)$.

  Since $\sigma_\beta(2b_0) \leq \sigma_\beta(a_0)$, there is some $u \in I_\beta$ such that $2b_0 \leq a_0 + u$. We make a refinement of this inequality

$$\begin{array}{cc} & \begin{array}{cc} \leq a_0 & \leq u \end{array} \\ \begin{array}{c} b_0 \\ b_0 \end{array} & \begin{pmatrix} a_3 & u_3 \\ a_4 & u_4 \end{pmatrix} \end{array}$$

  and then a further refinement of $b_0 = a_3 + u_3 = a_4 + u_4$:

$$\begin{array}{cc} & \begin{array}{cc} a_3 & u_3 \end{array} \\ \begin{array}{c} a_4 \\ u_4 \end{array} & \begin{pmatrix} d' & y_1 \\ y_2 & y_3 \end{pmatrix} \end{array}$$

  Set $b' = y_1 + y_2 + y_3$ so that $b_0 = d' + b'$. We will show that $b' \in \mathcal{B} \cap I_\beta \ldots$

  We have $2d' \leq a_3 + a_4 \leq a_0$, so there is some x such that $a_0 = 2d' + x$. Set $a' = d' + x$ so that $a_0 = d' + a'$. Then $a = d_0 + a_0 = (d_0 + d') + a'$ and $b = d_0 + b_0 = (d_0 + d') + b'$.

  Note that $d' \leq a_3 \leq b_0 \leq b$, so that $d' \in \{\prec b\}$ and also that, from 15.1, $\{\prec b\}$ is separative. From the equation $2a_0 = a_0 + b_0$ we get $2a' + 2d' = a' + b' + 2d'$ with $d' \leq 2a', a' + b'$. We can then use 8.14.3 to cancel $2d'$ and get $2a' = a' + b'$. Thus $b' \in \mathcal{B}$.

  Also $b' \leq u_3 + u_4 \leq u \in I_\beta$, so $b' \in I_\beta$, that is $b' \in I_\beta \cap \mathcal{B}$. Since $\beta < \alpha$, this contradicts α being the least ordinal such that $\mathcal{B} \cap I_\alpha \neq \emptyset$.

Since α is neither a limit ordinal or a successor ordinal, we must have $\alpha = 0$. Thus there is some $b_0 \in I_0 \cap \mathcal{B} = \{\leq 0\} \cap \mathcal{B}$. Let a_0, d_0 be the corresponding elements such that $a = d_0 + a_0$, $b = d_0 + b_0$ and $2a_0 = a_0 + b_0$. Since $b_0 \leq 0$, there is some b_0' such that $b_0 + b_0' = 0$. Set $e = a_0 + b_0'$. It is then easy to check that e is an idempotent such that $a_0 = b_0 + e$. Adding d_0 to this equation gives $a = b + e$. □

As an immediate corollary we have

Corollary 15.3. *Any semi-Artinian refinement monoid is midseparative.*

The next two results follow from 15.2 in the same way that 13.5 and 13.6 follow from 13.4:

Corollary 15.4. *Let I be an semi-Artinian order ideal in a refinement monoid M.*

1. *If $a, b, c \in M$ with $a + c = b + c$, $c \le a$ and $b \in I$, then there is an idempotent $e \le c$ such that $a = b + e$.*
2. *If $2[e]_I = [e]_I$ for some $e \in M$, then there is an idempotent $e' \le e$ such that $[e']_I = [e]_I$.*
3. *If $2[e]_I = [e]_I \le [a]_I$ for some $e, a \in M$, then there is an idempotent $e' \le a$ such that $[e']_I = [e]_I$.*

Proof. Exactly as in the proof of 13.5. □

Proposition 15.5. *Let I be a semi-Artinian order ideal in a refinement monoid M. Then if M/I is midseparative, so is M.*

Proof. Exactly as in the proof of 13.6. □

The next goal of this section is to show that semi-Artinian refinement monoids do not, in general, have weak cancellation or $\le$-multiplicative cancellation. Constructing a counterexample in the $\le$-multiplicative cancellation case is made more difficult by the fact that any weakly cancellative semi-Artinian refinement monoid has $\le$-multiplicative cancellation. To show this we will need the following two lemmas:

Lemma 15.6. [31, Lemma 1.9] *Let M be a refinement monoid and $a, b, c \in M$ such that $a + b = nc$ for some $n \in \mathbb{N}$. Then there are $x_0, x_1, \ldots, x_n \in M$ such that*

$$a = \sum_{k=0}^{n} k x_k \qquad b = \sum_{k=0}^{n} (n-k) x_k \qquad c = \sum_{k=0}^{n} x_k.$$

Proof. The $n = 1$ case is trivial. We will prove the other cases by induction...

Suppose the claim is true for some $n \in \mathbb{N}$ and we have $a + b = nc + c$ in M. Then we make a refinement of this equation

$$\begin{array}{cc} & \begin{matrix} nc & c \end{matrix} \\ \begin{matrix} a \\ b \end{matrix} & \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \end{array}$$

Since $a_1 + b_1 = nc$, we use the induction hypothesis to get $y_0, y_1, \ldots, y_n \in M$ such that $a_1 = \sum_{k=0}^{n} k y_k$, $b_1 = \sum_{k=0}^{n} (n-k) y_k$ and $c = \sum_{k=0}^{n} y_k$. We also have $a_2 + b_2 = c = \sum_{k=0}^{n} y_k$ so we get a refinement

$$\begin{array}{cc} & \begin{matrix} y_0 & y_1 & \ldots & y_n \end{matrix} \\ \begin{matrix} a_2 \\ b_2 \end{matrix} & \begin{pmatrix} u_0 & u_1 & \ldots & u_n \\ v_0 & v_1 & \ldots & v_n \end{pmatrix} \end{array}$$

For $k = 0, 1, 2, \ldots, n, n+1$, define

$$x_k = \begin{cases} v_0 & k = 0 \\ u_{k-1} + v_k & 1 \le k \le n \\ u_n & k = n+1 \end{cases}$$

It is then straight forward to check that $x_0, x_1, \ldots, x_{n+1}$ satisfy the equations

$$a = \sum_{k=0}^{n+1} kx_k \qquad b = \sum_{k=0}^{n+1}(n+1-k)x_k \qquad c = \sum_{k=0}^{n+1} x_k.$$

$\square$

Lemma 15.7. *Let M be a weakly cancellative refinement monoid, and $a_0, b_0 \in M$ such that $na_0 \leq nb_0$ for some $n \in \mathbb{N}$. Then there are a_1, b_1, d_1 such that $a_0 = d_1 + a_1$, $b_0 = d_1 + b_1$, $na_1 \leq nb_1$ and $na_1 \leq (n-1)a_0$.*

Proof. Let $u \in M$ be such that $na_0 + u = nb_0$. We write this equation as $a_0 + ((n-1)a_0 + u) = nb_0$ and then use the previous lemma to get $x_0, x_1, \ldots, x_n \in M$ such that $a_0 = \sum_{k=0}^{n} kx_k$ and $b_0 = \sum_{k=0}^{n} x_k$.

Set $d_1 = \sum_{k=1}^{n} x_k$ and $a_1 = \sum_{k=2}^{n}(k-1)x_k$ so that $a_0 = d_1 + a_1$ and $b_0 = d_1 + x_0$. We also have

$$na_1 = n\sum_{k=2}^{n}(k-1)x_k = \sum_{k=2}^{n}(nk-n)x_k \leq \sum_{k=2}^{n}(nk-k)x_k = (n-1)\sum_{k=2}^{n} kx_k \leq (n-1)a_0.$$

From $na_0 \leq nb_0$ we now get $nd_1 + na_1 \leq nd_1 + nx_0$. Using separativity, we cancel $(n-1)d_1$ from this to get $d_1 + na_1 \leq d_1 + nx_0$, and using 9.13.1, there is some $y_1 \ll d_1$ such that $na_1 \leq nx_0 + y_1$. Let y_2 be such that $d_1 = d_1 + y_1 + y_2$, and set $b_1 = x_0 + y_1 + y_2$. Then $d_1 + b_1 = d_1 + x_0 + y_1 + y_2 = d_1 + x_0 = b_0$ and $na_1 \leq nx_0 + y_1 \leq n(x_0 + y_1 + y_2) = nb_1$. $\square$

Theorem 15.8. *Any semi-Artinian refinement monoid which has weak cancellation, also has $\leq$-multiplicative cancellation.*

Proof. Suppose a and b are elements of a weakly cancellative semi-Artinian refinement monoid M such that $na \leq nb$ for some $n \geq 2$. We will show that $a \leq b$...

Let $I_0 \leq I_1 \leq \ldots \leq M$ be a semi-Artinian series such that $M = \bigcup_\alpha I_\alpha$. We define

$$\mathcal{A} = \{a' \mid \exists b', d' \text{ such that } a = d' + a', b = d' + b', \text{ and } na' \leq nb'\}.$$

Now $a \in \mathcal{A}$, and $a \in I_\gamma$ for some $\gamma \in \mathbf{Ord}$, so $\mathcal{A} \cap I_\gamma \neq \emptyset$.

Let $\alpha \in \mathbf{Ord}$ be the least ordinal such that $\mathcal{A} \cap I_\alpha \neq \emptyset$. We will prove that $\alpha = 0$ by showing that if α is either a limit or successor ordinal then we get a contradiction...

- <u>α is a limit ordinal</u> Let $a_0 \in \mathcal{A} \cap I_\alpha$. Since $I_\alpha = \bigcup_{\beta<\alpha} I_\beta$, there must be some $\beta < \alpha$ with $a_0 \in I_\beta$ and hence $a_0 \in \mathcal{A} \cap I_\beta$. This contradicts the minimality of α.
- <u>α is a successor ordinal</u> Suppose $\alpha = \beta + 1$. Let $\sigma_\beta \colon M \to M/I_\beta$ be the quotient homomorphism. Then $\sigma_\beta(I_\alpha) \cong I_\alpha/I_\beta$ is an Artinian monoid and $\mathcal{A}' = \sigma_\beta(\mathcal{A} \cap I_\alpha)$ is a nonempty subclass of an Artinian monoid.

  Let $a_0 \in \mathcal{A} \cap I_\alpha$ be chosen so that $\sigma_\beta(a_0)$ is minimal in $\mathcal{A}'$. Since $a_0 \in \mathcal{A}$, there are b_0, d_0 such that $a = d_0 + a_0$, $b = d_0 + b_0$, and $na_0 \leq nb_0$.

  By 15.7, there are a_1, b_1, d_1 such that $a_0 = d_1 + a_1$, $b_0 = d_1 + b_1$, $na_1 \leq nb_1$ and $na_1 \leq (n-1)a_0$. We have $a = (d_0 + d_1) + a_1$ and $b = (d_0 + d_1) + b_1$ and so $a_1 \in \mathcal{A}$. Since also $a_1 \leq a_0 \in I_\alpha$, we have $a_1 \in \mathcal{A} \cap I_\alpha$ and $\sigma_\beta(a_1) \in \mathcal{A}'$ with $\sigma_\beta(a_1) \leq \sigma_\beta(a_0)$. The minimality of $\sigma_\beta(a_0)$ then implies $\sigma_\beta(a_0) \leq \sigma_\beta(a_1)$. In particular, $\sigma_\beta(na_0) \leq \sigma_\beta(na_1) \leq \sigma_\beta((n-1)a_0)$.

Since $\sigma_\beta(na_0) \le \sigma_\beta((n-1)a_0)$, there is some $u \in I_\beta$ such that $na_0 \le (n-1)a_0 + u$. The monoid M is separative so we can cancel $(n-2)a$ from this to get $2a_0 \le a_0 + u$. Using 9.13.1, there is some $y_0 \ll a_0$ such that $a_0 \le u + y_0$. Decomposing this, we get $a_0 = a_2 + y_1$ with $a_2 \le u$ and $y_1 \le y_0 \ll a_0$, that is, $y_1 \ll a_0$. We will show that $a_2 \in \mathcal{A}$...

Let y_2 be such that $a_0 = a_0 + y_1 + y_2$. Since $na_0 \le nb_0$, we have $nb_0 = nb_0 + y_1 + y_2$, and since M is separative, $b_0 = b_0 + y_1 + y_2$. Set $b_2 = b_0 + y_2$. Then $b_0 = b_2 + y_1$, so $a = (d_0 + y_1) + a_2$ and $b = (d_0 + y_1) + b_2$. Also $na_2 \le na_0 \le nb_0 \le nb_2$, so $a_2 \in \mathcal{A}$.

We also have $a_2 \le u \in I_\beta$, and so $a_2 \in I_\beta$, that is $a_2 \in I_\beta \cap \mathcal{A}$. Since $\beta < \alpha$, this contradicts α being the least ordinal such that $\mathcal{A} \cap I_\alpha \ne \emptyset$.

Since α is neither a limit ordinal or a successor ordinal, we must have $\alpha = 0$, and so there is some $a_0 \in I_0 \cap \mathcal{A} = \{\le 0\} \cap \mathcal{A}$. Since $a_0 \in \mathcal{A}$, there are b_0, d_0 such that $a = d_0 + a_0$, $b = d_0 + b_0$. Finally $a_0 \le 0$, so we get $a = d_0 + a_0 \le d_0 \le b$. □

This proposition explains the complexity of the following example, which to exhibit failure of $\le$-multiplicative cancellation in a semi-Artinian refinement monoid must also fail weak cancellation. This example has many other interesting properties and serves also as a general purpose counterexample for many hoped-for-but-not-true claims about monoids.

Example 15.9. *A semi-Artinian refinement monoid containing two incomparable elements a_0, b_0 such that $2a_0 = 2b_0$.*

Let F be the free monoid on the generators, $c', d', a'_0, b'_0, a'_1, b'_1, a'_2, b'_2, \ldots$. Let $\sim$ be the congruence on F generated by

$$\begin{aligned} 2a'_n &\sim 2b'_n \\ a'_n &\sim a'_{n+1} + c' \sim b'_{n+1} + d' \\ b'_n &\sim b'_{n+1} + c' \sim a'_{n+1} + d' \end{aligned}$$

for all $n \in \mathbb{Z}^+$.

Let $M = F/\sim$ and $\sigma\colon F \to M$ the quotient homomorphism. We write $a_n = \sigma(a'_n)$, $b_n = \sigma(b'_n)$, $c = \sigma(c')$ and $d = \sigma(d')$. Then in M we have

$$\begin{aligned} 2a_n &= 2b_n \\ a_n &= a_{n+1} + c = b_{n+1} + d \\ b_n &= b_{n+1} + c = a_{n+1} + d. \end{aligned}$$

Note also that $a_n = a_{n+2} + 2c = a_{n+2} + 2d$, $b_n = b_{n+2} + 2c = b_{n+2} + 2d$, and also $a_n + 2c = a_n + 2d$ and $b_n + 2c = b_n + 2d$ for all $n \in \mathbb{Z}^+$.

- **Claim** *a_n and b_n are incomparable:*

  Define monoid homomorphism $\alpha'\colon F \to \mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$ by

$$\alpha'(a'_n) = (1, n, 0) \qquad \alpha'(b'_n) = (1, n, 1)$$
$$\alpha'(c') = (0, -1, 0) \qquad \alpha'(d') = (0, -1, 1)$$

for all $n \in \mathbb{Z}^+$. $\mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$ is a cancellative refinement monoid. Since

$$\begin{aligned} \alpha'(2a'_n) &= \alpha'(2b'_n) \\ \alpha'(a'_n) &= \alpha'(a'_{n+1}) + \alpha'(c') = \alpha'(b'_{n+1}) + \alpha'(d') \\ \alpha'(b'_n) &= \alpha'(b'_{n+1}) + \alpha'(c') = \alpha'(a'_{n+1}) + \alpha'(d'), \end{aligned}$$

there is an induced monoid homomorphism $\alpha\colon M \to \mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$ such that $\alpha' = \alpha \circ \sigma$, in particular,

$$\begin{aligned} \alpha(a_n) &= (1, n, 0) & \alpha(b_n) &= (1, n, 1) \\ \alpha(c) &= (0, -1, 0) & \alpha(d) &= (0, -1, 1) \end{aligned}$$

Since the images of $c, d, a_0, b_0, a_1, b_1, a_2, b_2, \ldots$ are distinct elements in $\mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$, these elements are distinct in M.

$\alpha(M)$ is $(\mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2) \setminus ((\{0\} \times \mathbb{N} \times \mathbb{Z}_2) \cup \{(0, 0, 1)\})$, which is, of course, a submonoid of $\mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$. Notice that $\alpha(c) + \alpha(c) = \alpha(d) + \alpha(d)$. This equation has no refinement in $\alpha(M)$, as is easily checked, so $\alpha(M)$ does not have refinement.

Now we show that for each $n \in \mathbb{N}$, the elements a_n and b_n are incomparable...

If $a_n \leq b_n$, then there would be some $u \in M$ such that $a_n + u = b_n$. Applying α to this equation gives $(1, n, 0) + \alpha(u) = (1, n, 1)$ and hence $\alpha(u) = (0, 0, 1)$. But $(0, 0, 1) \notin \alpha(M)$, so no such u can exist. Therefore $a_n \not\leq b_n$ and similarly, $b_n \not\leq a_n$.

- **Claim** *M is partially ordered:*

  A straight forward calculation shows that for two elements $u', v' \in F$, we have $\alpha'(u') + \alpha'(v') = (0, 0, 0)$ if and only if $u' = v' = 0$. Thus we have a similar result for two elements $u, v \in M$: $\alpha(u) + \alpha(v) = (0, 0, 0)$ if and only if $u = v = 0$.

  A consequence of this is that $u \ll x$ in M if and only if $u = 0$: If $x + u \leq x$ then there is v such that $x + u + v = x$ and hence $\alpha(x) = \alpha(u) + \alpha(v) + \alpha(x)$. Cancellation in $\mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$ gives $\alpha(u) + \alpha(v) = (0, 0, 0)$, and so $u = 0$.

  Now it is easy to show that M is partially ordered...

