

WORKSHEET 6.1: SMITH NORMAL FORM AND THE STRUCTURE THEOREM

Let R be a commutative ring. We say that R is an *integral domain*, or simply a *domain*, if $R \neq 0$ and $ab = 0$ implies $a = 0$ or $b = 0$. An element $u \in R$ is a *unit* if it has a multiplicative inverse in R . We say that a *divides* b , written $a \mid b$, if $b = ac$ for some $c \in R$, and that a and b are *associates* if $b = ua$ for some unit u . A *principal ideal domain*, or *PID*, is a domain in which every ideal is principal, that is, every ideal has the form $\langle d \rangle = Rd$ for some $d \in R$. The zero ideal is included, with generator 0.

You already know two examples, although we have not yet proved that they are PIDs: the integers $\mathbb{Z}$ and the polynomial ring $\mathbb{K}[x]$. In both cases the proof uses division with remainder. A *Euclidean domain* is a domain R together with a function $v: R \setminus \{0\} \rightarrow \mathbb{Z}_{\geq 0}$ such that for all $a, b \in R$ with $b \neq 0$ we can write

$$a = qb + r, \quad \text{with } r = 0 \text{ or } v(r) < v(b).$$

For $\mathbb{Z}$ we take $v(b) = |b|$, and for $\mathbb{K}[x]$ we take $v(f) = \deg(f)$. Notice that v need not measure any ordinary notion of size; all that matters is that it decreases when we pass to a remainder.

The goal of this worksheet is to classify finitely generated modules over a PID by studying their presentations. In Worksheet 4.1 we described modules by generators and relations. A finite list of generators of M gives a surjection $R^n \rightarrow M$, but its kernel need not be finitely generated, and even when M has a finite presentation its matrix can make the relations difficult to recognize. Over a PID both difficulties disappear. We will see that every submodule of R^n is free with a finite basis, and that every presentation matrix can be reduced by changes of basis to a diagonal matrix in *Smith normal form*. This resembles row reduction over a field, with one important difference: over a field we may divide a row by any nonzero pivot, but over $\mathbb{Z}$ dividing a row by 2 usually changes the subgroup generated by the columns, and therefore changes the module being presented. Instead, we will use division with remainder to replace several entries by their greatest common divisor. The diagonal entries which remain record the relations which cannot be removed by any change of basis. We will then show, using the ideals generated by minors, that these diagonal entries do not depend on any of our choices. This leads to the structure theorem for finitely generated modules over a PID, which we finish by applying to finitely generated abelian groups and to nilpotent operators.

Throughout this worksheet all rings are commutative with a multiplicative identity, and all modules are unital. As in the previous worksheets, the letter $\mathbb{K}$ will denote a field.

Let us recall why polynomial division works over a field. Suppose f has leading term ax^m , and g is a nonzero polynomial with leading term bx^n , where $m \geq n$. Subtracting $(a/b)x^{m-n}g$ from f cancels its leading term; this uses that b is invertible, since it belongs to a field. Repeating this step lowers the degree until the remainder is either zero or has degree less than $\deg(g)$, and the multiples of g which we subtracted accumulate into the quotient. For uniqueness, compare two expressions $f = qg + r$: their difference expresses a multiple of g as a difference of remainders, and a nonzero multiple of g cannot have

degree less than $\deg(g)$. We check these statements, and their consequences for ideals, in the following exercise.

1. Domains, Division, and Principal Ideals.

- (a) Prove cancellation in a domain: if $a \neq 0$ and $ab = ac$, then $b = c$. Prove that a product is a unit if and only if both factors are units. Use leading coefficients to show that $\mathbb{K}[x]$ is a domain, and identify the units of $\mathbb{Z}$ and of $\mathbb{K}[x]$.
 - (b) Let $a, b \in \mathbb{Z}$ with $b \neq 0$. Prove that the least nonnegative integer in $a + |b|\mathbb{Z}$ is less than $|b|$, and use it to obtain division with remainder. Deduce that $\mathbb{Z}$ is Euclidean with $v(b) = |b|$. Use the polynomial procedure above to prove existence and uniqueness of the quotient and remainder in $\mathbb{K}[x]$, and deduce that $\mathbb{K}[x]$ is Euclidean with $v(f) = \deg(f)$.
 - (c) Let I be a nonzero ideal in a Euclidean domain, and choose a nonzero $d \in I$ with $v(d)$ as small as possible. Divide an arbitrary $a \in I$ by d . Why does the remainder lie in I , and why must it be zero? Conclude that every Euclidean domain is a PID.
 - (d) Explain separately why the zero ideal causes no difficulty. Deduce that $\mathbb{Z}$ and $\mathbb{K}[x]$ are PIDs. Notice that you have not used anything about factorization into primes or irreducible polynomials.
 - (e) Show that $\langle a \rangle \subset \langle b \rangle$ if and only if $b \mid a$. For nonzero a and b in a domain, prove that $\langle a \rangle = \langle b \rangle$ if and only if a and b are associates. Explain why a nonzero ideal of $\mathbb{Z}$ has a unique positive generator, and a nonzero ideal of $\mathbb{K}[x]$ has a unique monic generator.
-

It is worth keeping in mind the direction of the containment in Problem 1(e): a larger ideal has a generator which divides the old generator. For example, $\langle 12 \rangle \subsetneq \langle 4 \rangle$. We will use this reversal twice, first to prove that factorizations terminate, and later to prove that a reduction of matrices terminates.