  If $x \equiv y$ in M then there is $u \ll x$ such that $x = y + u$. But $u \ll x$ implies $u = 0$, and so $x = y$.

- **Claim** *M is not weakly cancellative:*

  Let $I \subseteq M$ be the submonoid generated by c and d. In fact, I is an order ideal, since it is the inverse image under α of the order ideal $\{0\} \times \mathbb{Z} \times \mathbb{Z}_2 \leq \mathbb{Z}^+ \times \mathbb{Z} \times \mathbb{Z}_2$. We will also show that I is free...

  Write $(\mathbb{Z}^+ \times \mathbb{Z}^+)^\infty$ for the monoid obtained by adjoining an infinite element to $\mathbb{Z}^+ \times \mathbb{Z}^+$.

  We define a monoid homomorphism $\beta'\colon F \to (\mathbb{Z}^+ \times \mathbb{Z}^+)^\infty$

$$\begin{aligned} \beta'(a'_n) &= \infty & \beta'(b'_n) &= \infty \\ \beta'(c') &= (1, 0) & \beta'(d') &= (0, 1) \end{aligned}$$

Just as for the homomorphism α', β' induces a homomorphism $\beta\colon M \to (\mathbb{Z}^+ \times \mathbb{Z}^+)^\infty$ such that

$$\begin{aligned} \beta(a_n) &= \infty & \beta(b_n) &= \infty \\ \beta(c) &= (1, 0) & \beta(d) &= (0, 1) \end{aligned}$$

Using this homomorphism it is easy to see that $m_1c + n_1d = m_2c + n_2d$ *in* M *if and only if* $m_1 = m_2$ *and* $n_1 = n_2$*. Thus* c *and* d *generate a free submonoid in* M*. One immediate consequence is that* $2c \neq 2d$*, so the equation* $a_0 + 2c = a_0 + 2d$ *does not cancel. This means that* M *is not cancellative, and, since for any* u*,* $u \ll a_0$ *implies* $u = 0$*,* M *can not be weakly cancellative either.*

Note that $\alpha(M)$ *is cancellative but* M *is not, so* M *and* $\alpha(M)$ *are not isomorphic.*

- **Claim** *M is semi-Artinian:*

  We have already noted that the order ideal I *is isomorphic to* $\mathbb{Z}^+ \times \mathbb{Z}^+$*, so* I *is Artinian and* $I \subseteq \operatorname{Arad} M$*. In fact, we will show that* $I = \operatorname{Arad} M \ldots$

  We have $a_n = a_{n+1} + c$ *and* $b_n = b_{n+1} + c$ *for all* $n \in \mathbb{Z}^+$*, so* $a_0 \geq a_1 \geq a_2 \geq \ldots$ *and* $b_0 \geq b_1 \geq b_2 \geq \ldots$*. These sequences must be strictly decreasing since* M *is partially ordered and all the* a_n *and* b_n *are distinct. Thus* M *is not Artinian, no* a_n *or* b_n *is in* $\operatorname{Arad} M$*, and* $I = \operatorname{Arad} M$*.*

  In M/I*, it is easy to check that* $[a_0]_I = [a_m]_I = [b_n]_I$ *for all* $m, n \in \mathbb{Z}^+$*, so that* M/I *has a single generator. We will show, in fact, that* $M/I \cong \mathbb{Z}^+ \ldots$

  Let $\gamma' \colon F \to \mathbb{Z}^+$ *be the monoid homomorphism defined by* $\gamma'(a'_n) = \gamma'(b'_n) = 1$ *for all* $n \in \mathbb{Z}^+$ *and* $\gamma'(c') = \gamma'(d') = 0$*. This homomorphism is surjective. It is easily checked that* γ' *induces a surjective monoid homomorphism* $\gamma \colon M \to \mathbb{Z}^+$ *such that* $\gamma(a_n) = \gamma(b_n) = 1$ *for all* $n \in \mathbb{Z}^+$ *and* $\gamma(c) = \gamma(d) = 0$*. Since* $I \subseteq \ker \gamma$*, there is another induced homomorphism* $\bar{\gamma} \colon M/I \to \mathbb{Z}^+$ *such that* $\bar{\gamma}([a_0]_I) = 1$*. Since* $[a_0]_I$ *generates* M/I *and* $\bar{\gamma}$ *is surjective, we have* $M/I \cong \mathbb{Z}^+$*.*

  Thus both I *and* M/I *are Artinian and so* M *is semi-Artinian. Notice also that* I *and* M/I *are cancellative but* M *is not, and* I *and* M/I *are weakly cancellative but* M *is not.*

- **Claim** *M has refinement:*

  Let $x \in M$*. Then* x *can, in general, be written as a sum of the generators in many ways. Fix such an expression and let* $N \in \mathbb{N}$ *be a number greater than any subscript of* a *or* b *appearing in this expression. If there is no* a_n *or* b_n *in this expression, then any* $N \in \mathbb{N}$ *will do. Then using the rules* $a_n = a_{n+1} + c$ *and* $b_n = a_{n+1} + d$ *a sufficient number of times,* x *can be written in the form*

  $$x = la_N + mc + nd$$

  for some $l, m, n \in \mathbb{Z}^+$*. Note that* $\alpha(x) = (l,\ Nl - m - n,\ n \pmod 2)$*.*

  We now consider how to construct a refinement of the equation $x_1 + x_2 = x_3 + x_4$ *for elements* $x_1, x_2, x_3, x_4 \in M \ldots$

  If we happened to have $x_1, x_2, x_3, x_4 \in I$*, then, since* $I \cong \mathbb{Z}^+ \times \mathbb{Z}^+$ *which is a refinement monoid, we are done.*

  So it remains to deal with the case where not all of x_1, x_2, x_3, x_4 *are in* I*. We will assume, without loss of generality, that* $x_1 \notin I$*. This, of course, implies that either* x_3 *or* x_4 *is not in* I*. By choosing a suitably large* $N \in \mathbb{N}$*, we can write*

  $$x_i = l_i a_N + m_i c + n_i d \qquad i = 1, 2, 3, 4$$

  with $l_i, m_i, n_i \in \mathbb{Z}^+$*. Since* $x_1 \notin I$*, we have* $l_1 \geq 1$*.*

Applying the homomorphism α to the equation $x_1 + x_2 = x_3 + x_4$ we get

$$\begin{aligned} l_1 + l_2 &= l_3 + l_4 \\ N(l_1 + l_2) - (m_1 + m_2 + n_1 + n_2) &= N(l_3 + l_4) - (m_3 + m_4 + n_3 + n_4) \\ n_1 + n_2 &\equiv n_3 + n_4 \pmod 2 \end{aligned}$$

Set $\Delta = n_3 + n_4 - (n_1 + n_2)$. Then from the above equations we get $\Delta \equiv 0 \pmod 2$, that is, Δ is an even integer, and $\Delta = m_1 + m_2 - (m_3 + m_4)$. Without loss of generality, we will assume $\Delta \geq 0$. (If $\Delta < 0$ we can interchange the variables x_1, x_2 with x_3, x_4 and start over again.)

If $\Delta = 0$ then we have $l_1 + l_2 = l_3 + l_4$, $m_1 + m_2 = m_3 + m_4$ and $n_1 + n_2 = n_3 + n_4$. Refinements of these three equations in $\mathbb{Z}^+$ then provide coefficients of a_N, c, d for a refinement of the original equation, $x_1 + x_2 = x_3 + x_4$.

If $\Delta > 0$ then we will show that by replacing N by $N + 2$ we can reduce Δ by 2. Repetition of this process eventually gives $\Delta = 0$.

We make new expressions for x_1, x_2, x_3, x_4, using the rule $a_N = a_{N+2} + 2c = a_{N+2} + 2d \ldots$

Since $l_1 \geq 1$ we can write

$$\begin{aligned} x_1 &= (a_{N+2} + 2d) + (l_1 - 1)(a_{N+2} + 2c) + m_1 c + n_1 d \\ &= l_1 a_{N+2} + (m_1 + 2l_1 - 2)c + (n_1 + 2)d, \end{aligned}$$

and for $i = 2, 3, 4$ we write

$$\begin{aligned} x_i &= l_i(a_{N+2} + 2c) + m_i c + n_i d \\ &= l_i a_{N+2} + (2l_i + m_i)c + n_i d. \end{aligned}$$

A simple check shows that the Δ for these new expressions is $n_3 + n_4 - (n_1 + n_2) - 2$ as promised.

To close this section we combine the midseparativity of semi-Artinian refinement monoids with 14.29 and 14.33 to produce a proposition which is in a useful form for application to the monoids $M(R\text{-}\mathbf{Noeth})$ and $M(R\text{-}\mathbf{Art})$ in Section 17.

Corollary 15.10. *Let M and N be refinement monoids such that $N = \operatorname{Lrad} N$ and N has no proper regular elements. Let $\sigma\colon M \to N$ be an exact monoid homomorphism such that $\sigma^{-1}(\{\leq 0\}) = \{\leq 0\}$. Then*

1. *$M = \operatorname{Lrad} M$, in particular, M is semi-Artinian.*
2. *M is strongly separative.*
3. *For all $\alpha \in \mathbf{Ord}$, $\operatorname{soc}_\alpha M = \sigma^{-1}(\operatorname{soc}_\alpha N)$.*
4. *If, in addition, σ is surjective, then for all $\alpha \in \mathbf{Ord}$, $\operatorname{soc}_\alpha N = \sigma(\operatorname{soc}_\alpha M)$.*

Proof. N is a semi-Artinian refinement monoid with no proper regular elements, so by 15.3 and 9.6, it is strongly separative. Thus from 14.33, $\sigma^{-1}(\operatorname{soc}_\alpha N) \subseteq \operatorname{soc}_\alpha M$ for all $\alpha \in \mathbf{Ord}$. In particular, $M = \sigma^{-1}(N) = \sigma^{-1}(\operatorname{Lrad} N) \subseteq \operatorname{Lrad} M$, that is, $M = \operatorname{Lrad} M$.

Let $e \in M$ be regular, then $\sigma(e)$ is regular in N. But N has no proper regular elements, so we must have $\sigma(e) \leq 0$, and because $\sigma^{-1}(\{\leq 0\}) = \{\leq 0\}$, $e \leq 0$. Thus M is a semi-Artinian refinement monoid with no proper regular elements. As above, this implies M is strongly separative.

Since σ is exact, we get from 14.29, $\sigma(\operatorname{soc}_\alpha(M)) \subseteq \operatorname{soc}_\alpha N$ for all $\alpha \in \mathbf{Ord}$. Hence

$$\operatorname{soc}_\alpha M \subseteq \sigma^{-1}(\sigma(\operatorname{soc}_\alpha M)) \subseteq \sigma^{-1}(\operatorname{soc}_\alpha N) \subseteq \operatorname{soc}_\alpha M,$$

that is, $\operatorname{soc}_\alpha M = \sigma^{-1}(\operatorname{soc}_\alpha N)$.

If, in addition, σ is surjective, then we get

$$\operatorname{soc}_\alpha N = \sigma(\sigma^{-1}(\operatorname{soc}_\alpha N)) = \sigma(\operatorname{soc}_\alpha M) \subseteq \operatorname{soc}_\alpha N,$$

so $\operatorname{soc}_\alpha N = \sigma(\operatorname{soc}_\alpha M)$. $\square$

16 Monoids from Modules

The purpose of the current section is to construct monoids which will encode the properties of certain subcategories of R-**Mod** with respect to short exact sequences:

Definition 16.1. *A* **Serre subcategory of** R-**Mod***, is a full subcategory* $\mathbf{S}$ *of* R-**Mod** *such that for every short exact sequence* $0 \to A \to B \to C \to 0$ *in* R-**Mod***,* $B \in \mathbf{S}$ *if and only if* $A, C \in \mathbf{S}$.

In particular, a Serre subcategory is closed under taking submodules, factor modules and finite direct sums. The zero module is an object in any Serre subcategory.

Some standard Serre subcategories of R-**Mod** are

- R-**Noeth**, the full subcategory of R-**Mod** consisting of all Noetherian R-modules.
- R-**Art**, the full subcategory of R-**Mod** consisting of all Artinian R-modules.
- R-**len**, the full subcategory of R-**Mod** consisting of all R-modules of finite length.

For each Serre subcategory $\mathbf{S}$ of R-**Mod** we will construct a monoid $M(\mathbf{S})$ whose elements are equivalence classes of modules:

Definition 16.2. *A* **submodule series** *for a module A is a finite sequence of submodules of the form* $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$. *The* **factors** *of this series are the modules* A_i/A_{i-1} *for* $i = 1, 2, \ldots, n$. *A* **refinement** *of this series is a another submodule series* $0 = A'_0 \leq A'_1 \leq \cdots \leq A'_m = A$ *which contains all the* A_i*, that is,* $A_i \in \{A'_0, A'_1, \ldots, A'_m\}$ *for each* i.

Let A and B be R-modules. Then two submodule series $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ *and* $0 = B_0 \leq B_1 \leq \cdots \leq B_m = B$ *are* **isomorphic** *if* $n = m$ *and there is a permutation of the indices,* σ*, such that* $A_i/A_{i-1} \cong B_{\sigma(i)}/B_{\sigma(i)-1}$ *for* $i = 1, 2, \ldots, n$*. In this situation we will say A and B have* **isomorphic submodule series** *and write* $A \sim B$.

It is clear that isomorphism of submodule series is an equivalence relation, and that if $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ and $0 = B_0 \leq B_1 \leq \cdots \leq B_n = B$ are two isomorphic submodule series then any refinement of one of these series induces an isomorphic refinement of the other series.

If $\mathbf{S}$ is a Serre subcategory of R-**Mod** and $A \in \mathbf{S}$ then the factors of any submodule series for A are also in $\mathbf{S}$. So, in particular, if $B \in R$-**Mod** with $B \sim A$ then $B \in \mathbf{S}$. Thus $\mathbf{S}$ is a union of $\sim$-equivalence classes.

The most important property of submodule series is the Schreier refinement theorem which says that any two submodule series in a module have isomorphic refinements. This is exactly what is needed to make $\sim$ an equivalence relation:

Proposition 16.3. $\sim$ *is an equivalence relation on* R-**Mod**.

Proof. Reflexivity and symmetry are trivial, so it remains to check transitivity...

Suppose $A \sim B$ and $B \sim C$. From the first relation we get isomorphic submodule series $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ and $0 = B_0 \leq B_1 \leq \cdots \leq B_n = B$. From the second relation we get isomorphic submodule series $0 = B'_0 \leq B'_1 \leq \cdots \leq B'_m = B$ and

$0 = C_0 \leq C_1 \leq \cdots \leq C_m = C$. From the Schreier refinement theorem, the two series in B have isomorphic refinements. These new isomorphic submodule series in B induce isomorphic refinements in A and C. Hence $A \sim C$. $\square$

Lemma 16.4. *If $A, B, C \in R$-**Mod** then $A \sim B \implies A \oplus C \sim B \oplus C$.*

Proof. Let $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ and $0 = B_0 \leq B_1 \leq \cdots \leq B_n = B$ be isomorphic submodule series, then it is easily checked that $0 \leq A_0 \oplus C \leq A_1 \oplus C \leq \cdots \leq A_n \oplus C = A \oplus C$ and $0 \leq B_0 \oplus C \leq B_1 \oplus C \leq \cdots \leq B_n \oplus C = B \oplus C$ are isomorphic submodule series in $A \oplus C$ and $B \oplus C$. $\square$

This lemma has the immediate consequence that if $A \sim B$ and $C \sim D$ then $A \oplus C \sim B \oplus D$. That is, $\oplus$ induces a well defined operation on the $\sim$-equivalence classes. We formalize this in the following definition:

Definition 16.5. *Let $\mathbf{S}$ be a Serre subcategory of R-**Mod**. We will write $M(\mathbf{S})$ for $\mathbf{S}/\sim$, the class of $\sim$-equivalence classes of $\mathbf{S}$. We will write $[A] \in M(\mathbf{S})$ for the $\sim$-equivalence class containing $A \in \mathbf{S}$. Define the operation $+$ on $M(\mathbf{S})$ by $[A] + [B] = [A \oplus B]$ for all $A, B \in \mathbf{S}$.*

$(M(\mathbf{S}), +)$ is, in fact, a commutative monoid (and, by 16.10, a refinement monoid). Rather than proving this directly we will use the following more general and useful proposition.

Proposition 16.6. *Let $\mathbf{S}$ be a Serre subcategory of R-**Mod**, N a class with a binary operation $+$, and $\Lambda\colon \mathbf{S} \to N$, a function. Then the following properties of Λ are equivalent:*

(i) *$\Lambda(B) = \Lambda(A) + \Lambda(C)$ whenever $0 \to A \to B \to C \to 0$ is a short exact sequence in $\mathbf{S}$.*

(ii) *$\Lambda(A) = \Lambda(B)$ for any $A, B \in \mathbf{S}$ with $A \sim B$, and $\Lambda(A \oplus B) = \Lambda(A) + \Lambda(B)$ for all $A, B \in \mathbf{S}$.*

If either property is true, then $\Lambda(\mathbf{S})$ is a commutative monoid with identity element $\Lambda(0)$. Also, for $A \in \mathbf{S}$, we have $\Lambda(A) = \Lambda(A_1) + \Lambda(A_2) + \cdots + \Lambda(A_n)$ where $A_1, A_2, \ldots A_n$ are the successive factors of any submodule series for A.

Proof. We show first that (i) implies (ii), and at the same time we prove the other claims of the proposition:

1. For any $A, B \in \mathbf{S}$, the obvious exact sequence $0 \to A \to A \oplus B \to B \to 0$ implies $\Lambda(A \oplus B) = \Lambda(A) + \Lambda(B)$.
2. Let $A \in \mathbf{S}$, then the exact sequences $0 \to A \xrightarrow{\mathrm{id}} A \to 0 \to 0$ and $0 \to 0 \to A \xrightarrow{\mathrm{id}} A \to 0$ imply that $\Lambda(A) + \Lambda(0) = \Lambda(A) = \Lambda(0) + \Lambda(A)$. Thus $\Lambda(0)$ is an identity of $\Lambda(\mathbf{S})$.
3. Suppose $\sigma\colon A \to B$ is an isomorphism with $A, B \in \mathbf{S}$, then $0 \to A \xrightarrow{\sigma} B \to 0 \to 0$ is an exact sequence and so $\Lambda(B) = \Lambda(A) + \Lambda(0) = \Lambda(A)$. So we have shown that $A \cong B$ implies $\Lambda(A) = \Lambda(B)$.
4. The commutativity and associativity of the operation $+$ on $\Lambda(\mathbf{S})$ come directly from these same properties of $\oplus$ up to isomorphism. With 2, we have proved that $\Lambda(\mathbf{S})$ is a commutative monoid with identity $\Lambda(0)$.
5. Suppose $0 = A'_0 \leq A'_1 \leq \cdots \leq A'_n = A$ is a submodule series for $A \in \mathbf{S}$ with factors $A_i = A'_i/A'_{i-1}$. All A'_i and A_i are in $\mathbf{S}$. For each i we have the exact sequence

$0 \to A'_{i-1} \to A'_i \to A_i \to 0$, so $\Lambda(A'_i) = \Lambda(A_i) + \Lambda(A'_{i-1})$. A simple induction then shows that $\Lambda(A) = \Lambda(A_1) + \Lambda(A_2) + \cdots + \Lambda(A_n)$.

6. If $A, B \in \mathbf{S}$ have isomorphic submodule series, that is $A \sim B$, then using 3, 4 and 5, we get $\Lambda(A) = \Lambda(B)$.

To show that (ii) implies (i), suppose $0 \to A \xrightarrow{\sigma} B \to C \to 0$ is exact for some $A, B, C \in \mathbf{S}$. Then $C \cong B/\operatorname{im}(\sigma)$ with $\operatorname{im}(\sigma) \cong A$, so B has the submodule series $0 \le \operatorname{im}(\sigma) \le B$ with factors isomorphic to A and C. The module $A \oplus C$ has the submodule series $0 \le A \oplus 0 \le A \oplus C$ with these same factors, so $A \oplus C \sim B$. By (ii), $\Lambda(B) = \Lambda(A \oplus C) = \Lambda(A) + \Lambda(C)$.

□

Any function Λ which satisfies either of the conditions of this proposition will be said to **respect short exact sequences in S**.

Since the map $A \mapsto [A]$ from $\mathbf{S}$ to $M(\mathbf{S})$ satisfies condition (ii) and is surjective, $M(\mathbf{S})$ is a commutative monoid with identity $[0]$. We will use the notation $\overline{M}(\mathbf{S})$ and $\widetilde{M}(\mathbf{S})$ for the universal monoids constructed from $M(\mathbf{S})$ according to 6.3 and 6.18, respectively.

The monoid $(M(\mathbf{S}), +)$ has the following universal property:

Proposition 16.7. *Let $\mathbf{S}$ be a Serre subcategory of R-$\mathbf{Mod}$, N a class with a binary operation $+$, and $\Lambda\colon \mathbf{S} \to N$, a map which respects short exact sequences in $\mathbf{S}$. Then Λ factors uniquely through $M(\mathbf{S})$. Specifically, there exists a unique monoid homomorphism λ from $M(\mathbf{S})$ to $\Lambda(\mathbf{S})$ such that the following diagram commutes:*

$$\begin{tikzcd} \mathbf{S} \arrow[r, "{[\ ]}"] \arrow[dr, "\Lambda"'] & M(\mathbf{S}) \arrow[d, "\lambda"] \\ & \Lambda(\mathbf{S}) \end{tikzcd}$$

Proof. Define the map $\lambda\colon M(\mathbf{S}) \to \Lambda(\mathbf{S})$ by $\lambda([A]) = \Lambda(A)$ for all $A \in \mathbf{S}$. This is well defined because if $[A] = [B]$, then $A \sim B$ and, by 16.6, $\Lambda(A) = \Lambda(B)$. For any $[A], [B] \in M$, we have $\lambda([A]+[B]) = \lambda([A \oplus B]) = \Lambda(A \oplus B) = \Lambda(A) + \Lambda(B) = \lambda([A]) + \lambda([B])$. Also, $\lambda([0]) = \Lambda(0)$ which is the identity for $\Lambda(\mathbf{S})$. So λ is a monoid homomorphism. □

We note that, in this proposition, if N happened to be a monoid, the homomorphism λ would not be a monoid homomorphism when viewed as a map to N unless, in addition, $\Lambda(0) = 0$. This will indeed be the case in all the applications of the proposition we will make.

Proposition 16.7 provides a second characterization of the equivalence relation $\sim$ for modules $A, B \in R$-$\mathbf{Mod}$, namely, $A \sim B$ if and only if the modules A and B are indistinguishable by functions on R-$\mathbf{Mod}$ which respect short exact sequences.

By construction, $M(\mathbf{S})$ is a submonoid of $M(R\text{-}\mathbf{Mod})$ for any Serre subcategory $\mathbf{S}$. In fact, we will see that $M(\mathbf{S})$ is not just a submonoid but also an order ideal of $M(R\text{-}\mathbf{Mod})$, and further that every order ideal of $M(R\text{-}\mathbf{Mod})$ is $M(\mathbf{S})$ for some Serre subcategory $\mathbf{S}$:

Let $\mathbf{S}$ be a Serre subcategory of R-$\mathbf{Mod}$ and define a map Λ: R-$\mathbf{Mod} \to \{0, \infty\}$ by

$$\Lambda(A) = \begin{cases} 0 & A \in \mathbf{S} \\ \infty & A \notin \mathbf{S} \end{cases}$$

It is easy to check that Λ respects short exact sequences so that there is an induced monoid homomorphism λ: $M(R\text{-}\mathbf{Mod}) \to \{0, \infty\}$ such that $\lambda([A]) = 0$ if and only if $A \in \mathbf{S}$. From 6.12, $\lambda^{-1}(0) = M(\mathbf{S})$ is then an order ideal in $M(R\text{-}\mathbf{Mod})$.

Conversely, given an order ideal I of $M(R\text{-}\mathbf{Mod})$, it is easy to show that

$$\mathbf{S} = \{A \in R\text{-}\mathbf{Mod} \mid [A] \in I\}$$

is a Serre subcategory of R-$\mathbf{Mod}$.

These two constructions can be used to prove:

Proposition 16.8. *The map $\mathbf{S} \mapsto M(\mathbf{S})$ is a bijection from the class of all Serre subcategories of R-$\mathbf{Mod}$ to the class of all order ideals of $M(R\text{-}\mathbf{Mod})$.*

Proof. It remains only to check that the maps described above are inverses of each other. □

An immediate consequence of this proposition is that if Λ: $\mathbf{S} \to N$ respects short exact sequences, then the inverse image of any order ideal in the monoid $\Lambda(\mathbf{S})$ is a Serre subcategory in $\mathbf{S}$.

We collect in the next proposition some simple properties of $M(R\text{-}\mathbf{Mod})$:

Proposition 16.9. *Let $A, B \in R$-$\mathbf{Mod}$.*

1. *If A is a submodule of B, then*

$$[B] = [A] + [B/A].$$

2. *If A is a submodule, factor module or subfactor module of B, then $[A] \leq [B]$.*
3. *$[A] \leq [B]$ if and only if there are submodule series $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ and $0 = B_0 \leq B_1 \leq \cdots \leq B_m = B$ and an injection, σ: $\{1, 2, \ldots, n\} \to \{1, 2, \ldots, m\}$, such that $A_i/A_{i-1} \cong B_{\sigma(i)}/B_{\sigma(i)-1}$ for $i = 1, 2, \ldots, n$.*
4. *$[A] \prec [B]$ if and only if there is a submodule series $0 = A_0 \leq A_1 \leq \cdots \leq A_n = A$ such that A_i/A_{i-1} is isomorphic to a subfactor of B for $i = 1, 2, \ldots, n$.*

Proof.

1. Follows from the existence of the obvious exact sequence $0 \to A \to B \to B/A \to 0$.
2. Immediate from 1.
3. If $[A] \leq [B]$, then there is some module C such that $[B] = [A] + [C] = [A \oplus C]$, that is B and $A \oplus C$ have isomorphic submodule series. The submodules series in $A \oplus C$ can be chosen as a refinement of the series $0 \leq A \oplus 0 \leq A \oplus C$. The permutation of factors of the two submodule series restricted to the factors in A gives the map σ as required.

   The converse is easy.
4. Proved in a similar way to 3.

□

We have, by definition, that a short exact sequence $0 \to A \to B \to C \to 0$ in a Serre subcategory $\mathbf{S}$ gives rise to the equation $[B] = [A] + [C]$ in $M(\mathbf{S})$. For an exact sequence with four terms,

$$0 \to A \to B \xrightarrow{\beta} C \xrightarrow{\gamma} D \to 0,$$

we get the equation $[A] + [C] = [B] + [D]$. This is proved by making the two short exact sequences

$$0 \to A \to B \xrightarrow{\beta} \operatorname{im}\beta \to 0,$$

and

$$0 \to \ker\gamma \to C \xrightarrow{\gamma} D \to 0.$$

Since $\operatorname{im}\beta = \ker\gamma$, we get $[A]+[C] = [A]+[\ker(\gamma)]+[D] = [A]+[\operatorname{im}(\beta)]+[D] = [B]+[D]$. Induction using this process gives the general rule that if

$$0 \to A_1 \to B_1 \to A_2 \to B_2 \to \ldots \to A_n \to B_n \to 0$$

is an exact sequence in $\mathbf{S}$, then $[A_1] + [A_2] + \ldots + [A_n] = [B_1] + [B_2] + \ldots + [B_n]$.

The most important algebraic property of the monoid $M(R\text{-}\mathbf{Mod})$ is that it has refinement...

Suppose a submodule series is given for a module A and there is another module B such that $A \sim B$, then the Schreier refinement theorem implies that there is a refinement of the existing series in A which is isomorphic to a submodule series in B. If B also happened to have a submodule series given, then a second application of the theorem would give refinements of the two given series which are isomorphic. This principle is used in showing that $M(\mathbf{S})$ is a refinement monoid:

Proposition 16.10. *$M(\mathbf{S})$ is a refinement monoid for any Serre subcategory $\mathbf{S}$ of R-$\mathbf{Mod}$.*

Proof. Suppose there are modules $A, B, C, D \in \mathbf{S}$ such that $[A] + [B] = [C] + [D]$ in $M(\mathbf{S})$. Then $[A \oplus B] = [C \oplus D]$, that is, $A \oplus B \sim C \oplus D$. From the above discussion, there are isomorphic submodule series for these two modules which are refinements of the series $0 \le A \oplus 0 \le A \oplus B$ and $0 \le C \oplus 0 \le C \oplus D$. That is, there are submodule series $0 \le A_1 \le \cdots \le A$, $0 \le B_1 \le \cdots \le B$, $0 \le C_1 \le \cdots \le C$, and $0 \le D_1 \le \cdots \le D$ such that the series

$$0 \le A_1 \oplus 0 \le \cdots \le A \oplus 0 \le A \oplus B_1 \le \cdots \le A \oplus B$$

and

$$0 \le C_1 \oplus 0 \le \cdots \le C \oplus 0 \le C \oplus D_1 \le \cdots \le C \oplus D$$

are isomorphic.

The permutation that matches isomorphic factors in these submodule series divides them into four types: (1) $A_i/A_{i-1} \cong C_j/C_{j-1}$; (2) $A_i/A_{i-1} \cong D_j/D_{j-1}$; (3) $B_i/B_{i-1} \cong C_j/C_{j-1}$; or (4) $B_i/B_{i-1} \cong D_jD_{j-1}$ for suitable indices i, j. If we let $W, X, Y, Z \in \mathbf{S}$ be the direct sums of the factors of type 1,2,3,4 respectively, then one can easily check that $[W] + [X] = \sum_i [A_i/A_{i-1}] = [A]$ and, similarly, $[W] + [Y] = [C]$, $[X] + [Z] = [D]$, $[Y] + [Z] = [B]$, that is, we have the refinement matrix

$$\begin{array}{cc} & \begin{array}{cc} [A] & [B] \end{array} \\ \begin{array}{c} [C] \\ [D] \end{array} & \begin{pmatrix} [W] & [Y] \\ [X] & [Z] \end{pmatrix} \end{array}$$

□

We can now identify the atoms and the socle of the monoid $M(R\text{-}\mathbf{Mod})$:

Proposition 16.11. *Let $A \in R$-**Mod**.*

1. *$[A] \leq 0 = [0]$ if and only if $A = 0$. In particular, $M(R\text{-}\mathbf{Mod})$ is a conical monoid.*
2. *If $[A] \leq [M]$ with M a simple module, then either $A = 0$ or $A \cong M$.*
3. *If $[A] \neq 0$, then there is a simple module $M \in R$-**Mod** such that $[M] \leq [A]$.*
4. *$[A]$ is an atom of $M(R\text{-}\mathbf{Mod})$ if and only if A is a simple module.*
5. $\mathrm{soc}(M(R\text{-}\mathbf{Mod})) = M(R\text{-}\mathbf{len})$
6. $\mathrm{soc}(M(R\text{-}\mathbf{Mod}))$ *is a free monoid with the atoms of $M(R\text{-}\mathbf{Mod})$ as its basis.*

Proof.

1. Direct from the definition.
2. Follows from 16.9.3 and the scarcity of submodule series for M.
3. Since $[A] \neq 0$, the module A is nonzero. Let $Ra \leq A$ be a nonzero cyclic submodule. We have $Ra \cong R/\mathrm{ann}(a)$ with $\mathrm{ann}(a)$ a proper ideal of R. Let I be a maximal (one-sided) ideal of R containing $\mathrm{ann}(a)$, and $M = R/I$. Then M is simple and
$$[M] \leq [M] + [I/\mathrm{ann}(a)] = [R/I] + [I/\mathrm{ann}(a)] = [R/\mathrm{ann}(a)] = [Ra] \leq [A].$$
4. Suppose $[A]$ is an atom of $M(R\text{-}\mathbf{Mod})$. Since $[A] \neq 0$, there is a simple module M such that $0 \neq [M] \leq [A]$. Since $[A]$ is an atom, this implies that $[A] \leq [M]$, and hence, from 2, that $A \cong M$.

   Conversely, if A is simple, then 2 shows immediately that $[A]$ is an atom.
5. If A has finite length, then it is the zero module or it has a composition series whose factors are all simple modules. In the first case $[A] = 0 \in \mathrm{soc}(M(R\text{-}\mathbf{Mod}))$. In the second case, from 4, we get that $[A]$ is a sum of atoms of $M(R\text{-}\mathbf{Mod})$ and hence is in $\mathrm{soc}(M(R\text{-}\mathbf{Mod}))$.

   Conversely, if $a \in \mathrm{soc}(M(R\text{-}\mathbf{Mod}))$ then by 14.23, a is zero or a sum of atoms. (Here we need that $M(R\text{-}\mathbf{Mod})$ has refinement.) If $a = 0$ then $a = [0] \in M(R\text{-}\mathbf{len})$. Otherwise, using 4, there are simple modules $A_1, A_2, \ldots A_n$, such that
$$a = [A_1] + [A_2] + \ldots + [A_n] = [A_1 \oplus A_2 \oplus \ldots \oplus A_n].$$
   Since $A_1 \oplus A_2 \oplus \ldots \oplus A_n$ has finite length, we have $a \in M(R\text{-}\mathbf{len})$.
6. If $[A] \in \mathrm{soc}(M(R\text{-}\mathbf{Mod}))$ is nonzero, then A has finite length and so
$$[A] = [A_1] + [A_2] + \ldots + [A_n]$$
   where $A_1, A_2, \ldots, A_n$ are the simple factor modules of a composition series for A. By the Jordan-Hölder Theorem [1, 11.3] these simple modules are uniquely determined by A. Thus, with 4, $[A]$ is uniquely a sum of atoms of $M(R\text{-}\mathbf{Mod})$.

   Let B be the class of all atoms of $M(R\text{-}\mathbf{Mod})$ and $\Lambda\colon B \to M'$ a map from B to an arbitrary monoid M'. From the above discussion, there is a unique way to extend Λ to a monoid homomorphism from $\mathrm{soc}(M(R\text{-}\mathbf{Mod}))$ to M'. Thus from Definition 5.13, $\mathrm{soc}(M(R\text{-}\mathbf{Mod}))$ is a free monoid with basis B.

□

For finite length modules, we have the composition series length function len from R-**len** to $\mathbb{Z}^+$. If $0 \to A \to B \to C \to 0$ is a short exact sequence in R-**len**, then $\mathrm{len}\, B = \mathrm{len}\, A + \mathrm{len}\, C$. Thus len respects short exact sequences as a map to the monoid $(\mathbb{Z}^+, +)$. From 16.7, there is an induced monoid homomorphism, which we will also call len, from

$M(R\text{-}\mathbf{len})$ to $\mathbb{Z}^+$ such that $\operatorname{len}[A] = \operatorname{len} A$ for all $A \in R\text{-}\mathbf{len}$. So, for example, $[A] \le [B]$ implies $\operatorname{len}[A] \le \operatorname{len}[B]$ implies $\operatorname{len} A \le \operatorname{len} B$ for all $A, B \in R\text{-}\mathbf{len}$. In Section 17, this map will be extended to all of R-**Noeth** using the Krull length function.