Let $a, b \in R$, not both zero. A *greatest common divisor* of a and b is an element $g \in R$ which divides both a and b , and which is divisible by every common divisor of a and b . Here “greatest” refers to divisibility, rather than to a numerical order. In a PID we can find a greatest common divisor by choosing a generator of the ideal

$$\langle a, b \rangle = \{ua + vb \mid u, v \in R\}.$$

We will see that any such generator g is a greatest common divisor, and that it satisfies a *Bézout identity* $g = ua + vb$. In particular, if a and b are *relatively prime*, meaning that their greatest common divisor is a unit, then $ua + vb = 1$ for some $u, v \in R$. The same construction applies to any finite nonempty list of elements, using the ideal generated by the list, and we use 0 as the greatest common divisor of a list whose entries are all zero.

It is natural to ask how to find the coefficients u and v . Over a Euclidean domain, division supplies an algorithm. Replacing the pair (a, b) by $(b, a - qb)$ does not change the ideal it generates. Repeating this with successive remainders eventually produces a nonzero remainder followed by zero, and back substitution then expresses that last nonzero remainder as a linear combination of the original pair. Thus the Euclidean algorithm both finds a common divisor and explains why it generates the same ideal. We check these statements in the following exercise.

2. Greatest Common Divisors and Bézout Identities. Throughout this exercise let R be a PID.

- (a) Suppose $\langle a, b \rangle = \langle g \rangle$. Prove that g divides a and b , that every common divisor of a and b divides g , and that $g = ua + vb$ for some $u, v \in R$. Prove that a greatest common divisor is unique up to multiplication by a unit.
 - (b) Check that $\langle a, b \rangle = \langle b, a - qb \rangle$. In a Euclidean domain, explain why the sequence of nonzero remainders terminates, and why the last one generates the original ideal.
 - (c) Run the Euclidean algorithm on the integers 252 and 198. Back substitute to obtain a Bézout identity for their positive greatest common divisor, and check the identity by direct multiplication.
 - (d) Over $\mathbb{Q}[x]$, run the Euclidean algorithm on $x^3 - 1$ and $x^2 - 1$. Find their monic greatest common divisor, and express it as a polynomial linear combination of the original pair.
 - (e) Suppose a and b are relatively prime and $a \mid bc$. Use a Bézout identity to prove that $a \mid c$. Explain why this conclusion can fail without the assumption that a and b are relatively prime.
-

Before turning to matrices, we prove the factorization result which generalizes prime factorization in $\mathbb{Z}$. A nonzero nonunit p in a domain is *irreducible* if $p = ab$ implies that a or b is a unit, and it is *prime* if $p \mid ab$ implies that $p \mid a$ or $p \mid b$. Notice that these two definitions describe different things. Irreducibility concerns the possible factorizations of p itself, while primality concerns how p divides other products. We will see that the two notions agree in a PID.

There are two separate issues in unique factorization. First, repeatedly factoring a nonunit must eventually stop at irreducibles. Second, different choices of factors must lead to the same list, up to order and multiplication by units. The first issue is a question about termination, and we can settle it using ideals, without a Euclidean function.

Lemma 1 (Stabilization of Ascending Chains of Ideals). *If R is a PID and $I_1 \subset I_2 \subset \cdots$ is an ascending chain of ideals, then there is an index N such that $I_j = I_N$ for every $j \geq N$.*

The idea is that the union of the chain is again an ideal. Since it has a single generator, that generator already appears at some stage, and then that stage contains the entire union.

Theorem 2 (Unique Factorization in a PID). *Every nonzero nonunit in a PID is a finite product of irreducibles. Moreover, if*

$$a = up_1 \cdots p_r = vq_1 \cdots q_s,$$

where u and v are units and all p_i and q_j are irreducible, then $r = s$, and after reordering p_i is associate to q_i for every i .

We prove the lemma and the theorem in the following exercise.

3. Stabilization and Unique Factorization. Throughout this exercise let R be a PID.

- (a) Prove that $I = \bigcup_{j \geq 1} I_j$ is an ideal. Write $I = \langle d \rangle$, choose a stage containing d , and prove the stabilization lemma. Include the case $I = 0$.
- (b) Suppose a nonzero element factors as $a = bc$, where b and c are nonunits. Prove that $\langle a \rangle \subsetneq \langle b \rangle$. Hint: for strictness, suppose $b = at$, and use cancellation to contradict the assumption on c .
- (c) Suppose a nonzero nonunit a_0 is not a finite product of irreducibles. Since a_0 is not irreducible, factor it as a product of two nonunits. Explain why at least one of those factors, say a_1 , is also not a finite product of irreducibles. Repeat this to construct a strictly ascending chain $\langle a_0 \rangle \subsetneq \langle a_1 \rangle \subsetneq \cdots$, and use part (a) to obtain a contradiction.
- (d) Let p be irreducible. If $p \nmid a$, prove that a greatest common divisor of p and a is a unit. Use a Bézout identity to show that if $p \mid ab$ and $p \nmid a$, then $p \mid b$. Deduce that every irreducible element of a PID is prime. Conversely, if a prime element satisfies $p = ab$, use $p \mid a$ or $p \mid b$, together with cancellation, to show that one factor is a unit. Deduce that prime elements are irreducible.
- (e) In the two factorizations in the theorem, use primality to show that p_1 divides some q_j , and explain why p_1 and q_j are associates. Reorder, absorb the resulting unit into v , and cancel a nonzero factor. Finish the proof of uniqueness by induction, treating an empty list of irreducibles as a unit.
- (f) Which part of the proof gives existence, and which gives uniqueness? State the conclusions for $\mathbb{Z}$ and for $\mathbb{K}[x]$, including the ambiguity coming from units in each case. Why did the proof not require R to be Euclidean?