Having established the main properties of $M(R\text{-}\mathbf{Noeth})$, we calculate this monoid in the simplest case:

Example 16.12. *Suppose R is a field. Then, up to isomorphism, any R-module A is determined by the cardinality of a basis, that is, by its dimension,* $\dim A$*. Further, for any short exact sequence, $0 \to A \to B \to C \to 0$ we have $\dim B = \dim A + \dim C$. Here the operation is cardinal addition. Thus the map* $\dim$ *respects short exact sequences as a map from R-**Mod** to $(\mathbf{Card}, +)$. By 16.7, there is a monoid homomorphism from $M(R\text{-}\mathbf{Mod})$ to **Card**. This homomorphism must be injective since elements of $M(R\text{-}\mathbf{Mod})$ are unions of isomorphism classes of modules and each isomorphism class maps to a different cardinal via* $\dim$*. The homomorphism is surjective because for every cardinal $\mathfrak{a}$, the free module $R^{\mathfrak{a}}$ has $\dim R^{\mathfrak{a}} = \mathfrak{a}$. Thus $M(R\text{-}\mathbf{Mod})$ is isomorphic to **Card**.*

Note that, in this example, $M(R\text{-}\mathbf{Mod})$ is not a set. This is, in fact, always the case:

Proposition 16.13. *For any non-trivial ring R, $M(R\text{-}\mathbf{Mod})$ is a proper class.*

Proof. Let $(\mathbf{Card}, \cdot)$ be the commutative monoid whose elements are the cardinal numbers and whose operation is cardinal multiplication [25, Section 5.4]. Define a map Φ from R-**Mod** to $(\mathbf{Card}, \cdot)$ by $\Phi(A) = |A|$ for all $A \in R\text{-}\mathbf{Mod}$, that is, $\Phi(A)$ is the cardinality of A as a set. If $0 \to A \to B \to C \to 0$ is a short exact sequence in R-**Mod**, then, as sets, we have $B \cong A \times C$. Thus Φ respects short exact sequences. Let $\phi\colon M(R\text{-}\mathbf{Mod}) \to (\mathbf{Card}, \cdot)$ be the monoid homomorphism induced from Φ.

We proceed by contradiction...

Suppose $M(R\text{-}\mathbf{Mod})$ is a set. Then $\phi(M(R\text{-}\mathbf{Mod}))$ is a set of cardinal numbers, so must have an upper bound $\mathfrak{a} \in \mathbf{Card}$ [25, 5.2.7]. The free module $R^{\mathfrak{a}}$ has cardinality $|R|^{\mathfrak{a}}$, and since R has at least two distinct elements, we get [25, 5.4.2p]

$$\mathfrak{a} < 2^{\mathfrak{a}} \le |R|^{\mathfrak{a}} = \phi([R^{\mathfrak{a}}]).$$

Thus $\mathfrak{a}$ can not be an upper bound for $\phi(M(R\text{-}\mathbf{Mod}))$. □

As partial compensation for this proposition we have

Proposition 16.14. *For any ring R and $a \in M(R\text{-}\mathbf{Mod})$, the order ideal $\{\prec a\}$ is a set.*

Proof. Let $a = [A]$ for some $A \in R\text{-}\mathbf{Mod}$. Since A is a set, so are the classes of all submodules, factor modules, and subfactor modules of A. By 16.9, $\{\prec [A]\}$ consists of all elements of the form $[A_1] + [A_2] + \ldots + [A_n] \in M(R\text{-}\mathbf{Mod})$ where A_i is a subfactor of A for $i = 1, 2, \ldots, n$. Since the class of finite sets of subfactor modules of A is also a set, this implies that $\{\prec [A]\}$ is a set. □

Corollary 16.15. $\operatorname{Lrad} M(R\text{-}\mathbf{Mod}) = \operatorname{srad} M(R\text{-}\mathbf{Mod})$

Proof. This follows from 14.27. □

Another feature of Example 16.12 is that each $\sim$-equivalence class is an isomorphism class of modules. That is, $M(R\text{-}\mathbf{Mod})$ contains as much information as the classification of the R-modules up to isomorphism given by the dimension function. To investigate when this situation happens we define another monoid based on the category R-**Mod**:

Definition 16.16. *For any module $A \in R\text{-}\mathbf{Mod}$, let $\{\cong A\}$ be its isomorphism class. Let $V(R\text{-}\mathbf{Mod})$ be the monoid whose elements are the isomorphism classes of R-**Mod** with operation $+$ induced by the direct sum, that is,*

$$\{\cong A\} + \{\cong B\} = \{\cong A \oplus B\}$$

for all $A, B \in R\text{-}\mathbf{Mod}$.

It is easy to see that $V(R\text{-}\mathbf{Mod})$ is a well defined commutative monoid. Further, there is a well defined monoid homomorphism $\tau\colon V(R\text{-}\mathbf{Mod}) \to M(R\text{-}\mathbf{Mod})$ given by $\tau(\{\cong A\}) = [A]$ for $A \in R\text{-}\mathbf{Mod}$.

Proposition 16.17. *For a ring R, the following are equivalent:*

1. *R is semisimple*
2. *$\tau\colon V(R\text{-}\mathbf{Mod}) \to M(R\text{-}\mathbf{Mod})$ is a monoid isomorphism*
3. *For $A, B \in R\text{-}\mathbf{Mod}$, $A \cong B$ if and only if $A \sim B$*

Proof. See [26, Theorem 4.13] for the relevant properties of semisimple rings.

$1 \Rightarrow 2$ Let $0 \to A \to B \to C \to 0$ be an exact sequence in R-**Mod**. Since R is semisimple, this sequence splits to give $B \cong A \oplus C$, and hence $\{\cong B\} = \{\cong A\} + \{\cong C\}$. Thus the map which takes a module A to its isomorphism class $\{\cong A\}$ respects short exact sequences. By the universal property of $M(R\text{-}\mathbf{Mod})$, there is a monoid homomorphism $\sigma\colon M(R\text{-}\mathbf{Mod}) \to V(R\text{-}\mathbf{Mod})$ such that $\sigma([A]) = \{\cong A\}$ for all $A \in R\text{-}\mathbf{Mod}$. Clearly σ is the inverse of τ, so τ is an isomorphism.

$2 \Rightarrow 3$ Since τ is injective, $[A] = [B]$ if and only if $\{\cong A\} = \{\cong B\}$.

$3 \Rightarrow 1$ We show that every R-module is projective...

Let $A \in R\text{-}\mathbf{Mod}$. Then there is an exact sequence $0 \to K \to F \to A \to 0$ in R-**Mod** such that F is a free module and $K \leq F$. Thus $[F] = [A] + [K] = [A \oplus K]$, and $F \sim A \oplus K$. By 3, $F \cong A \oplus K$, that is, A is a direct summand of a free module. This implies that A is projective.

□

17 Noetherian and Artinian Modules

In this section we will investigate the cancellation properties of the monoid $M(R\text{-}\mathbf{Noeth})$. Since modules in R-**Noeth** satisfy the descending chain condition, one might expect that there is some corresponding chain condition appearing in $M(R\text{-}\mathbf{Noeth})$. This is indeed true. For example, in Section 19, we will see that for FBN rings and, in particular, for commutative Noetherian rings, $M(R\text{-}\mathbf{Noeth})$ is Artinian.

Our goal in this section is to show that, though $M(R\text{-}\mathbf{Noeth})$ may not be Artinian for all rings, it is always semi-Artinian. Since $M(R\text{-}\mathbf{Noeth})$ has no proper regular elements, this then implies that $M(R\text{-}\mathbf{Noeth})$ is strongly separative.

We start by providing an example that shows that $M(R\text{-}\mathbf{Noeth})$ may not be Artinian:

Example 17.1. [9] *A Noetherian R-module S such that $[S] \notin \mathrm{Arad}(M(R\text{-}\mathbf{Noeth}))$:*

Let F be a field of characteristic zero, and $S = F[[t]]$, the formal power series ring over F. For each $n \in \mathbb{N}$, define the ideal $S_n = St^n$. Every nonzero ideal of S is S_n for some $n \in \mathbb{N}$, and these ideals form a descending chain $S = S_0 > S_1 > S_2 > \dots$ of S-submodules of S.

Define a derivation δ on S according to the rule

$$\delta(s) = t\frac{d}{dt}(s)$$

for all $s \in S$. Let $R = S[\theta; \delta]$, the skew polynomial ring over S with multiplication defined so that

$$\theta s = s\theta + \delta(s)$$

for all $s \in S$. This ring is Noetherian [11, Theorem 1.12]*, as is the left R-module $R/R\theta$. To discuss the properties of $R/R\theta$ it is convenient to make S into a left R-module isomorphic to $R/R\theta$:*

Since $R/R\theta$ is already isomorphic to S as an S-module, we need only define the action of θ on elements of s. In $R/R\theta$ we have $\theta(s + R\theta) = \delta(s) + R\theta$, for all $s \in S$. Accordingly, we define a new module multiplication $\cdot$ on S by

$$s' \cdot s = s's$$

and

$$\theta \cdot s = \delta(s)$$

for all $s, s' \in S$. In particular, $\theta \cdot t^m = \delta(t^m) = mt^m$ for all $m \in \mathbb{N}$.

It is easily checked that S_n is an R-module for all $n \in \mathbb{N}$. Since any R-submodule of S must also be an S-submodule, $S_0, S_1, \dots$ are all the R-submodules of S. Each of these submodules has infinite length, whereas for any $n \leq k$, $\mathrm{len}(S_n/S_k) = k - n$.

- **Claim** *S_m and S_n are isomorphic as R-modules if and only if $m = n$.*

  Without loss of generality we assume $m \leq n$. Let $\phi\colon S_m \to S_n$ be an R-module isomorphism. Since $S_n = St^n$, there is some $u \in S$ such that $\phi(t^m) = ut^n$. Set

$v = ut^{n-m}$ so that $\phi(t^m) = vt^m$. Since ϕ is an R-module homomorphism, we have $\phi(\theta \cdot t^m) = \theta \cdot \phi(t^m)$, and so,

$$\begin{aligned} mvt^m &= \phi(mt^m) = \phi(\theta \cdot t^m) = \theta \cdot \phi(t^m) \\ &= \theta \cdot vt^m = \delta(vt^m) = \delta(v)t^m + v\delta(t^m) \\ &= \delta(v)t^m + mvt^m. \end{aligned}$$

Thus $\delta(v)t^m = 0$, and, since S is a domain, $\delta(v) = 0$. From the definition of δ, it is easy to see that $\delta(v) = 0$ implies $v \in F$. Since ϕ is an isomorphism, $v = 0$ is not possible, so we have $v = ut^{n-m} \in F^*$. This implies $m = n$.

Consider the module S_n for some $n \in \mathbb{N}$. From the above discussion it is clear that any submodule series for S_n has exactly one factor which has infinite length, and this factor is S_k for some $k \geq n$. All other factors in the series have finite length.

- **Claim** $[S_m] \leq [S_n]$ *in $M(R$-**Noeth**$)$ if and only if $m \geq n$.*

  Suppose $[S_m] \leq [S_n]$. Then, from 16.9.3, S_m and S_n have submodule series such that every factor in the series for S_m is isomorphic to a factor in the series for S_n. Both of these series have exactly one factor with infinite length. From the previous claim, these infinite factors must coincide in both series: S_k, say, for some $k \geq m, n$. The remaining factors in the two series can then be used to show that $[S_m/S_k] \leq [S_n/S_k]$. But S_m/S_k and S_n/S_k are finite length modules, so $\operatorname{len}(S_m/S_k) \leq \operatorname{len}(S_n/S_k)$*, that is, $k - m \leq k - n$ and $m \geq n$.*

  The converse is trivial since if $m \geq n$, then S_m is a submodule of S_n.

*From this claim we get immediately that $[S] = [S_0] \geq [S_1] \geq [S_2] \geq \ldots$ is a decreasing sequence in $M(R$-**Noeth**$)$ which has no minimal element. Therefore $[S]$ is not in* $\operatorname{Arad}(M(R\text{-}\mathbf{Noeth}))$*. In particular, $M(R$-**Noeth**$)$ is not an Artinian monoid.*

Having shown that $M(R$-**Noeth**$)$ may not be Artinian, we turn to the positive result that $M(R$-**Noeth**$)$ is semi-Artinian. The main tool for proving this claim is the function $\mathrm{Klen}^\circ \colon R\text{-}\mathbf{Noeth} \to \mathbf{Krull}$ which we defined and discussed in Section 4. Thus it will be necessary to reformulate what we know about Krull length and Krull dimension in terms of Serre categories, monoids and monoid homomorphisms...

Recall from 3.21 that $\mathbf{Krull} = (\mathbf{Ord} \times \mathbb{N}) \cup \{0\}$ with operation $+$ given by

1. $0 + 0 = 0$
2. $0 + (\gamma, n) = (\gamma, n) + 0 = (\gamma, n)$ for all $(\gamma, n) \in \mathbf{Ord} \times \mathbb{N}$
3. $$(\gamma_1, n_1) + (\gamma_2, n_2) = \begin{cases} (\gamma_1, n_1) & \text{if } \gamma_2 < \gamma_1 \\ (\gamma_2, n_2) & \text{if } \gamma_1 < \gamma_2 \\ (\gamma_1, n_1 + n_2) & \text{if } \gamma_1 = \gamma_2 \end{cases}$$

   for all $(\gamma_1, n_1), (\gamma_2, n_2) \in \mathbf{Ord} \times \mathbb{N}$.

We also defined $\mathbf{Ord}^* = \mathbf{Ord} \cup \{-1\}$ and the function $\kappa \colon \mathbf{Krull} \to \mathbf{Ord}^*$ such that $\kappa(\gamma, n) = \gamma$ for $(\gamma, n) \in \mathbf{Ord} \times \mathbb{N}$ and $\kappa(0) = -1$. We will henceforth consider **Krull** and $\mathbf{Ord}^*$ to be monoids with the operations $+$ and max, respectively, and in the next proposition we collect their basic properties.

Proposition 17.2.

1. $(\mathbf{Krull}, +)$ *and* $(\mathbf{Ord}^*, \max)$ *are commutative monoids.* $\mathbf{Ord}^*$ *is a semilattice monoid whose identity element is* -1*, rather than* 0.
2. *The minimum order on* $\mathbf{Krull}$ *is given by*

$$\begin{aligned}(\gamma_1, n_1) \leq (\gamma_2, n_2) &\iff (\gamma_1 < \gamma_2 \text{ or } (\gamma_1 = \gamma_2 \text{ and } n_1 \leq n_2)) \\ &\iff \omega^{\gamma_1} n_1 \leq \omega^{\gamma_2} n_2 \textit{ in } \mathbf{Ord}\end{aligned}$$

for all $(\gamma_1, n_1), (\gamma_2, n_2) \in \mathbf{Ord} \times \mathbb{N}$*. The minimum order on* $\mathbf{Ord}^*$ *is the usual one. Both of these monoids are totally ordered and Artinian.*
3. $\mathbf{Krull}$ *is strongly separative and* $\mathbf{Ord}^*$ *is separative. Both monoids have refinement.*
4. *For all* $\alpha \in \mathbf{Ord}$*,* $\operatorname{soc}_\alpha \mathbf{Ord}^* = \{< \alpha\} = \{\beta \in \mathbf{Ord}^* \mid \beta < \alpha\}$
5. *For all* $\alpha \in \mathbf{Ord}$*,* $\operatorname{soc}_\alpha \mathbf{Krull} = \kappa^{-1}(\{< \alpha\}) = \{(\gamma, n) \mid \gamma < \alpha \text{ and } n \in \mathbb{N}\} \cup \{0\}$.

Proof.

1. Trivial.
2. Trivial.
3. That $\mathbf{Krull}$ is strongly separative is an easy check, and that $\mathbf{Ord}^*$ is separative is trivial. Since these monoids are totally ordered, by 9.4, they also have refinement.
4. This is an easy induction.
5. It is easy to see that the map κ: $\mathbf{Krull} \to \mathbf{Ord}^*$ is exact, so by 14.29,

$$\kappa(\operatorname{soc}_\alpha \mathbf{Krull}) \subseteq \operatorname{soc}_\alpha \mathbf{Ord}^* = \{< \alpha\}$$

for all $\alpha \in \mathbf{Ord}$. That is, $\operatorname{soc}_\alpha \mathbf{Krull} \subseteq \kappa^{-1}(\{< \alpha\})$.

To prove the converse inclusion we use induction:

- $\underline{\alpha = 0}$ Trivial.
- $\underline{\alpha \text{ is a successor ordinal}}$ Suppose $\alpha = \beta + 1$ and $\operatorname{soc}_\beta \mathbf{Krull} = \kappa^{-1}(\{< \beta\})$. Then $(\beta, 1)$ is the minimum element of $\mathbf{Krull} \backslash (\operatorname{soc}_\beta \mathbf{Krull})$. It must therefore map to an atom of $\mathbf{Krull}/(\operatorname{soc}_\beta \mathbf{Krull})$. This implies that $(\beta, 1) \in \operatorname{soc}_\alpha \mathbf{Krull}$. Since $\operatorname{soc}_\alpha \mathbf{Krull}$ is an order ideal, this means that $(\beta, n) \in \operatorname{soc}_\alpha \mathbf{Krull}$ for all $n \in \mathbb{N}$, and so $\kappa^{-1}(\{< \alpha\}) \subseteq \operatorname{soc}_\alpha \mathbf{Krull}$.
- $\underline{\alpha \text{ is a limit ordinal}}$ Trivial.

□

The order in $\mathbf{Krull}$ is such that

$$0 < (0,1) < (0,2) < \ldots < (1,1) < (1,2) < \ldots < (2,1) < \ldots$$

Thus $\operatorname{len}\{\leq (2,1)\} = \omega 2$, $\operatorname{Klen}\{\leq (2,1)\} = (1,2)$ and $\operatorname{Kdim}\{\leq (2,1)\} = 1$. So it is not true, in general, that $\operatorname{Klen}\{\leq (\gamma, n)\} = (\gamma, n)$ or $\operatorname{Kdim}\{\leq (\gamma, n)\} = \gamma$.

Now we consider the Serre category R-$\mathbf{Noeth}$ for some ring R. Of course, what we will prove here applies equally well to R-$\mathbf{Art}$. From 4.1, we have the Krull length function $\operatorname{Klen}^\circ : R\text{-}\mathbf{Noeth} \to \mathbf{Krull}$ such that if

$$0 \to A \to B \to C \to 0$$

is an exact sequence in R-**Noeth**, then

$$\operatorname{Klen}^\circ B = \operatorname{Klen}^\circ A + \operatorname{Klen}^\circ C.$$

That is, Klen° respects exact sequences. From 16.7, there is an induced monoid homomorphism, which we will also call Klen°, from $M(R\text{-}\mathbf{Noeth})$ to **Krull**. Since **Krull** is partially ordered, there is another induced homomorphism, again called Klen°, from $\overline{M}(R\text{-}\mathbf{Noeth})$ to **Krull**.

We also have the monoid homomorphism $\operatorname{Kdim}^\circ = \kappa \circ \operatorname{Klen}^\circ$ from $M(R\text{-}\mathbf{Noeth})$ to the semilattice monoid $\mathbf{Ord}^*$. So, from 6.22, there is an induced monoid homomorphism, also called Kdim°, from $\widetilde{M}(R\text{-}\mathbf{Noeth})$ to $\mathbf{Ord}^*$. Thus we get the following commutative diagram:

$$\begin{tikzcd}
R\text{-}\mathbf{Noeth} \arrow[r, "{[\ ]}"] \arrow[drr, "\operatorname{Klen}^\circ"'] & M(R\text{-}\mathbf{Noeth}) \arrow[r, "\{\equiv\ \}"] \arrow[dr, "\operatorname{Klen}^\circ"] & \overline{M}(R\text{-}\mathbf{Noeth}) \arrow[r, "\{\asymp\ \}"] \arrow[d, "\operatorname{Klen}^\circ"] & \widetilde{M}(R\text{-}\mathbf{Noeth}) \arrow[d, "\operatorname{Kdim}^\circ"] \\
 & & \mathbf{Krull} \arrow[r, "\kappa"] & \mathbf{Ord}^*
\end{tikzcd}$$

For all $A \in R$-**Noeth** we have

$$\operatorname{Klen}^\circ A = \operatorname{Klen}^\circ[A] = \operatorname{Klen}^\circ\{\equiv [A]\} \in \mathbf{Krull}$$

and

$$\operatorname{Kdim}^\circ A = \operatorname{Kdim}^\circ[A] = \operatorname{Kdim}^\circ\{\equiv [A]\} = \operatorname{Kdim}^\circ\{\asymp [A]\} \in \mathbf{Ord}^*$$

where in all but the last case, Kdim° is the composition of the maps Klen° and κ.

Lemma 17.3. *The homomorphism* $\operatorname{Klen}^\circ : M(R\text{-}\mathbf{Noeth}) \to \mathbf{Krull}$ *is exact.*

Proof. We need to show that $\{\le \operatorname{Klen}^\circ[A]\} \subseteq \operatorname{Klen}^\circ\{\le [A]\}$ for all $A \in R$-**Noeth**. If $A = 0$ then this is trivially true.

So suppose $A \ne 0$ and we have some $(\gamma, n) \in \mathbf{Krull}$ with $(\gamma, n) \le \operatorname{Klen}^\circ[A] = \operatorname{Klen}^\circ A$. Then $\omega^\gamma n \le \operatorname{len}^\circ A$, so using 4.7.1, there is some submodule $A' \le A$ such that $\operatorname{len}(A/A') = \omega^\gamma n$. Thus we have $[A/A'] \le [A]$ and $\operatorname{Klen}^\circ[A/A'] = \operatorname{Klen}^\circ A/A' = (\gamma, n)$. ☐

We now have all the ingredients in place to prove the main theorem of this section:

Theorem 17.4.

1. $\operatorname{Lrad} M(R\text{-}\mathbf{Noeth}) = M(R\text{-}\mathbf{Noeth})$. *In particular,* $M(R\text{-}\mathbf{Noeth})$ *is semi-Artinian.*
2. $M(R\text{-}\mathbf{Noeth})$ *is strongly separative.*
3. *For all* $A \in R$-**Noeth** *and* $\alpha \in \mathbf{Ord}$,

$$[A] \in \operatorname{soc}_\alpha(M(R\text{-}\mathbf{Mod})) \iff \operatorname{Kdim} A < \alpha.$$

Proof. We apply 15.10, to the homomorphism $\operatorname{Klen}^\circ : M(R\text{-}\mathbf{Noeth}) \to \mathbf{Krull}$...

We have already noted that **Krull** is a refinement monoid without proper regular elements. $M(R\text{-}\mathbf{Noeth})$ is an order ideal in the refinement monoid, $M(R\text{-}\mathbf{Mod})$, so is itself a refinement monoid. The homomorphism Klen° is exact by the lemma, and, since the only module with zero Krull length is the zero module, we have $(\operatorname{Klen}^\circ)^{-1}(\{\le 0\}) = \{\le 0\}$.

Therefore, using 15.10, we get immediately 1 and 2 above. From 14.26,

$$\operatorname{soc}_\alpha(M(R\text{-}\mathbf{Noeth})) = \operatorname{soc}_\alpha(M(R\text{-}\mathbf{Mod})) \cap M(R\text{-}\mathbf{Noeth}),$$

and from 15.10.3, we get

$$\operatorname{soc}_\alpha(M(R\text{-}\mathbf{Noeth})) = (\mathrm{Klen}^\circ)^{-1}(\operatorname{soc}_\alpha(\mathbf{Krull})).$$

Thus for $A \in R$-**Noeth**,

$$\begin{aligned}[A] \in \operatorname{soc}_\alpha(M(R\text{-}\mathbf{Mod})) &\iff [A] \in \operatorname{soc}_\alpha(M(R\text{-}\mathbf{Noeth})) \\ &\iff \mathrm{Klen}^\circ[A] \in \operatorname{soc}_\alpha(\mathbf{Krull}) \\ &\iff \mathrm{Kdim}(\mathrm{Klen}^\circ[A]) < \alpha \\ &\iff \mathrm{Kdim}\, A < \alpha.\end{aligned}$$

□

The dual proposition for the Serre category R-**Art** is

Theorem 17.5.

1. $\operatorname{Lrad} M(R\text{-}\mathbf{Art}) = M(R\text{-}\mathbf{Art})$. *In particular,* $M(R\text{-}\mathbf{Art})$ *is semi-Artinian.*
2. $M(R\text{-}\mathbf{Art})$ *is strongly separative.*
3. *For all* $A \in R$-**Art** *and* $\alpha \in$ **Ord**,
$$[A] \in \operatorname{soc}_\alpha(M(R\text{-}\mathbf{Mod})) \iff \mathrm{Kdim}_\circ(A) < \alpha.$$

From these two propositions and 14.8, we have

$$M(R\text{-}\mathbf{Noeth}) + M(R\text{-}\mathbf{Art}) \subseteq \operatorname{srad} M(R\text{-}\mathbf{Mod}).$$

Before discussing the consequences of these theorems, we will provide a more direct way of proving that $M(R\text{-}\mathbf{Noeth})$ and $M(R\text{-}\mathbf{Art})$ are strongly separative which avoids needing to understand semi-Artinian monoids. The disadvantage of this method is that nothing is learned about the relationship between the Loewy series in the monoid and the Krull dimensions of modules as seen in 17.4.3 and 17.5.3.

First we prove a monoid theoretic lemma:

Lemma 17.6. *Let M be a refinement monoid and K an Artinian monoid. If there is a monoid homomorphism $\sigma\colon M \to K$ such that $\sigma(2a) \le \sigma(a)$ implies $a \le 0$ for any $a \in M$, then M is strongly separative.*

Proof. Suppose $a, b, c \in M$ such that $a + c = b + c$ and $c \le a$. We will show that $a = b$...

Define

$$\mathcal{T} = \{(a', b', c', d') \in M^4 \mid a' + c' = b' + c',\ a = d' + a',\ b = d' + b' \text{ and } c' \le a'\}.$$

Let $\mathcal{C} \subseteq M$ be the projection of $\mathcal{T}$ onto the third component. $\mathcal{C}$ is not empty since $(a, b, c, 0)$ is in $\mathcal{T}$. Let $c_0 \in \mathcal{C}$ be chosen such that $\sigma(c_0)$ is minimal in $\sigma(\mathcal{C})$, and let a_0, b_0, d_0 be such that $(a_0, b_0, c_0, d_0) \in \mathcal{T}$.

From Lemma 8.5.1, there is a refinement of $a_0 + c_0 = b_0 + c_0$,

$$\begin{array}{cc} & \begin{array}{cc} b_0 & c_0 \end{array} \\ \begin{array}{c} a_0 \\ c_0 \end{array} & \begin{pmatrix} d_1 & a_1 \\ b_1 & c_1 \end{pmatrix} \end{array}$$

such that $c_1 \le a_1$. Thus $a_1 + c_1 = b_1 + c_1$, $a = (d_0 + d_1) + a_1$, $b = (d_0 + d_1) + b_1$, that is, $(a_1, b_1, c_1, d_0 + d_1) \in \mathcal{T}$ and $c_1 \in \mathcal{C}$. Since $c_1 \le c_0$, we have $\sigma(c_1) \le \sigma(c_0)$, and then the minimality of $\sigma(c_0)$ implies $\sigma(c_0) \le \sigma(c_1)$.

From $c_1 \leq a_1$, we get $2\sigma(c_0) \leq 2\sigma(c_1) \leq \sigma(c_1) + \sigma(a_1) = \sigma(c_1 + a_1) = \sigma(c_0)$. By our hypotheses, this implies $c_0 \leq 0$. Thus $a_0 = b_0$ and $a = d_0 + a_0 = d_0 + b_0 = b$. □

Theorem 17.7. *The monoids $M(R\text{-}\mathbf{Noeth})$ and $M(R\text{-}\mathbf{Art})$ are strongly separative.*

Proof. For these refinement monoids we have the maps Klen° and $\mathrm{Klen}_\circ$ which satisfy the hypotheses of Lemma 17.6. □

For the remainder of this section we will investigate the consequences of Theorem 17.4 for Noetherian modules. Similar results can, of course, be obtained by duality for Artinian modules.

Many simple results can be obtained by reinterpreting a relationship among modules as an equation in the monoid $M(R\text{-}\mathbf{Noeth})$, and then applying strong separativity. For example, the existence of any of the following types of exact sequences in R-**Noeth** implies that $A \sim B$:

$$0 \to A \to A \oplus B \to A \to 0,$$

$$0 \to A \to A \oplus A \to B \to 0,$$

$$0 \to A \oplus A \to A \oplus A \oplus B \to A \to 0,$$

$$0 \to A \to A \to A \to B \to 0,$$

$$0 \to A \to B \to A \to A \to 0,$$

$$0 \to A \to A \to A \to B \to A \to A \to 0.$$

We prove that $A \sim B$ only for the last case: From the given exact sequence we get the equation $3[A] = 2[A] + [B]$ in $M(R\text{-}\mathbf{Noeth})$. Since this monoid is strongly separative, 8.12.4 implies that $[A] = [B]$, that is, $A \sim B$.

We can also apply Theorem 17.4 in a similar way to direct sums of Noetherian modules. For example, if $A, B \in R\text{-}\mathbf{Noeth}$, then

$$A \oplus A \oplus A \sim A \oplus A \oplus B \implies A \sim B.$$

Since $M(R\text{-}\mathbf{Noeth})$ is a strongly separative order ideal in $M(R\text{-}\mathbf{Mod})$, we can use 8.14 to get stronger cancellation properties which involve modules which are not Noetherian. For example, if

$$0 \to C \to A \to B \to C \to 0$$

is an exact sequence in R-**Mod** with $C \in R\text{-}\mathbf{Noeth}$, then $A \sim B$. Here we use the fact that C is isomorphic to a submodule of A and so $[C] \leq [A]$ as well as $[A] + [C] = [B] + [C]$.

For comparison with Theorem 17.10 we single out one particular result of this type:

Proposition 17.8. *Let $A, B \in R\text{-}\mathbf{Mod}$, $C \in R\text{-}\mathbf{Noeth}$ and $n \in \mathbb{N}$. If C is a submodule, factor module or subfactor module of $\bigoplus^n (A \oplus B)$, and $A \oplus C \sim B \oplus C$, then $A \sim B$.*

Proof. In the monoid $M(R\text{-}\mathbf{Mod})$ we have $[A] + [C] = [B] + [C]$ with $[C] \prec [A] + [B]$. Since $[C] \in M(R\text{-}\mathbf{Noeth})$, we also have that $\{\prec [C]\}$ is strongly separative. Thus from 8.14.2, $[A] = [B]$. □

The final aim of this section is to show that, in this proposition, we can drop the hypothesis on C if we have $A \oplus C \cong B \oplus C$ instead of $A \oplus C \sim B \oplus C$. To do this we need a way of cutting down the size of C in the relation $A \oplus C \cong B \oplus C$ so that C is comparable to $A \oplus B$:

For any R-module X we define a map $T_X \colon R\text{-}\mathbf{Mod} \to R\text{-}\mathbf{Mod}$ by

$$T_X(C) = \sum \{\operatorname{im} \gamma \mid \gamma \in \operatorname{Hom}_R(X, C)\},$$

that is, $T_X(C)$ is the sum of all submodules of C which are isomorphic to factor modules of X. We note that if X_1 is a direct summand of X then $T_X(X_1) = X_1$.

Lemma 17.9. *For all $C_1, C_2, X \in R$-$\mathbf{Mod}$, $T_X(C_1 \oplus C_2) = T_X(C_1) \oplus T_X(C_2)$.*

Proof. See [1, Proposition 8.18]. □

Theorem 17.10. *If $A, B \in R$-$\mathbf{Mod}$ and $C \in R$-$\mathbf{Noeth}$ such that $A \oplus C \cong B \oplus C$, then $A \sim B$.*

Proof. We apply the map $T_{A \oplus B}$ to the equation $A \oplus C \cong B \oplus C$...

We have $T_{A\oplus B}(A) = A$ and $T_{A\oplus B}(B) = B$, so using 17.9, we get $A \oplus C' \cong B \oplus C'$ where $C' = T_{A\oplus B}(C)$.

The module C is Noetherian, so C' is a finite sum of images of $A \oplus B$, that is, there is an $n \in \mathbb{N}$ such that C' is a factor module of $\bigoplus^n (A \oplus B)$. Since C' is Noetherian, 17.8 implies that $A \sim B$. □

We should remark that this theorem is not true with the weaker hypothesis that $A \oplus C \sim B \oplus C$. For example, let $R = \mathbb{Z}$. Then from the short exact sequence

$$0 \to \mathbb{Z} \xrightarrow{2} \mathbb{Z} \to \mathbb{Z}_2 \to 0$$

we get $[\mathbb{Z}] = [\mathbb{Z}_2] + [\mathbb{Z}] = [\mathbb{Z}_2 \oplus \mathbb{Z}]$. Hence $0 \oplus \mathbb{Z} \sim \mathbb{Z}_2 \oplus \mathbb{Z}$ but $0 \not\sim \mathbb{Z}_2$.

It is an interesting question to ask what further cancellation properties $M(R\text{-}\mathbf{Noeth})$ may have. We will see in Section 19 that if the ring R is FBN, or, in particular, commutative Noetherian, then $M(R\text{-}\mathbf{Noeth})$ is Artinian. Thus in this case, $M(R\text{-}\mathbf{Noeth})$ has $\leq$-multiplicative cancellation and weak cancellation in addition to strong separativity. Is this true in general?

It is clear from Example 15.9 that a refinement monoid can be semi-Artinian but not have $\leq$-multiplicative cancellation or weak cancellation. So if $M(R\text{-}\mathbf{Noeth})$ has these properties, they arise from some module property that has not yet been taken into account. Also clear from 15.8, is that these two cancellation properties are closely related in semi-Artinian monoids.

18 The Radann Map

The radann map takes R-modules to semiprime ideals of the ring R and respects exact sequences. With it we will be able to link the prime elements of the monoid $M(R\text{-}\mathbf{Mod})$ with the prime ideals of the ring.

Definition 18.1. *Let* $\operatorname{Spec} R$ *be the set of all prime ideals of* R*, and* $\operatorname{SSpec} R$ *the set of all semiprime ideals of* R*, that is, all intersections of sets of prime ideals.* R *is a semiprime ideal by this definition since it is the intersection of the empty set of prime ideals.*

For a two-sided ideal $I \subseteq R$ *we define*

$$\mathcal{K}(I) = \{P \in \operatorname{Spec} R \mid I \subseteq P\} \subseteq \operatorname{Spec} R$$

and the **(prime) radical** *of* I *by*

$$\operatorname{rad} I = \bigcap \mathcal{K}(I) = \bigcap \{P \in \operatorname{Spec} R \mid I \subseteq P\} \in \operatorname{SSpec} R.$$

Clearly $I \subseteq \operatorname{rad} I$, and if $I_1 \subseteq I_2$ then $\mathcal{K}(I_1) \supseteq \mathcal{K}(I_2)$ and $\operatorname{rad} I_1 \subseteq \operatorname{rad} I_2$. Also if I is a semiprime ideal then $\operatorname{rad} I = \bigcap \mathcal{K}(I) = I$.

Lemma 18.2. *If* I_1 *and* I_2 *are two-sided ideals of* R *then*

$$\mathcal{K}(I_1) \cup \mathcal{K}(I_2) = \mathcal{K}(I_1 \cap I_2) = \mathcal{K}(I_1 I_2)$$

$$\operatorname{rad} I_1 \cap \operatorname{rad} I_2 = \operatorname{rad}(I_1 \cap I_2) = \operatorname{rad}(I_1 I_2)$$

Proof. Since $I_1 I_2 \subseteq I_1 \cap I_2 \subseteq I_1, I_2$, we get $\mathcal{K}(I_1 I_2) \supseteq \mathcal{K}(I_1 \cap I_2) \supseteq \mathcal{K}(I_1) \cup \mathcal{K}(I_2)$. Conversely, if $P \in \mathcal{K}(I_1 I_2)$ so $P \supseteq I_1 I_2$, then $P \supseteq I_1$ or $P \supseteq I_2$, that is, $P \in \mathcal{K}(I_1)$ or $P \in \mathcal{K}(I_2)$. This implies $P \in \mathcal{K}(I_1) \cup \mathcal{K}(I_2)$. Thus $\mathcal{K}(I_1 I_2) \subseteq \mathcal{K}(I_1) \cup \mathcal{K}(I_2)$. The equation $\operatorname{rad} I_1 \cap \operatorname{rad} I_2 = \operatorname{rad}(I_1 \cap I_2) = \operatorname{rad}(I_1 I_2)$ follows directly. □

Next we prove a simple property of annihilators of modules. Reminder: The annihilator of a module A, $\operatorname{ann} A = \{r \in R \mid rA = 0\}$, is a two-sided ideal in R.