We now return to the kernel of a free presentation. A submodule of R is simply an ideal. Over a PID, every nonzero ideal $\langle a \rangle$ is isomorphic to R by $r \mapsto ra$: this map is surjective by the definition of $\langle a \rangle$, and injective by cancellation. This is the rank one case of a much stronger statement.

Theorem 3 (Submodules of Finite Free Modules). *If R is a PID, then every submodule $N \subset R^n$ is free, with a basis of at most n elements.*

The idea is to project onto the last coordinate. This sends N to a submodule of R , which is an ideal, while the kernel of the projection lies in the first $n - 1$ coordinates, where we can apply induction. If the image ideal is nonzero, we choose one element of N projecting to its generator. Subtracting multiples of that element moves every element of N into the kernel, and the same subtraction will also show that the resulting sum is direct.

Notice that we have not assumed that N is finitely generated. This matters when N is the kernel of a surjection $R^n \rightarrow M$: the theorem produces finitely many relations from finitely many generators of M . Notice also that a basis of N need not extend to a basis of R^n . The inclusion $2\mathbb{Z} \subset \mathbb{Z}$ already shows this difference from vector spaces. We prove the theorem in the following exercise.

4. Bases for Submodules by Projection. Throughout this exercise let R be a PID.

- (a) Prove the theorem for $n = 0$ and $n = 1$, including the zero submodule. In the nonzero rank one case, explain precisely where you use that R is a domain.
- (b) For the induction step, let $p : R^n \rightarrow R$ be projection onto the last coordinate, and put $N_0 := N \cap \ker(p)$. Identify N_0 with a submodule of R^{n-1} , and choose a basis of it using induction. What happens if $p(N) = 0$?
- (c) Suppose $p(N) = \langle a \rangle$ with $a \neq 0$, and choose $v \in N$ with $p(v) = a$. Given $w \in N$, write $p(w) = ca$, and prove that $w - cv \in N_0$. Show that the coefficient c is unique.
- (d) Prove that $N = N_0 \oplus Rv$ and that $Rv \cong R$. Adjoin v to the basis of N_0 , check that the result is linearly independent, and obtain the bound on the size of the basis. Finish the induction.
- (e) Carry out this argument for

$$N = \{(a, b, c) \in \mathbb{Z}^3 \mid a + 2b + 3c \text{ is divisible by } 6\}.$$

First project onto the third coordinate, and then project the kernel onto the second coordinate. Find three basis vectors for N , and express an arbitrary element of N in this basis.

- (f) Suppose M is a finitely generated module over a PID. Choose a surjection $R^n \rightarrow M$, and apply the theorem to its kernel to construct an exact sequence $0 \rightarrow R^m \rightarrow R^n \rightarrow M \rightarrow 0$ with $m \leq n$. Conclude that every finitely generated module over a PID is finitely presented.

Recall from Worksheet 4.1 how to describe a presentation by a matrix. After choosing bases, a homomorphism $R^m \rightarrow R^n$ is given by an $n \times m$ matrix A acting on coordinate columns, and its cokernel is $R^n/\text{img}(A)$. Thus the n rows of A correspond to generators of the cokernel, while each column specifies one relation among them. We also allow redundant relations, so a presentation matrix need not have linearly independent columns, even though the theorem we just proved supplies one which does.

Changing the basis of either free module changes the matrix A without changing its cokernel. More precisely, if $U \in \text{GL}_n(R)$ and $V \in \text{GL}_m(R)$, then A and UAV have isomorphic cokernels. We will carry out these changes of basis using three kinds of row and column operations: interchanging two rows or two columns, adding a multiple of one row or column to another, and multiplying a row or column by a unit. All three operations are invertible over R . Notice that over $\mathbb{Z}$ the last operation allows multiplication by -1 , but not by 2 or by $1/2$.

To keep track of the isomorphism, begin with $U = I_n$, $V = I_m$, and a working matrix $C = A$, and maintain the identity $C = UAV$. A row operation with matrix E replaces C by EC and U by EU , while a column operation with matrix F replaces C by CF and V by VF . Equivalently, we perform the same row operations on I_n and the same column operations on I_m . An invertible integer matrix has determinant 1 or -1 , and is called *unimodular*. We check these statements in the following exercise.

5. Changing a Presentation and Recording Its Maps.

- Write down the inverse of each allowed row and column operation. Using the multiplicativity of the determinant and the adjugate identity from Worksheet 3.2, show that a square matrix over a commutative ring is invertible if and only if its determinant is a unit. Deduce the criterion for invertible integer matrices stated above.
- Prove that $\text{img}(UAV) = U(\text{img}(A))$, and construct the isomorphism of cokernels $[y] \mapsto [Uy]$. Check the rules for updating U and V . Why does V change the relations, but not appear in this formula for elements of the cokernel?
- Compare the 1×1 matrices (2) and (1) over $\mathbb{Z}$. Identify their cokernels, and explain why dividing the first matrix by 2 is not an allowed change of basis over $\mathbb{Z}$, although it is one over $\mathbb{Q}$.
- Let D be an $n \times m$ matrix whose only nonzero entries are diagonal entries $a_1, \dots, a_r$. Prove directly from the coordinate description of the cokernel that

$$\text{coker}(D) \cong R^{n-r} \oplus \bigoplus_{i=1}^r R/\langle a_i \rangle.$$

Explain why a unit diagonal entry contributes a zero summand. Distinguish the effect of a zero row from the effect of an extra zero column.

A rectangular matrix is in *Smith normal form* if its only nonzero entries are diagonal entries $s_1, \dots, s_r$ satisfying

$$s_1 \mid s_2 \mid \cdots \mid s_r.$$

The zero matrix is included, with an empty list of diagonal entries. Over $\mathbb{Z}$ we choose the s_i to be positive, and over $\mathbb{K}[x]$ we choose them to be monic. Notice that these normalizations only multiply the diagonal entries by units.

Theorem 4 (Existence of Smith Normal Form). *For every $n \times m$ matrix A over a PID, there are invertible matrices U and V such that UAV is in Smith normal form.*

We first prove the theorem when division with remainder is available. Move a nonzero entry of smallest Euclidean value to the upper left corner, and call it the *pivot* a . Dividing an entry in the first column by a gives a row operation which either clears that entry or leaves a nonzero remainder of smaller value. In the second case, we interchange rows to make the remainder the new pivot. Column operations do the same for the first row. Thus a nonzero remainder always improves the pivot, and the Euclidean value of the pivot cannot decrease forever.

Clearing the first row and column produces a block matrix $\begin{pmatrix} a & 0 \\ 0 & B \end{pmatrix}$. It is tempting to move on immediately to B . This would produce a diagonal matrix, but its diagonal entries need not satisfy the required divisibility conditions. Instead, before fixing the first pivot we check whether it divides every entry of B . If it does not, we bring an entry it fails to divide into the first row and divide again. This additional step is what makes Smith reduction more than diagonalization. We carry out this argument in the following exercise.

6. Smith Reduction over a Euclidean Domain. Throughout this exercise let R be a Euclidean domain.

- (a) Suppose the pivot is $a \neq 0$, and an entry below it is $b = qa + r$. Use a row operation to replace that entry by r . If $r \neq 0$, use a row interchange to obtain a pivot of strictly smaller Euclidean value. Give the corresponding column operations for an entry in the first row.
- (b) Explain why repeated division eventually clears every other entry in the first row and the first column. Account both for the finitely many improvements of the pivot and for the finitely many clearing operations between improvements.
- (c) Suppose the first row and column have been cleared, but a does not divide an entry b_{ij} of the remaining block. Add the row containing b_{ij} to the first row. Check that the pivot is still a , while the first row now contains the entry b_{ij} . Divide this entry by a , and make its nonzero remainder the new pivot.
- (d) Deduce that eventually the first row and column are cleared and a divides every entry of the remaining block. Explain why operations confined to that block preserve this divisibility.
- (e) Use induction on the smaller of the numbers of rows and columns to reduce the remaining block. Prove that the resulting diagonal entries satisfy the divisibility conditions, including the condition involving the first pivot. Handle a matrix which is entirely zero, as well as a matrix with no rows or no columns.

- (f) Show how to accumulate the operations into U and V , as in Problem 5. Over $\mathbb{Z}$ or $\mathbb{K}[x]$, explain how to make the diagonal entries positive or monic after the reduction.

There is also a useful way to repair a diagonal matrix after it has been produced. Let a and b be nonzero elements of a PID, and let g generate $\langle a, b \rangle$. Write $a = ga_0$ and $b = gb_0$, and choose $u, v \in R$ with $ua + vb = g$. Cancellation gives $ua_0 + vb_0 = 1$. An invertible combination of two rows then replaces the column $(a, b)^T$ by $(g, 0)^T$:

$$P := \begin{pmatrix} u & v \\ -b_0 & a_0 \end{pmatrix}, \quad P \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}.$$

The determinant of P is 1, so we have not divided by a nonunit. You can think of P as the matrix form of the Bézout identity.

Applied to a diagonal pair, the same idea replaces the diagonal entries a and b by g and ab/g . Notice that ab/g means the element ga_0b_0 of R , rather than a quotient taken in a larger field. The new first entry divides the second. For instance, a diagonal matrix with positive integer entries need not be in Smith normal form: divisibility is an additional condition. These two-by-two operations are useful both for calculations by hand and for the proof over a PID which is not Euclidean. We check these statements in the following exercise.

7. Bézout Operations and Diagonal Pairs. Throughout this exercise let R be a PID, and let a, b, g, a_0, b_0, u, v , and P be as above.