Lemma 18.3. *If* $0 \to A \to B \to C \to 0$ *is a short exact sequence in* R-$\mathbf{Mod}$*, then*

$$(\operatorname{ann} A)(\operatorname{ann} C) \subseteq \operatorname{ann} B \subseteq \operatorname{ann} A \cap \operatorname{ann} C.$$

Proof. Without loss of generality we can assume that $A \subseteq B$ and $C = B/A$. If $r \in \operatorname{ann} A$, $s \in \operatorname{ann} C$ and $b \in B$, then $b + A \in B/A$ so $s(b + A) = 0$, that is, $sb \in A$. But then $rsb = 0$. Therefore $\operatorname{ann} A \operatorname{ann} C \subseteq \operatorname{ann}(B)$.

Now let $r \in \operatorname{ann} B$. Since $A \subseteq B$, $r \in \operatorname{ann} A$. Also, if $b + A \in C = B/A$ then $r(b + A) = rb + A = 0 + A$ so that $r \in \operatorname{ann} C$. Therefore $\operatorname{ann}(B) \subseteq \operatorname{ann} A \cap \operatorname{ann} C$. □

If we had used right modules instead of left modules, the first inequality would become $(\operatorname{ann} C)(\operatorname{ann} A) \subseteq \operatorname{ann} B$.

Definition 18.4. *We define the map* $\operatorname{radann}\colon R\text{-}\mathbf{Mod} \to \operatorname{SSpec} R$ *as the composition of the radical and annihilator maps,*

$$\operatorname{radann} A = \operatorname{rad}(\operatorname{ann} A)$$

for all $A \in R\text{-}\mathbf{Mod}$.

From Lemmas 18.2 and 18.3 we get immediately

Proposition 18.5. *If* $0 \to A \to B \to C \to 0$ *is a short exact sequence in* $R\text{-}\mathbf{Mod}$, *then*

$$\operatorname{radann} B = \operatorname{radann} A \cap \operatorname{radann} C.$$

This proposition suggests that we should consider $\operatorname{SSpec} R$ to be a monoid with the operation $\cap$ and identity R, so that radann respects exact sequences as a map to $(\operatorname{SSpec} R, \cap)$. The radann map is surjective since for any semiprime ideal I we have $\operatorname{radann}(R/I) = \operatorname{rad} I = I$.

We investigate the properties of the monoid $\operatorname{SSpec} R$...

Proposition 18.6.

1. *The minimum preorder on* $\operatorname{SSpec} R$ *is reverse inclusion:*
$$S_1 \le S_2 \iff S_1 \supseteq S_2$$
for all $S_1, S_2 \in \operatorname{SSpec} R$.
2. $\operatorname{SSpec} R$ *is a partially ordered monoid.*
3. $\operatorname{SSpec} R$ *with its minimum order is a distributive lattice in which*
$$S_1 \vee S_2 = S_1 \cap S_2$$
$$S_1 \wedge S_2 = \operatorname{rad}(S_1 \cup S_2)$$
for all $S_1, S_2 \in \operatorname{SSpec} R$.
4. $\operatorname{SSpec} R$ *has refinement.*

Proof.

1. If $S_1 \le S_2$, then there exists some $S_3 \in \operatorname{SSpec} R$ such that $S_1 \cap S_3 = S_2$. So $S_1 \supseteq S_2$. Conversely, if $S_1 \supseteq S_2$, then $S_2 \cap S_1 = S_2$ so $S_1 \le S_2$.
2. Immediate from 1.
3. The claims that $S_1 \vee S_2 = S_1 \cap S_2$ and $S_1 \wedge S_2 = \operatorname{rad}(S_1 \cup S_2)$ are easy. To show distributivity we calculate using 18.2
$$\begin{aligned} S_1 \wedge (S_2 \vee S_3) &= \operatorname{rad}(S_1 \cup (S_2 \cap S_3)) \\ &= \operatorname{rad}((S_1 \cup S_2) \cap (S_1 \cup S_3)) \\ &= \operatorname{rad}(S_1 \cup S_2) \cap \operatorname{rad}(S_1 \cup S_3) \\ &= (S_1 \wedge S_2) \vee (S_1 \wedge S_3) \end{aligned}$$
4. Follows from the distributivity of the lattice as in Example 7.7.

$\square$

Proposition 18.7. *$S \in \operatorname{SSpec} R$ is a prime element of the monoid if and only if $S = R$ or S is a prime ideal of R.*

Proof.

⇒ Suppose S is a prime element of $\operatorname{SSpec} R$ and $I_1, I_2 \subseteq R$ are two-sided ideals such that $I_1 I_2 \subseteq S$. Using 18.2 , we get $\operatorname{rad} I_1 \cap \operatorname{rad} I_2 = \operatorname{rad}(I_1 I_2) \subseteq \operatorname{rad}(S) = S$, or, as elements of the monoid, $\operatorname{rad} I_1 \cap \operatorname{rad} I_2 \geq S$. Since S is a prime element of the monoid, either $\operatorname{rad} I_1 \geq S$ or $\operatorname{rad} I_2 \geq S$. Hence, either $I_1 \subseteq \operatorname{rad} I_1 \subseteq S$ or $I_2 \subseteq \operatorname{rad} I_2 \subseteq S$. This makes S either a prime ideal of R or R itself.

⇐ Since R is the identity of the monoid, it is also a prime element.

Suppose S is a prime ideal and $S_1, S_2 \in \operatorname{SSpec} R$ such that $S_1 \cap S_2 \geq S$, that is, $S_1 \cap S_2 \subseteq S$. Then $S_1 S_2 \subseteq S$ so either $S_1 \subseteq S$ or $S_2 \subseteq S$, that is, either $S_1 \geq S$ or $S_2 \geq S$. Thus S is a prime element of $\operatorname{SSpec} R$.

□

Since the map radann respects exact sequences, it induces a monoid homomorphism from $M(R\text{-}\mathbf{Mod})$ to $\operatorname{SSpec} R$, and since the monoid $\operatorname{SSpec} R$ is a semilattice, there are further induced monoid homomorphisms from $\overline{M}(R\text{-}\mathbf{Mod})$ and $\widetilde{M}(R\text{-}\mathbf{Mod})$ to $\operatorname{SSpec} R$. These induced maps we will also call radann so we get the following commutative diagram:

$$\begin{tikzcd}
R\text{-}\mathbf{Mod} \arrow[r, "{[\ ]}"] \arrow[drr, "\mathrm{radann}"'] & M(R\text{-}\mathbf{Mod}) \arrow[r, "\{\equiv\ \}"] \arrow[dr, "\mathrm{radann}"] & \overline{M}(R\text{-}\mathbf{Mod}) \arrow[r, "\{\asymp\ \}"] \arrow[d, "\mathrm{radann}"] & \widetilde{M}(R\text{-}\mathbf{Mod}) \arrow[dl, "\mathrm{radann}"] \\
 & & \mathrm{SSpec} &
\end{tikzcd}$$

We next will show that the homomorphism $\mathrm{radann}\colon \widetilde{M}(R\text{-}\mathbf{Mod}) \to \operatorname{SSpec} R$ has a right inverse, so that $\operatorname{SSpec} R$ embeds in $\widetilde{M}(R\text{-}\mathbf{Mod})$. First we need an easy lemma:

Lemma 18.8. *Let A_1 and A_2 be submodules of $A \in R\text{-}\mathbf{Mod}$. Then*

1. $\{\asymp [A_1 + A_2]\} = \{\asymp [A_1]\} + \{\asymp [A_2]\}$
2. $\{\asymp [A/(A_1 \cap A_2)]\} = \{\asymp [A/A_1]\} + \{\asymp [A/A_2]\}$

Proof.

1. Since A_1, A_2 are submodules of $A_1 + A_2$, we get $[A_1], [A_2] \leq [A_1 + A_2]$. Thus $[A_1] + [A_2] \leq 2[A_1 + A_2]$, that is, $[A_1] + [A_2] \prec [A_1 + A_2]$. For the converse we use the module isomorphism $(A_1 + A_2)/A_1 \cong A_2/(A_1 \cap A_2)$ and the submodule series $0 \leq A_1 \leq A_1 + A_2$:
$$[A_1 + A_2] = [A_1] + [(A_1 + A_2)/A_1] = [A_1] + [A_2/(A_1 \cap A_2)] \leq [A_1] + [A_2].$$
Thus $[A_1 + A_2] \prec [A_1] + [A_2]$.
2. Proof is very similar to the proof of 1.

□

Proposition 18.9. *Let* $\Psi\colon \operatorname{SSpec} R \to \widetilde{M}(R\text{-}\mathbf{Mod})$ *be defined by* $\Psi(S) = \{\asymp [R/S]\}$ *for all* $S \in \operatorname{SSpec} R$. *Then* Ψ *is a monoid homomorphism, and* $\operatorname{radann} \circ \Psi$ *is the identity homomorphism on* $\operatorname{SSpec} R$.

Proof. First we check that Ψ is a monoid homomorphism…

The identity element of $\operatorname{SSpec} R$ is R and $\Psi(R) = \{\asymp [R/R]\} = 0$. Also, if $S_1, S_2 \in \operatorname{SSpec} R$ then, using the lemma,

$$\Psi(S_1 \cap S_2) = \{\asymp [R/(S_1 \cap S_2)]\} = \{\asymp [R/S_1]\} + \{\asymp [R/S_1]\} = \Psi(S_1) + \Psi(S_2).$$

Thus Ψ is a monoid homomorphism.

To prove the second claim, suppose $S \in \operatorname{SSpec} R$. Then

$$\operatorname{radann} \Psi(S) = \operatorname{radann}\{\asymp [R/S]\} = \operatorname{radann}(R/S) = \operatorname{rad} S = S.$$

$\square$

In particular, this lemma implies that $\operatorname{SSpec} R$ is embedded in $\widetilde{M}(R\text{-}\mathbf{Mod})$ by the homomorphism Ψ. Since $[R/S] \le [R]$ for any $S \in \operatorname{SSpec} R$, we have more exactly that $\operatorname{SSpec} R$ is embedded in the order ideal $\{\le \{\asymp [R]\}\}$ of $\widetilde{M}(R\text{-}\mathbf{Mod})$. We will see later that for FBN rings, and in particular, for commutative Noetherian rings, Ψ becomes an isomorphism between $\operatorname{SSpec} R$ and $\widetilde{M}(R\text{-}\mathbf{Noeth})$.

19 Noetherian and FBN Rings

The main purpose of this section is to show that if R is a (left) fully bounded Noetherian (FBN) ring, then the monoid $M(R\text{-}\mathbf{Noeth})$ is Artinian. The proof of this fact is primarily a reinterpretation of theorems about FBN rings that are available in the literature. We will use Chapter 8 of K. R. Goodearl and R. B. Warfield Jr., *An Introduction to Noncommutative Noetherian Rings* [11], as a source of such results, and we refer the reader to this book for further details.

All commutative Noetherian rings are FBN, and in fact, FBN rings are studied because they share many properties with commutative Noetherian rings which do not occur in the general non-commutative case.

Before discussing FBN rings we will prove some properties of arbitrary (left) Noetherian rings...

For the first lemma, we do not require that R be a Noetherian ring:

Lemma 19.1. *Let $A \in R\text{-}\mathbf{Mod}$ be finitely generated module. Then $[A] \prec [R/\operatorname{ann} A]$.*

Proof. Suppose $A = Ra_1 + Ra_2 + \ldots + Ra_n$ for generators $a_1, a_2, \ldots, a_n \in A$. For each i, we have $\operatorname{ann} A \subseteq \operatorname{ann}(a_i)$ so that $r + \operatorname{ann} A \mapsto ra_i$ is a homomorphism from $R/\operatorname{ann} A$ to A. Combining these, we get a homomorphism ϕ: $(R/\operatorname{ann} A)^n \to A$ defined by

$$(r_1 + \operatorname{ann} A, r_2 + \operatorname{ann} A, \ldots, r_n + \operatorname{ann} A) \mapsto r_1a_1 + r_2a_2 + \ldots + r_na_n.$$

ϕ is surjective so $[A] \leq [(R/\operatorname{ann} A)^n] = n[R/\operatorname{ann} A]$, that is, $[A] \prec [R/\operatorname{ann} A]$. $\square$

For a Noetherian ring R, a module $A \in R\text{-}\mathbf{Mod}$ is finitely generated if and only if it is Noetherian. In $M(R\text{-}\mathbf{Mod})$, we can add another equivalent condition:

Lemma 19.2. *Let R be a Noetherian ring and $A \in R\text{-}\mathbf{Mod}$. Then*

$$[A] \prec [R] \iff A \text{ is finitely generated} \iff A \text{ is Noetherian.}$$

Proof. If A is finitely generated then from 19.1, $[A] \prec [R/\operatorname{ann} A]$. But $[R/\operatorname{ann} A] \leq [R]$ so $[A] \prec [R]$.

Conversely, suppose $[A] \prec [R]$. Since R is Noetherian, $[R]$ is an element of the order ideal $M(R\text{-}\mathbf{Noeth}) \subseteq M(R\text{-}\mathbf{Mod})$, and so $\{\prec [R]\} \subseteq M(R\text{-}\mathbf{Noeth})$. Thus $[A] \in M(R\text{-}\mathbf{Noeth})$, and A is Noetherian. $\square$

In particular, $M(R\text{-}\mathbf{Noeth}) = \{\prec [R]\} \subseteq M(R\text{-}\mathbf{Mod})$.

Lemma 19.1 says in particular, that if A is finitely generated and $\operatorname{ann} A$ is semiprime then $[A] \prec [R/\operatorname{radann} A]$. If R is Noetherian, then we get the same result without the restriction on the annihilator...

Proposition 19.3. *Let R be a Noetherian ring. Then for all $A \in R\text{-}\mathbf{Noeth}$ we have $[A] \prec [R/\operatorname{radann} A]$.*

Proof. If $A = 0$ then the claim is trivial. Otherwise, from [11, 2.13], there is a submodule series $0 = A_0 < A_1 < \ldots < A_n = A$ such that $P_i = \operatorname{ann}(A_i/A_{i-1})$ is a prime ideal for each i. For each i, $\operatorname{ann} A \subseteq \operatorname{ann}(A_i/A_{i-1}) = P_i$, so we have $\operatorname{radann} A \subseteq P_i$.

Each quotient module A_i/A_{i-1} is Noetherian and so finitely generated. From 19.1,

$$[A_i/A_{i-1}] \prec [R/P_i] \leq [R/\operatorname{radann} A].$$

Thus

$$[A] = \sum_{i=1}^{n} [A_i/A_{i-1}] \prec n[R/\operatorname{radann} A],$$

and $[A] \prec [R/\operatorname{radann} A]$ □

We next consider prime elements in $M(R\text{-}\mathbf{Mod})$ for Noetherian rings.

Proposition 19.4. *Let R be a Noetherian ring, $U \in R$-**Mod** such that $[U]$ is a prime element of $M(R\text{-}\mathbf{Mod})$ and $P = \operatorname{radann} U$. Then P is a prime ideal or $P = R$, and there is a subfactor U_0 of U such that $[U_0] \equiv [U]$ and $P = \operatorname{ann} U_0$.*

Proof. Let

$$\mathcal{U} = \{U' \mid U' \text{ is a subfactor of } U \text{ and } [U'] \equiv [U]\}.$$

Since R is Noetherian, there is some $U_0 \in \mathcal{U}$ such that the ideal $P_0 = \operatorname{ann} U_0$ is maximal among annihilators of elements of $\mathcal{U}$. We will show that P_0 is either a prime ideal or $P_0 = R$.

Suppose I and J are left ideals such that $IJ \subseteq P_0$. The modules JU_0 and U_0/JU_0 are subfactors of U, with $\operatorname{ann}(JU_0) \supseteq \operatorname{ann} U_0 = P_0$ and $\operatorname{ann}(U_0/JU_0) \supseteq \operatorname{ann} U_0 = P_0$.

Since $[U_0] \equiv [U]$, the element $[U_0]$ is prime. The equation

$$[U_0] = [JU_0] + [U_0/JU_0]$$

implies that either $[U_0] \equiv [JU_0]$ or $[U_0] \equiv [U_0/JU_0]$.

In the first case we have $[JU_0] \equiv [U]$, so $JU_0 \in \mathcal{U}$ and the maximality of $\operatorname{ann} U_0$ implies that $\operatorname{ann}(JU_0) = P_0$. From $I(JU_0) \subseteq P_0 U_0 = 0$, we then get $I \subseteq \operatorname{ann}(JU_0) = P_0$.

In the second case, the maximality of $\operatorname{ann} U_0$ implies similarly that $\operatorname{ann}(U_0/JU_0) = P_0$ and so $J \subseteq \operatorname{ann}(U_0/JU_0) = P_0$.

Therefore P_0 is either a prime ideal or $P_0 = R$. Further, since $U_0 \equiv U$, we have $P = \operatorname{radann} U = \operatorname{radann} U_0 = \operatorname{rad} P_0 = P_0$ completing the proof. □

From this proposition and 18.7, we see that radann function takes prime elements of $M(R\text{-}\mathbf{Noeth})$ to prime elements of $\operatorname{SSpec} R$.