- (a) Verify the displayed formula for P , compute its determinant, and write down its inverse. By transposing, obtain an invertible column operation which sends the row (a, b) to $(g, 0)$.
- (b) Put $Q := \begin{pmatrix} 1 & -vb_0 \\ 0 & ua_0 \end{pmatrix}$. Compute its determinant and prove that

$$P \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} Q = \begin{pmatrix} g & 0 \\ 0 & ab/g \end{pmatrix}.$$

Check explicitly that $g \mid ab/g$.

- (c) Apply this formula to the pair 4, 6 over $\mathbb{Z}$. Choose a Bézout identity, and write down the two resulting invertible matrices. Use Problem 5 to identify the cokernels before and after the operation.
- (d) Suppose an earlier diagonal entry c divides both a and b . Prove that c divides both new entries. Explain why repairing a later diagonal pair preserves divisibility by the entries which have already been fixed.

We now carry out a complete example over $\mathbb{Z}$. The rows and columns of the following matrix have useful differences, which make the relations easier to see before we repair the divisibility conditions. Keep a record of the operations, and accumulate U and V as you go. Notice that the final equation $UAV = D$ checks the entire calculation, while the isomorphism $[y] \mapsto [Uy]$ interprets the answer in terms of the original generators. We carry out this calculation in the following exercise.

8. A Three-by-Four Integer Presentation. Throughout this exercise let

$$A = \begin{pmatrix} 6 & 6 & 6 & 12 \\ 6 & 16 & 16 & 22 \\ 6 & 16 & 31 & 37 \end{pmatrix} : \mathbb{Z}^4 \longrightarrow \mathbb{Z}^3.$$

- (a) Start with $U = I_3$ and $V = I_4$. Subtract the second row from the third, and then subtract the first row from the second. Record the resulting matrix and the updated matrix U .
 - (b) Inspect the differences between the columns. Eliminate the fourth column by subtracting suitable other columns, and then work backward through the third and second columns to obtain a diagonal matrix. Record the column operations and update V . Which divisibility conditions still fail?
 - (c) Repair the first two nonzero diagonal entries, using either the Euclidean algorithm or Problem 7. If the first pivot still fails to divide a later diagonal entry, repair that pair as well. Finish the reduction to Smith normal form, keeping all diagonal entries positive.
 - (d) Write down the resulting matrices U , V , and D . Check that $UAV = D$ and that $\det(U), \det(V) \in \{1, -1\}$. Notice that different valid sequences of operations may produce different matrices U and V .
 - (e) Identify $\text{coker}(A)$ using the diagonal presentation. What happens to a unit diagonal entry? Why does the extra column not contribute a free summand to the cokernel? Using your matrix U , write down the isomorphism of cokernels on a vector $y \in \mathbb{Z}^3$.
-

For a general PID we cannot assume that division produces a smaller remainder. We do, however, still have greatest common divisors and Bézout matrices. Suppose the current pivot is a , and an entry b in its row or column is not divisible by a . Replacing the pair by their greatest common divisor changes the ideal generated by the pivot from $\langle a \rangle$ to $\langle a, b \rangle$, which is strictly larger. The stabilization lemma then plays the role that the decreasing Euclidean value played before. Notice that no numerical size needs to decrease: each necessary improvement enlarges a particular ideal. Once the pivot divides the other entries in its row and column, the rest of the argument is the same as in the Euclidean case. We carry out this argument in the following exercise.

9. Smith Reduction over an Arbitrary PID. Throughout this exercise let R be a PID.

- (a) Place a nonzero entry a in the upper left corner. If an entry b below it is not divisible by a , apply the Bézout row operation from Problem 7 to replace the pivot by a generator of $\langle a, b \rangle$ and clear that entry. Prove that $\langle a \rangle \subsetneq \langle a, b \rangle$. Give the corresponding operation for an entry in the first row.
- (b) If a divides an entry in its row or column, clear that entry by subtracting a multiple of the first row or column. Organize the process so that only finitely many clearing operations occur between strict enlargements of the ideal generated by the pivot. Explain why, once a divides all of these entries, clearing the first column and then the first row works.
- (c) Suppose the first row and column have been cleared, but some entry of the remaining block is not divisible by a . Use the row addition from Problem 6(c), followed by a Bézout column operation, to strictly enlarge the ideal generated by the pivot again.
- (d) Use the stabilization lemma to prove that these improvements cannot continue forever. Deduce that eventually the first row and column are cleared and the pivot divides every entry of the remaining block.
- (e) Reduce the remaining block by induction. Explain why divisibility by the fixed pivot survives every later operation, and finish the proof of the existence of Smith normal form for every rectangular matrix over a PID.
- (f) Compare the two arguments for termination. What decreases over a Euclidean domain, and what increases over a general PID? Explain why neither proof uses unique factorization of the entries, or any classification of modules.

We now carry out a second, larger example over $\mathbb{Z}$. The following matrix will appear twice in this worksheet: here we reduce it by hand and keep track of the changes of basis, and later we will compute its minors independently as a check on the result. We carry out the first calculation in the following exercise.

10. A Four-by-Five Presentation by Hand. Throughout this exercise let

$$B = \begin{pmatrix} 4 & 4 & 4 & 8 & 0 \\ 4 & 10 & 10 & 20 & -6 \\ 4 & 10 & 40 & 50 & -36 \\ 8 & 14 & 44 & 58 & -36 \end{pmatrix} : \mathbb{Z}^5 \longrightarrow \mathbb{Z}^4.$$

- (a) Before reducing, compare the fourth row with the first and third rows, and find a row operation which makes one row zero. What does this suggest about a free summand of the cokernel? Do not yet assume that the other rows are linearly independent.

- (b) Use differences of rows and columns to obtain a diagonal matrix. A useful order for the column operations is to simplify the fifth and fourth columns before changing the second and third. Keep a record of the operations, and update $U \in \text{GL}_4(\mathbb{Z})$ and $V \in \text{GL}_5(\mathbb{Z})$ starting from identity matrices.
- (c) Repair any failures of divisibility using two-by-two operations, and then normalize the signs. Write down the final matrices D , U , and V with $UBV = D$. Check that U and V are invertible over $\mathbb{Z}$, and not only over $\mathbb{Q}$.
- (d) Identify $\text{coker}(B)$ as a direct sum of a free abelian group and finite cyclic groups. Use U to write down the isomorphism from the original cokernel, reducing the appropriate coordinates modulo the nonunit diagonal entries.
- (e) Use the columns of V corresponding to the zero columns of D to find a basis for $\ker(B)$. Prove that these vectors span the kernel and are linearly independent. Explain why the same calculation gives information about both the generators and the relations of the presentation.

The existence of Smith normal form describes every finitely generated module over a PID as a direct sum: choose a finite presentation, reduce its matrix, and read off the cokernel one coordinate at a time. The divisibility conditions organize the nonzero diagonal entries, and after discarding the units these are called the *invariant factors*. It is natural to ask whether a different sequence of row and column operations could produce different diagonal entries. More seriously, could a different presentation of the same module produce a different list of cyclic summands? To answer these questions, we need quantities which can be read off from a matrix without choosing a reduction.

For the moment, let R be an arbitrary commutative ring and let A be an $n \times m$ matrix over R . For $k \geq 1$, a $k \times k$ *minor* of A is the determinant of a $k \times k$ submatrix, obtained by choosing k rows and k columns of A . The k th *determinantal ideal* $I_k(A)$ is the ideal of R generated by all the $k \times k$ minors of A . We extend this definition to every integer k using the conventions

$$I_k(A) := R \quad (k \leq 0) \quad \text{and} \quad I_k(A) := \langle 0 \rangle \quad (k > \min(n, m)).$$

The first convention reflects the fact that the empty matrix has determinant 1. Notice that $I_k(A)$ can also be zero when $1 \leq k \leq \min(n, m)$, if all of the corresponding minors vanish.

You already know what minors say over a field. In Worksheet 3.2 we saw that the rank of a linear map can be read off from which of its exterior powers are nonzero, and the matrix entries of $\Lambda^k(\phi)$ are exactly the $k \times k$ minors of the matrix of ϕ . Thus over a field, $I_k(A)$ is either zero or the whole field, and it is nonzero exactly when $k \leq \text{rank}(A)$. Over a general commutative ring a minor carries more information than whether it vanishes: the ideals generated by minors also record divisibility. Determinantal ideals will first show that Smith normal form is unique. Then, once we account for presentations of different sizes, they will lead to invariants of the module itself, called Fitting ideals.

Expanding a determinant along a row shows that

$$I_{k+1}(A) \subset I_k(A),$$

so the determinantal ideals become smaller as the size of the minors increases. Changing bases replaces A by UAV , where U and V are invertible. Individual minors usually change, but we will see that the ideals they generate do not. The reason is a version of the product formula for determinants of rectangular matrices. If P is a $k \times l$ matrix and Q is an $l \times k$ matrix, then the *Cauchy–Binet formula* says that

$$\det(PQ) = \sum_{\substack{J \subset \{1, \dots, l\} \\ |J|=k}} \det(P_{*,J}) \det(Q_{J,*}),$$

where $P_{*,J}$ consists of the columns of P indexed by J , and $Q_{J,*}$ consists of the rows of Q indexed by J , in increasing order. If $k > l$, the sum is empty and equals zero. Notice that the formula follows from the multilinearity and alternation of the determinant, so it holds over every commutative ring, including rings with zero divisors. We prove these statements in the following exercise.

11. Determinantal Ideals and Changes of Basis. Throughout this exercise let A be an $n \times m$ matrix over a commutative ring R .

- (a) Expand a $(k+1) \times (k+1)$ minor along a row to prove that $I_{k+1}(A) \subset I_k(A)$ for $k \geq 1$. Check that the conventions above extend this inclusion to every integer k .
- (b) Let P and Q be as in the Cauchy–Binet formula. Write the j th column of PQ as $\sum_{a=1}^l q_{aj} P_{*,a}$, and expand $\det(PQ)$ multilinearly in all k columns. Explain why the terms in which some column of P is chosen twice vanish.
- (c) Group the remaining terms according to the set J of chosen columns. Reorder those columns increasingly, keeping track of the sign of the permutation, and show that the total coefficient of $\det(P_{*,J})$ is $\det(Q_{J,*})$. Deduce the Cauchy–Binet formula, including the case $k > l$.
- (d) Let U be an $n \times n$ matrix. Fix a set of k rows and a set of k columns of UA , and apply the Cauchy–Binet formula to the corresponding $k \times n$ submatrix of U and $n \times k$ submatrix of A . Deduce that $I_k(UA) \subset I_k(A)$. If U is invertible, apply the same argument to $A = U^{-1}(UA)$ to obtain equality.
- (e) Prove similarly that $I_k(AV) = I_k(A)$ for every invertible $m \times m$ matrix V . Conclude that $I_k(UAV) = I_k(A)$ for every integer k , including the cases covered by the conventions.
- (f) Suppose every entry of an integer matrix B is divisible by c . Explain why every $k \times k$ minor of B is divisible by c^k . Why is this useful for proving that a proposed greatest common divisor divides every minor, but not enough by itself to determine that greatest common divisor?