Proposition 19.5. *Let U be a nonzero Noetherian R-module such that $[U]$ is a prime element of $M(R\text{-}\mathbf{Mod})$ and $P = \operatorname{ann} U$ is a prime ideal. Then there is a uniform cyclic subfactor U' of U such that $[U'] \equiv [U]$ and $\operatorname{ann} U' = P$.*

Proof. Any Noetherian module contains a uniform submodule [11, 4.15], and any nonzero cyclic submodule of a uniform module is again uniform. So any Noetherian module contains a cyclic uniform submodule. Using this fact and the Noetherian hypothesis, there is a submodule series $0 = U_0 < U_1 < \ldots < U_n = U$ in U whose factors are cyclic uniform modules.

In $M(R\text{-}\mathbf{Noeth})$ we get

$$[U] = [U_1/U_0] + [U_2/U_1] + \ldots + [U_n/U_{n-1}].$$

Since $[U]$ is prime, it is indecomposable and there is some index i such that $[U] \equiv [U_i/U_{i-1}]$. Set $U' = U_i/U_{i-1}$. Then, since P is prime,

$$P = \operatorname{ann} U \leq \operatorname{ann} U' \leq \operatorname{radann} U' = \operatorname{radann} U = \operatorname{ann} U = P.$$

Thus $\operatorname{ann} U' = P$. □

The module U' is isomorphic to R/I for some left ideal I, so combining these two propositions we get

Corollary 19.6. *Let R be a Noetherian ring and $U \in R$-**Noeth** such that $[U]$ is a nonzero prime element of $M(R$-**Mod**$)$. Then there is a left ideal $I \le R$ such that R/I is a uniform module, $[R/I] \equiv [U]$ and $\operatorname{ann} R/I = \operatorname{radann} U$.*

In this corollary, $P = \operatorname{radann} U$ is, by 19.4, a prime ideal. Since also $\operatorname{ann} R/I = P$, we have $P \le I$. Thus every nonzero prime element of $M(R$-**Noeth**$)$ is $\equiv$-equivalent to a uniform factor module of R/P for some prime ideal $P \le R$.

Our next goal is to show that every uniform submodule of R/P for a prime ideal $P \le R$ gives a corresponding prime element in $M(R$-**Noeth**$)$.

Lemma 19.7. *Let R be a Noetherian ring, P a prime ideal of R, and U, V uniform left ideals of R/P. Then V is isomorphic to a submodule of U and vice versa. In particular, $[U] \equiv [V]$.*

Proof. The ring $S = R/P$ is prime left Noetherian (hence prime left Goldie). From [23, 3.3.4], V is isomorphic to a submodule of U and vice versa. This immediately implies $[U] \le [V]$ and $[V] \le [U]$, that is, $[U] \equiv [V]$. □

Proposition 19.8. *Let R be a Noetherian ring, P a prime ideal of R, and U a uniform left ideal of R/P. Then $[U]$ is a prime element of $M(R$-**Mod**$)$.*

Proof. Suppose we have $A, B \in R$-**Mod** such that $[U] \le [A] + [B] = [A \oplus B]$. From 16.9.3, there are submodule series for U and $A \oplus B$ such that every factor in the series for U is isomorphic to a factor in the series for $A \oplus B$. Using the Shreier refinement theorem, we can assume that the series for $A \oplus B$ is a refinement of the series $0 \le A \oplus 0 \le A \oplus B$. In particular, considering the first factor of the series for U, there is a nonzero submodule $V \le U$ which is isomorphic to a subfactor of the series for $A \oplus B$. Thus we have either $[V] \le [A]$ or $[V] \le [B]$.

Any nonzero submodule of a uniform module is also uniform, so by 19.7, $[U] \equiv [V]$, and we have either $[U] \le [V] \le [A]$ or $[U] \le [V] \le [B]$. Therefore $[U]$ is prime. □

Of course, if $U \in R$-**Noeth** as in this proposition then $\operatorname{radann} U = \operatorname{ann} U = P$.

Given a prime ideal P in a Noetherian ring, the nonzero module R/P is Noetherian and so, by [11, 4.15] has a uniform submodule, which by this proposition maps to a prime element of $M(R$-**Mod**$)$. Thus for Noetherian rings we have:

- Every nonzero prime element of $M(R$-**Mod**$)$ is $\equiv$-equivalent to a uniform factor module of R/P for a prime ideal P.
- For every prime ideal P of R there is a uniform submodule U of R/P such that $[U]$ is a prime element of $M(R$-**Mod**$)$.

Without further hypotheses, it is not possible to fill the gap between uniform factor modules of R/P and uniform submodules of R/P :

Example 19.9. *Let F be a field with characteristic zero and $R = A_1(F)$ the Weyl algebra over F. See* [11, pages 14-16] *for details. This ring is simple* [11, 1.15]*, so* 0 *is the only prime ideal of R. On the other hand, R is not a division ring. In particular, R has nonzero* *maximal left ideals. If I is such a maximal left ideal, then $S = R/I$ is a simple module and $[S]$ is an atom, and hence a prime element, of $M(R\text{-}\mathbf{Mod})$. It is also trivially, a factor module of R/P where $P = 0$ is the only prime ideal.*

Since R is a domain but not a division ring, it has no simple submodules: If I were a simple left ideal, then for any $0 \neq x \in I$ we would have $R \cong Rx \leq I$. Since I is simple, this would imply that $R \cong I$, and R is a simple left module. This in turn would imply that R is a division ring.

Since R has no simple submodules, the prime element $[S]$ cannot be constructed from a submodule of R/P.

In FBN rings we will be able to bridge this gap between between uniform factor modules of R/P and uniform submodules of R/P.

Definition 19.10.

1. *A prime ring R is* **left bounded** *if every essential left ideal contains a non-zero two-sided ideal.*
2. *A ring R is* **left fully bounded** *if every prime factor ring of R is left bounded.*
3. *A ring R is* **left fully bounded Noetherian***(left FBN) if it is left fully bounded and left Noetherian.*

Right bounded, right fully bounded and right FBN rings are defined in the obvious way. We will not need to discuss these right-handed variations, and so following the pattern already established for Noetherian rings, any ring labeled as FBN, will be assumed to be <u>left</u> FBN.

Proposition 19.11. *Let R be an FBN ring and $A \in R\text{-}\mathbf{Noeth}$. Then there exist submodules $0 = A_0 < A_1 < \ldots < A_n = A$ such that, for $i = 1, 2, \ldots, n$, $P_i = \operatorname{ann}(A_i/A_{i-1})$ is a prime ideal of R, and A_i/A_{i-1} is isomorphic to a uniform left ideal of R/P_i.*

Proof. See [11, Theorem 8.6]. □

If R is a commutative Noetherian ring then we may take A_i/A_{i-1} to be isomorphic to R/P_i in this proposition. See, for example, [22, 6.4].

Reinterpreting this as a property of $M(R\text{-}\mathbf{Noeth})$ we get

Corollary 19.12. *Let R be an FBN ring and $A \in R\text{-}\mathbf{Noeth}$. Then*

$$[A] = [U_1] + [U_2] + \ldots + [U_n]$$

where, for $i = 1, 2, \ldots, n$, $P_i = \operatorname{ann} U_i$ is a prime ideal of R, and U_i is a uniform left ideal of R/P_i.

In this proposition we include the case $A = 0$ by defining the sum of an empty set of terms to be 0. If R is commutative Noetherian, then we can take $U_i = R/P_i$, so that if $A \in R\text{-}\mathbf{Noeth}$ then

$$[A] = [R/P_1] + [R/P_2] + \ldots + [R/P_n]$$

where, for $i = 1, 2, \ldots, n$, P_i is a prime ideal of R.

Since, by 19.8, each term in the summation of 19.12 is prime we have

Proposition 19.13. *If R is an FBN ring, then $M(R\text{-}\mathbf{Noeth})$ is primely generated.*

Proposition 19.14. *Let R be an FBN ring, $A \in R$-**Mod** such that $[A]$ is a prime element of $M(R\text{-}\mathbf{Mod})$ and $P = \operatorname{radann} A$. Then $A = 0$ or $[A] \equiv [U]$ where U is a uniform left ideal of R/P.*

Proof. If $A \neq 0$, then, from 19.12, $[A] = [U_1] + [U_2] + \ldots + [U_n]$ where for $i = 1, 2, \ldots, n$, $P_i = \operatorname{ann} U_i$ is a prime ideal of R, and U_i is a uniform left ideal of R/P_i. Since $[A]$ is prime, it is also indecomposable, so there is some index I such that $[U_I] \equiv [A]$. This then implies that $P = \operatorname{radann} A = \operatorname{radann} U_I = P_I$ □

In FBN rings we have therefore that <u>every</u> prime element in $M(R\text{-}\mathbf{Noeth})$ is, up to $\equiv$, in the form discussed in 19.8.

Corollary 19.15. *Let R be an FBN ring. Then the restriction of the* radann *map to prime elements of $M(R\text{-}\mathbf{Noeth})$ is strictly increasing.*

Proof. Suppose $[A]$ and $[B]$ are prime elements of $M(R\text{-}\mathbf{Noeth})$ such that $[A] \leq [B]$ and $\operatorname{radann}[A] \geq \operatorname{radann}[B]$. From $[A] \leq [B]$ we get $\operatorname{radann}[A] \leq \operatorname{radann}[B]$, hence $\operatorname{radann}[A] = \operatorname{radann}[B]$. Set $P = \operatorname{radann} A$. If $P = R$, then $A = B = 0$ and $[A] = [B] = 0$. Otherwise, from 19.14, there are uniform left ideals U and V of R/P such that $[A] \equiv [U]$ and $[B] \equiv [V]$. From 19.7 we get $[U] \equiv [V]$, so finally $[A] \equiv [B]$. □

Now we prove the main theorem of this section:

Theorem 19.16. *If R is an FBN ring then $M(R\text{-}\mathbf{Noeth})$ is Artinian.*

Proof. Since R is a Noetherian module, $\operatorname{SSpec} R$ is an Artinian monoid (the minimum order in $\operatorname{SSpec} R$ is <u>reverse</u> inclusion). The subset of prime elements of $M(R\text{-}\mathbf{Noeth})$ is mapped to $\operatorname{SSpec} R$ by the radann map which, by the corollary, is strictly increasing. So, by 2.17, the subset of prime elements of $M(R\text{-}\mathbf{Noeth})$ is Artinian. Since $M(R\text{-}\mathbf{Noeth})$ has refinement and is primely generated (19.13), we can apply 12.13, to get that $M(R\text{-}\mathbf{Noeth})$ is Artinian. □

From 13.1, 13.2 and 17.4 we get

Corollary 19.17. *If R is an FBN ring, then $M(R\text{-}\mathbf{Noeth})$ and $\overline{M}(R\text{-}\mathbf{Noeth})$ have weak cancellation, strong separativity and $\leq$-multiplicative cancellation. $\overline{M}(R\text{-}\mathbf{Noeth})$ is a join-semilattice when viewed as a poclass.*

By the last claim we mean that any pair of elements $a, b \in \overline{M}(R\text{-}\mathbf{Noeth})$ has a supremum $a \vee b \in \overline{M}(R\text{-}\mathbf{Noeth})$. It may not be true that $a \vee b = a + b$.

We consider again the relationship between $\operatorname{SSpec}(R)$ and $\widetilde{M}(R\text{-}\mathbf{Mod})$ that we began in Section 18. Recall from 18.9 that the monoid homomorphism Ψ: $\operatorname{SSpec} R \to \widetilde{M}(R\text{-}\mathbf{Mod})$ defined by $\Psi(S) = \{\asymp [R/S]\}$ for all $S \in \operatorname{SSpec} R$, is a right inverse for radann, that is, $\operatorname{radann} \circ \Psi$ is the identity map on $\operatorname{SSpec} R$.

From 19.3, we have, for R Noetherian and $A \in R$-**Noeth**, that

$$[A] \prec [R/\operatorname{radann} A].$$

Equivalently, this says that

$$\{\asymp [A]\} \leq \{\asymp [R/\operatorname{radann} A]\} = \Psi(\operatorname{radann} A) = \Psi(\operatorname{radann}\{\asymp [A]\})$$

for all $\{\asymp [A]\} \in \widetilde{M}(R\text{-}\mathbf{Noeth})$. For FBN rings this inequality becomes an equality:

Proposition 19.18. *Let R be an FBN ring, and $A \in R$-**Noeth**. Then*

$$\{\asymp A\} = \{\asymp [R/\operatorname{radann} A]\}.$$

Proof. From the above discussion we have $\{\asymp [A]\} \leq \{\asymp [R/\operatorname{radann} A]\}$. So it remains to prove the opposite inequality...

From [11, Theorem 8.9] there exists a finite subset $\{a_1, a_2, \ldots, a_n\} \subseteq A$ such that $\operatorname{ann} A = \bigcap_{i=1}^n \operatorname{ann}(a_i)$. Let $\phi\colon R \to A^n$ be the homomorphism defined by $\phi(r) = (ra_1, ra_2, \ldots, ra_n)$. Then $\ker \phi = \operatorname{ann} A$, and $R/\operatorname{ann} A$ is isomorphic to a submodule of A^n. In addition, $\operatorname{ann} A \leq \operatorname{radann} A$, so we get

$$[R/\operatorname{radann} A] \leq [R/\operatorname{ann} A] \leq [A^n] = n[A],$$

That is, $[R/\operatorname{radann} A] \prec [A]$ and $\{\asymp [R/\operatorname{radann} A]\} \leq \{\asymp [A]\}$ $\square$

Corollary 19.19. *Let R be an FBN ring. Then* $\operatorname{radann}\colon \widetilde{M}(R\text{-}\mathbf{Noeth}) \to \operatorname{SSpec}(R)$ *is an isomorphism with inverse Ψ.*

Proof. We know already that the homomorphism $\operatorname{radann} \circ \Psi$ is the identity homomorphism on $\operatorname{SSpec}(R)$. The proposition says that $\Psi \circ \operatorname{radann}$ is the identity homomorphism on $\widetilde{M}(R\text{-}\mathbf{Noeth})$. $\square$

In the remainder of this section we will calculate $M(R\text{-}\mathbf{Noeth})$ for commutative principal ideal domains (PIDs) and Dedekind domains. These are commutative Noetherian rings and so are FBN.

First we note a few simple properties of domains:

Proposition 19.20. *Let R be a (possibly noncommutative) domain and $r = [R]$.*

1. *If $I = Rx$ is a nonzero principal left ideal of R, then*
$$[R] + [R/I] = [R].$$
2. *For every nonzero left ideal I we have $[I] \equiv [R]$.*
3. *$[R]$ is a prime element of $M(R\text{-}\mathbf{Mod})$.*

*If, in addition, R is Noetherian and $A \in R$-**Noeth**, then*

4. $n_r([A]) < \infty$
5. $n_r([A]) = 0 \iff [A] \ll [R]$
6. $(\forall n \in \mathbb{N})\ (n_r([A]) = n \iff [A] \equiv n[R])$

Proof.

1. As left modules we have $I \cong R$, so $[I] = [R]$ and $[R] + [R/I] = [I] + [R/I] = [R]$.
2. Any ideal I of R is a submodule so we get $[I] \leq [R]$. For the opposite inequality, let x be a nonzero element of I. Then, as in 1, $[R] = [Rx] \leq [I]$.
3. To show that $[R]$ is prime, it suffices to show that it is indecomposable:
Suppose $[R] = [A] + [B] = [A \oplus B]$. Then R and $A \oplus B$ have isomorphic submodule series. Without loss of generality, we can assume that the submodule series for $A \oplus B$ is a refinement of the series $0 \leq A \oplus 0 \leq A \oplus B$ which has factors A and B. The

first nonzero factor module in the series for R is a nonzero ideal, I, say, which by 2 satisfies $[R] \leq [I]$. Using the isomorphism of the submodule series, I is isomorphic to a factor in the series for $A \oplus B$. Thus either $[R] \leq [A]$ or $[R] \leq [B]$.

4. Since $M(R\text{-}\mathbf{Noeth}) = \{\prec [R]\}$ is a strongly separative refinement monoid, $[R]$ is free. Also A is Noetherian, so we have $[A] \prec [R]$ and Lemma 11.32.1 applies.
5. From 11.32.
6. From 11.32.

□

We can now describe precisely the structure of $M(R\text{-}\mathbf{Noeth})$ when R is a PID:

Proposition 19.21. *Let R be a commutative PID and $r = [R] \in M(R\text{-}\mathbf{Mod})$.*

1. *$r + [S] = r$ for all simple modules S, and so $r + a = r$ for all $a \in \operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$.*
2. *For all $a \in M(R\text{-}\mathbf{Noeth})$ we have $n_r(a) < \infty$ and*
   - *$n_r(a) = 0 \iff a \ll r \iff a \in \operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$*
   - *$(\forall n \in \mathbb{N})\ (n_r(a) = n \iff a = nr)$*

Proof.

1. Any simple module S is isomorphic to R/I for some maximal ideal I. Since I is principal, we get, as in 19.20.1, $[R] + [S] = [R]$.
2. The prime ideals of R are 0 and all maximal ideals. From 19.12, every element of $M(R\text{-}\mathbf{Noeth})$ is then a sum of elements of the form $[R]$ and $[R/I]$ where I is a maximal ideal, that is, R/I is simple and $[R/I]$ is an atom. Using 1, an element $a \in M(R\text{-}\mathbf{Noeth})$ is then either in $\operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$ or $a = nr$ for some $n \in \mathbb{N}$.

   From 1, $a \in \operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$ implies $a \ll r$ implies $n_r(a) = 0$. In the converse direction, since $n_r(nr) = n$ for any $n \in \mathbb{N}$, $n_r(a) = 0$ implies $a \in \operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$.

   The rest of the claim is then trivial or is proved in 19.20.

□

This proposition, and the fact that $\operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$ is isomorphic to the free monoid generated by the atoms of the monoid, suffices to determine $M(R\text{-}\mathbf{Noeth})$ for PIDs. Notice that $\{\equiv r\} = \{r\}$ and so G_r is the trivial group.