Now suppose R is a PID. Then each ideal $I_k(A)$ has a generator, which is called the k th *determinantal divisor* of A and is denoted $d_k(A)$. It is determined up to multiplication by a unit. We choose $d_0(A) = 1$, and $d_k(A) = 0$ when $I_k(A)$ is zero. Over $\mathbb{Z}$ we may choose the nonzero determinantal divisors to be positive, and over $\mathbb{K}[x]$ we may choose them to be monic. Notice that these choices are normalizations, rather than additional invariants.

Let D be an $n \times m$ matrix in Smith normal form, with nonzero diagonal entries $s_1 \mid \cdots \mid s_r$. Its nonzero $k \times k$ minors are products of k of its diagonal entries. The divisibility conditions make the product of the first k entries divide every other such product, and since this product is itself a minor, it generates the whole determinantal ideal:

$$I_k(D) = \langle s_1 \cdots s_k \rangle \quad (1 \leq k \leq r) \quad \text{and} \quad I_k(D) = \langle 0 \rangle \quad (k > r).$$

Thus consecutive nonzero determinantal divisors determine each s_k up to a unit. This also explains why an arbitrary diagonal matrix does not give a canonical description: it is the divisibility conditions which single out the smallest product of each size. We prove that Smith normal form is unique in the following exercise.

12. Uniqueness of Smith Normal Form. Throughout this exercise let R be a PID.

- (a) Determine the minors of a rectangular diagonal matrix. If $1 \leq j_1 < \cdots < j_k \leq r$, show that $s_1 \cdots s_k$ divides $s_{j_1} \cdots s_{j_k}$. Prove the displayed formula for $I_k(D)$.
- (b) Suppose $UAV = D$ is in Smith normal form. Use Problem 11 to show that $d_k(A)$ is associate to $s_1 \cdots s_k$ for $k \leq r$, and that $d_k(A) = 0$ for $k > r$.
- (c) Prove that r and the nonzero diagonal entries of D are determined up to units by A . More precisely, show that s_k is associate to $d_k(A)/d_{k-1}(A)$ for $1 \leq k \leq r$. Explain why this quotient exists in R , and why it never requires dividing by zero.
- (d) Conclude that any two Smith normal forms of A agree after normalizing associates. Include the zero matrix and the possibility of unit diagonal entries. What does a unit diagonal entry contribute to the cokernel?
- (e) Over $\mathbb{Z}$, compute the determinantal divisors of $\text{diag}(4, 6, 30)$ and $\text{diag}(2, 6, 60)$ directly. Which of these matrices is in Smith normal form? Explain why knowing only that the products of the diagonal entries agree would not show that two diagonal matrices have the same Smith normal form.

The determinantal divisors also give a way to check a calculation which involved many row and column operations. It is usually unnecessary to list every minor. To prove that an integer is the greatest common divisor of all $k \times k$ minors, it is enough to show that it divides every such minor, and then to exhibit a few minors whose greatest common divisor is no larger.

We return to the matrix B of Problem 10:

$$B = \begin{pmatrix} 4 & 4 & 4 & 8 & 0 \\ 4 & 10 & 10 & 20 & -6 \\ 4 & 10 & 40 & 50 & -36 \\ 8 & 14 & 44 & 58 & -36 \end{pmatrix}.$$

The following calculation uses divisibility of the entries, reduction modulo a prime, and multilinearity of the determinant. Notice that it never uses the Smith normal form you found in Problem 10, so it gives an independent check of that reduction. We carry out this calculation in the following exercise.

13. An Independent Computation of Determinantal Divisors. Throughout this exercise let B be the matrix displayed above.

- (a) Compute $d_1(B)$ directly from the entries of B . Find a small set of entries whose greatest common divisor is already $d_1(B)$.
- (b) Every entry of B is even. Reduce the integer matrix $\frac{1}{2}B$ modulo 3, and show that its rows span a one-dimensional subspace of $\mathbb{F}_3^5$. Deduce that every 2×2 minor of B is divisible by 12.
- (c) Compute the 2×2 minor of B using rows 1,2 and columns 1,2, and the one using rows 2,3 and columns 2,3. Use their greatest common divisor, together with part (b), to determine $d_2(B)$.
- (d) Set

$$u = (1, 1, 1, 2, 0), \quad v = (0, 1, 1, 2, -1), \quad w = (0, 0, 1, 1, -1).$$

Check that the first three rows of B are $4u$, $4u + 6v$, and $4u + 6v + 30w$, and that the fourth row is the sum of the first and third rows. Restrict to any three columns, and expand the determinant multilinearly in the rows to show that every 3×3 minor of B is divisible by $4 \cdot 6 \cdot 30$.