The next most complicated rings after PIDs are Dedekind domains. What distinguishes $M(R\text{-}\mathbf{Noeth})$ when R is a Dedekind domain from the same monoid when R is a PID, is that in the former case, $G_{[R]}$ may be non-trivial. In fact, we will show that $G_{[R]}$ is isomorphic to the ideal class group of the ring.

Before discussing this fact, we recall the definition and some of the properties of Dedekind domains. For more details, see [26, Chapter VIII.6]. There are many equivalent ways to define Dedekind domains – we will choose one which will be useful in the proofs of the upcoming theorems:

Definition 19.22. *A* **Dedekind domain** *is a hereditary commutative domain. A* **fractional ideal** *of a Dedekind domain R is a nonzero finitely generated R-submodule of the quotient field of R.*

Proposition 19.23. *Let R be a Dedekind domain.*

1. *R is Noetherian.*
2. *All nonzero prime ideals of R are maximal.*
3. *Every fractional ideal is isomorphic to a nonzero ideal of R.*
4. *For all fractional ideals I, J of R*
$$I \oplus J \cong R \oplus IJ.$$
5. *For every fractional ideal I of R there is a unique fractional ideal, written I^{-1}, such that $I(I^{-1}) = R$.*
6. *Every nonzero projective module $P \in R$-**Noeth** is isomorphic to $nR \oplus I$ for some $n \in \mathbb{Z}^+$ and nonzero ideal I. The number n is uniquely determined by P and the fractional ideal I is determined up to isomorphism.*

Proof. See [26, Chapter VIII.6] and [24, Chapter 7]. □

Definition 19.24. *Let R be a Dedekind domain. For a fractional ideal I of R we will write $\langle I \rangle$ for the set of all fractional ideals of R which are isomorphic to I. Then the* **ideal class group** *of R, $\mathcal{I}(R)$, is the set of all isomorphism classes of fractional ideals with operation defined by*
$$\langle I \rangle + \langle J \rangle = \langle IJ \rangle.$$
Proposition 19.23.5 ensures that $\mathcal{I}(R)$ is a group with identity $0 = \langle R \rangle$.

Suppose that R is a Dedekind domain and $r = [R] \in M(R\text{-}\mathbf{Noeth})$. In the next proposition we will show that $G_r \cong \mathcal{I}(R)$. To do so, we will need to think of $\{\equiv r\} \subseteq M$ itself as the group G_r. Recall from 10 that, since $M(R\text{-}\mathbf{Noeth})$ is separative, we can do this if we define the group operation $+_r$ in terms of the monoid operation as follows: If $r_1, r_2 \in \{\equiv r\}$ then $r_1 +_r r_2 = r_3$ where $r_3 \in \{\equiv r\}$ is the unique element such that $r_1 + r_2 = r + r_3$. With this group structure on $\{\equiv r\}$, r is the identity element.

Theorem 19.25. *Let R be a Dedekind domain, $M = M(R\text{-}\mathbf{Noeth})$, and $r = [R] \in M$. Then the map $\sigma\colon \mathcal{I}(R) \to G_r = (\{\equiv r\}, +_r)$ defined by $\sigma(\langle I \rangle) = [I]$, is a group isomorphism.*

Notation: In this proof it will be convenient to write nR, instead of R^n, for the direct sum of n copies of R.

Proof. First we note that, since R is a Noetherian domain, Proposition 19.20 applies. In particular, any fractional ideal I is isomorphic to a nonzero ideal of R, and so $\sigma(\langle I \rangle) = [I]$ is in $\{\equiv r\} = \{\equiv [R]\}$. The identity $\langle R \rangle$ of $\mathcal{I}(R)$ maps to $[R] = r$ which is the identity of $(\{\equiv r\}, +_r)$. Further, for fractional ideals I, J we have
$$\sigma(\langle I \rangle + \langle J \rangle) = \sigma(\langle IJ \rangle) = [IJ]$$
and also from 19.23.4,
$$\sigma(\langle I \rangle) + \sigma(\langle J \rangle) = [I] + [J] = [I \oplus J] = [R \oplus IJ] = [R] + [IJ].$$
Thus $\sigma(\langle I \rangle + \langle J \rangle) = \sigma(\langle I \rangle) +_r \sigma(\langle J \rangle)$, which makes σ a group homomorphism.

We will show that σ is an isomorphism by constructing the inverse homomorphism. To do this we define first a map $\Lambda\colon R\text{-}\mathbf{Noeth} \to \mathcal{I}(R)$ which respects short exact sequences...

Let A be a Noetherian R-module. Since A is finitely generated, there is a short exact sequence of the form
$$0 \to P \to mR \to A \to 0$$

where $m \in \mathbb{N}$, and P is a submodule of mR. The ring R is hereditary, so P is projective. It will be convenient to assume that $P \neq 0$. This we can do since if $P = 0$, we can replace the above sequence in an obvious way by

$$0 \to R \to (m+1)R \to A \to 0.$$

By 19.23.6, there is a fractional ideal I and $n \in \mathbb{Z}^+$ such that $P \cong nR \oplus I$. Using 19.23.5, we set $\Lambda(A) = \langle I^{-1} \rangle$.

We need to check that $\Lambda(A)$ does not depend on the short exact sequence used. Suppose we have a second short exact sequence for A,

$$0 \to P' \to m'R \to A \to 0,$$

with $P \cong n'R \oplus I'$. Then Schanuel's Lemma, [26, Theorem 3.62], implies that $P \oplus m'R \cong P' \oplus mR$, that is $(n+m')R \oplus I \cong (n'+m)R \oplus I'$. The uniqueness part of 19.23.6 implies that $n + m' = n' + m$ and $I \cong I'$. Thus $I^{'-1} \cong I^{-1}$, that is, $\langle I^{'-1} \rangle = \langle I^{-1} \rangle$.

Note that $\Lambda(0) = \Lambda(R) = \langle R \rangle = 0$.

Suppose we have the exact sequence

$$0 \to A_1 \to A_2 \to A_3 \to 0$$

in R-**Noeth**. Using the method described above we can find projective resolutions for A_1 and A_3 to make up the top and bottom rows of the following diagram:

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & n_1R \oplus I_1 \arrow[r] \arrow[d] & m_1R \arrow[r] \arrow[d] & A_1 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & (n_1+n_3)R \oplus I_1 \oplus I_3 \arrow[r] \arrow[d] & (m_1+m_3)R \arrow[r] \arrow[d] & A_2 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & n_3R \oplus I_3 \arrow[r] \arrow[d] & m_3R \arrow[r] \arrow[d] & A_1 \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

Here I_1 and I_2 are fractional ideals, $m_1, m_2 \in \mathbb{N}$ and $n_1, n_3 \in \mathbb{Z}^+$. Using the Horseshoe Lemma, [26, Lemma 6.20], we can fill in the middle row so that all rows and columns are exact. Each new entry in the middle row is the direct sum of the corresponding entries in the top and bottom rows. From 19.23.4, we get $(n_1+n_3)R \oplus I_1 \oplus I_3 \cong (n_1+n_3+1)R \oplus I_1I_3$, so $\Lambda(A_2) = \langle (I_1I_3)^{-1} \rangle = \langle I_1^{-1}I_3^{-1} \rangle = \langle I_1^{-1} \rangle + \langle I_3^{-1} \rangle = \Lambda(A_1) + \Lambda(A_3)$.

Since Λ respects short exact sequences, it induces a unique monoid homomorphism $\lambda\colon M(R\text{-}\mathbf{Noeth}) \to \mathcal{I}(R)$. We will show that the restriction of λ to $\{\equiv r\} \subseteq M$ is the inverse map to σ...

First we show that $\lambda \circ \sigma$ is the identity on $\mathcal{I}(R)$:

Let $\langle I\rangle \in \mathcal{I}(R)$ with I a fractional ideal. Then $\sigma(\langle I\rangle) = [I]$. From 4 and 5 of 19.23, we have $I \oplus I^{-1} \cong R \oplus R$, so there is a short exact sequence

$$0 \to I^{-1} \to 2R \to I \to 0.$$

Thus $\lambda([I]) = \langle (I^{-1})^{-1}\rangle = \langle I\rangle$ and so $\lambda(\sigma(\langle I\rangle)) = \langle I\rangle$.

Finally we show that $\sigma \circ \lambda$ is the identity on $\{\equiv r\} \subseteq M$:

Let $[A] \in \{\equiv r\}$ for some $A \in R$-**Mod**. To find $\sigma(\lambda([A]))$ we find a short exact sequence of the form

$$0 \to nR \oplus I \to mR \to A \to 0$$

with $m \in \mathbb{N}$, $n \in \mathbb{Z}^+$ and I a fractional ideal. Then $\sigma(\lambda([A])) = [I^{-1}]$. From the short exact sequence we get $m[R] = [A]+n[R]+[I]$, and from $I \oplus I^{-1} \cong R \oplus R$, we get $[I]+[I^{-1}] = 2[R]$. Combining these gives $m[R] + [I^{-1}] = (n+2)[R] + [A]$.

From 19.20 we have $n_r([A]) = n_r([I^{-1}]) = 1$, so applying the monoid homomorphism n_r to the above equation gives $m = n+2$ and hence $m[R] + [I^{-1}] = m[R] + [A]$.

We have $[R] \leq [A]$ and $[R] \leq [I]$ so that the separativity of M allows us to cancel $m[R]$ from the equation to get $[I^{-1}] = [A]$. Thus $\sigma(\lambda([A])) = [I^{-1}] = [A]$.

□

We can now describe the structure of $M(R$-**Noeth**$)$ when R is a Dedekind domain:

Theorem 19.26. *Let R be a Dedekind domain and $r = [R] \in M(R$-**Mod**$)$.*

1. *For all $a \in M(R$-**Noeth**$)$ we have $n_r(a) < \infty$ and*
   - $n_r(a) = 0 \iff a \ll r \iff a \in \mathrm{soc}(M(R\text{-}\mathbf{Noeth}))$
   - $(\forall n \in \mathbb{N})$ $(n_r(a) = n \iff a = (n-1)r + [I]$ *for some nonzero ideal I of R)*
2. *If $a \in M(R$-**Noeth**$)$ with $n_r(a) = n \in \mathbb{N}$, then in the expression above for a, $a = (n-1)r + [I]$, the nonzero ideal I is uniquely determined up to module isomorphism.*

*We consider next how to add two elements of $M(R$-**Noeth**$)$:*

4. *If $a = [S]$ for a simple module S, then $S \cong R/I$ for some nonzero ideal I of R. If J is a nonzero ideal of R then*
$$a + [J] = [JI^{-1}].$$
In particular,
$$a + r = [I^{-1}].$$
5. *If I_1 and I_2 are nonzero ideals of R, then*
$$[I_1] + [I_2] = r + [I_1 I_2].$$
In particular,
$$[I] + [I^{-1}] = 2r$$
for any nonzero ideal I.

Proof. We consider first the addition rules 4 and 5...

Item 5, of course, is just a restatement of 19.23.4.

To prove 4, suppose $a \in M(R$-**Noeth**$)$ is an atom, that is, $a = [S]$ for a simple module S, and J is a nonzero ideal. We have $S \cong R/I$ for some nonzero ideal I, so $a+[I] = [S]+[I] = r$. Adding $[J] + [I^{-1}] = r + [JI^{-1}]$ to both sides gives

$$a + [J] + 2r = [JI^{-1}] + 2r.$$

We also have $r \leq [J], [JI^{-1}]$ and so we can use the separativity of $M(R\text{-}\mathbf{Noeth})$ to cancel $2r$ from this equation to get $a + [J] = [JI^{-1}]$.

From 19.23.2, we have that the prime ideals of R are 0 and all maximal ideals. Exactly as in 19.21, this implies that $M(R\text{-}\mathbf{Noeth})$ is generated by r and the atoms of $M(R\text{-}\mathbf{Noeth})$, and also that $n_r(a) = 0 \iff a \ll r \iff a \in \operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$.

Suppose $n_r(a) = n \in \mathbb{N}$. Then, by 19.20, $a \equiv nr$ so $a = nr + s$ for some $s \ll r$, that is, $s \in \operatorname{soc}(M(R\text{-}\mathbf{Noeth}))$. Using 4, we can write $a = (n-1)r + [I]$ for some nonzero ideal I of R. This ideal is unique up to isomorphism, since if $a = (n-1)r + [I] = (n-1)r + [I']$, then, using $r \leq [I], [I']$ and the separativity of $M(R\text{-}\mathbf{Noeth})$ we can cancel $(n-1)r$ from this equality to get $[I] = [I']$. By 19.25, this implies that $\langle I \rangle = \langle I' \rangle$ in $\mathcal{I}(R)$, and hence $I \cong I'$. $\square$

We have already noted in 19.17 that, for an FBN ring R, the monoid $M(R\text{-}\mathbf{Noeth})$ will have $\leq$-multiplicative cancellation. With our new understanding of the structure of $M(R\text{-}\mathbf{Noeth})$ for such rings, we can show that $M(R\text{-}\mathbf{Noeth})$ may not have multiplicative cancellation. Specifically, we will show that $2a = 2b$ does not, in general, imply that $a = b$ in $M(R\text{-}\mathbf{Noeth})$:

Example 19.27. *Let R be a Dedekind domain whose ideal class group is isomorphic to $\mathbb{Z}_2$. An example of such a ring is $R = \mathbb{R}[X,Y]/(X^2+Y^2-1)$. See* [23, 12.1.6]. *Let I be a nonzero ideal of R such that $\langle I \rangle \neq \langle R \rangle = 0$, and $2\langle I \rangle = \langle I^2 \rangle = \langle R \rangle = 0$ in the ideal class group. This implies in particular that $I^2 \cong R$, so, using 19.25 and 19.23.4, we get*

$$2[I] = [I \oplus I] = [R \oplus I^2] = [R] + [I^2] = 2[R],$$

and $[I] \neq [R]$.

Thus $M(R\text{-}\mathbf{Noeth})$ does not have multiplicative cancellation.

References

[1] K. W. Anderson and K. R. Fuller, *Rings and Categories of Modules,* Springer Verlag, (1974).

[2] P. Ara, K. R. Goodearl, K. C. O'Meara and E. Pardo, *Separative Cancellation for Projective Modules over Exchange Rings,* to appear in Israel J. Math., (1997).

[3] G. Birkhoff, *Lattice Theory, 3rd ed.,* A.M.S., (1967).

[4] B. Blackadar, *K-Theory for Operator Algebras,* MRSI Publications, (1986).

[5] B. Blackadar, *C^*-Algebras and Nonstable K-Theory,* Rocky Mountain J. of Math, **20** (1990), 285-316.

[6] G. Brookfield, *Direct Sum Cancellation of Noetherian Modules,* to appear in J. Algebra, (1998).

[7] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups, Vol. I,* Math. Surveys of the A.M.S. **7**, (1961)

[8] H. Dobbertin, *Primely Generated Regular Refinement Monoids,* J. Algebra **91**, (1984), 166-175.

[9] K. R. Goodearl, *Incompressible Critical Modules,* Comm. in Algebra **8** (19), (1980), 1845-1851.

[10] K. R. Goodearl, *Partially Ordered Abelian Groups with Interpolation,* Math. Surveys and Monographs, A.M.S., (1986).

[11] K. R. Goodearl and R. B. Warfield Jr., *An Introduction to Noncommutative Noetherian Rings,* London Math. Soc., (1989).

[12] R. Gordon and J. C. Robson, *Krull Dimension,* Mem. of the A.M.S. **133**, (1973).

[13] T. H. Gulliksen, *A Theory of Length for Noetherian Modules,* J. of Pure and App. Algebra **3**, (1973), 159-170.

[14] R. Guralnick, L. S. Levy and R. B. Warfield Jr., *Cancellation Counterexamples in Krull Dimension 1,* Proc. of the A.M.S. **109** no.2, (1990), 323-326.

[15] P. Halmos, *Naive Set Theory,* Springer Verlag, (1974).

[16] G. Hessenberg, *Grundbegriffe der Mengenlehre,* Göttingen, (1906).

[17] G. Higman, *Ordering by Divisibility in Abstract Algebras,* Proc. London Math. Soc., Ser. 3, **2**, (1952), 326-336.

[18] J. M. Howie, *Fundamentals of Semigroup Theory,* Oxford Univ. Press, (1995).

[19] T. W. Hungerford, *Algebra,* Springer Verlag, (1974).

[20] G. Krause, *Additive Rank Functions in Noetherian Rings,* J. Algebra **130**, (1990), 451-461.

[21] L. S. Levy, *Krull-Schmidt Uniqueness Fails Dramatically over Subrings of $\mathbb{Z}\oplus\mathbb{Z}\oplus\ldots\oplus\mathbb{Z}$,* Rocky Mountain J. Math. **13**, (1983), 659-678.

[22] H. Matsumura, *Commutative Ring Theory,* Cambridge Univ. Press, (1989).

[23] J. C. McConnell and J. C. Robson, *Noncommutative Noetherian Rings,* Wiley, (1987).

[24] D. Passman, *A Course in Ring Theory,* Wadsworth and Brooks, (1991).

[25] M. D. Potter, *Sets, An Introduction,* Oxford Univ. Press, (1990).

[26] J. J. Rotman, *An Introduction to Homolgical Algebra,* Academic Press, (1979).

[27] L. Rowen, *Ring Theory, Volume 1,* Academic Press, (1988).

[28] W. Sierpinski, *Cardinal and Ordinal Numbers,* P. W. N., (1958).

[29] R. G. Swan, *Vector Bundles and Projective Modules,* Trans. A.M.S. **105**, (1962), 264-277.

[30] A. Tarski, *Cardinal Algebras,* Oxford Univ. Press, (1949).

[31] F. Wehrung, *Injective Positively Ordered Monoids I,* J. of Pure and App. Algebra **83**, (1992), 43-82.

[32] F. Wehrung, *Injective Positively Ordered Monoids II,* J. of Pure and App. Algebra **83**, (1992), 83-100.

[33] N. E. Wegge-Olsen, *K-Theory and C^*-Algebras,* Oxford Science Publications, (1993).