- (e) Compute the leading 3×3 minor of B , and determine $d_3(B)$. Use the relation among the rows to determine $d_4(B)$.
 - (f) Recover the diagonal entries of the Smith normal form of B from these determinantal divisors, and compare with your answer to Problem 10. Identify $\text{coker}(B)$, including its free summand. Explain why the vanishing of $d_4(B)$ records a different phenomenon from a nonzero nonunit diagonal entry.
-

We have proved that determinantal ideals do not change under changes of basis in a fixed presentation. A module, however, can have presentations of different sizes. For example, adding a generator together with a relation specifying its value adds both a row and a column to the presentation matrix. To obtain an invariant of the module, we must let the size of the minors change with the number of generators.

Let R again be an arbitrary commutative ring, and suppose M has a finite presentation

$$R^m \xrightarrow{A} R^n \xrightarrow{\pi} M \longrightarrow 0.$$

For $i \geq 0$, the i th Fitting ideal of M is

$$\text{Fitt}_i(M) := I_{n-i}(A).$$

For now this definition depends on the chosen presentation. Notice that the index counts down from the number of generators, rather than recording the size of a minor. Thus $\text{Fitt}_0(M)$ uses minors whose size is the number of generators, while $\text{Fitt}_i(M) = R$ for $i \geq n$. If the requested minors are larger than either dimension of the matrix, the corresponding ideal is zero.

Theorem 5 (Independence of the Presentation). *For a finitely presented module M over a commutative ring, the ideals $\text{Fitt}_i(M)$ do not depend on the chosen finite presentation. In particular, isomorphic modules have the same Fitting ideals.*

We will prove the theorem in two steps. First we check the elementary changes of presentation which we met in Worksheet 4.1. Then we show that any two finite presentations can be compared using those changes. The second step is essential: invariance under row and column operations only compares different choices of bases, not arbitrary choices of generators and relations. We carry out the first step in the following exercise.

14. Elementary Changes of Presentation. Throughout this exercise let A be an $n \times m$ presentation matrix of M .

- (a) Explain why replacing A by UAV does not change $I_{n-i}(A)$. Notice that the number n of generators does not change.
- (b) A redundant relation has the form Ab for some $b \in R^m$. Let $A' := [A \mid Ab]$. Expand each minor of A' which uses the new column linearly in the old columns, and show that $I_k(A') = I_k(A)$ for every k . Why does a term in which some column is chosen twice vanish? Include the cases in which the size of the minors exceeds the number of old columns.
- (c) Add a generator h with the relation $h = \sum_{j=1}^n c_j g_j$, where $g_1, \dots, g_n$ are the images of the standard basis vectors. The new presentation matrix is

$$\hat{A} = \begin{pmatrix} A & -c \\ 0 & 1 \end{pmatrix}.$$

Check that its cokernel is still M , and use an invertible row operation to replace $\hat{A}$ by $A \oplus (1) := \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$.

- (d) For $k \geq 0$, sort the $(k+1) \times (k+1)$ minors of $A \oplus (1)$ according to whether they use its final row and its final column. Prove that

$$I_{k+1}(A \oplus (1)) = I_{k+1}(A) + I_k(A) = I_k(A).$$

Treat minors which use exactly one of the final row and the final column, and check the equality when k exceeds the dimensions of A .

- (e) The new presentation has $n + 1$ generators. Use part (d) with $k = n - i$ to show that $I_{n+1-i}(\hat{A}) = I_{n-i}(A)$ for $i \leq n$, and check the case $i > n$ using the conventions. Deduce that adding a generator together with its defining relation does not change any Fitting ideal.
- (f) Explain why the inverse changes also preserve the Fitting ideals, whenever the generator or relation being removed is redundant in the sense above. Include permuting generators or relations, and adding a zero relation, among the allowed changes.

To compare two presentations of M , we enlarge both generating sets to the union of the two finite lists. If $\pi : R^n \rightarrow M$ and $\sigma : R^q \rightarrow M$ are the two surjections, the common generating map is

$$\tau : R^n \oplus R^q \longrightarrow M, \quad \tau(u, v) := \pi(u) + \sigma(v).$$

Starting from either presentation, each additional generator can be expressed in terms of the old ones, so this enlargement is one of the changes from Problem 14. Once the generating map is the same, the two lists of relations generate the same kernel, and adding one list to the other only adds redundant relations. We carry out the comparison in the following exercise.

15. Comparing Arbitrary Finite Presentations. Throughout this exercise suppose

$$R^m \xrightarrow{A} R^n \xrightarrow{\pi} M \longrightarrow 0 \quad \text{and} \quad R^p \xrightarrow{A'} R^q \xrightarrow{\sigma} M \longrightarrow 0$$

are finite presentations of the same module M .

- (a) Lift the images of the standard basis vectors of R^q along π to obtain a matrix $C : R^q \rightarrow R^n$ with $\pi C = \sigma$, and similarly obtain $C' : R^n \rightarrow R^q$ with $\sigma C' = \pi$. Why do these lifts exist even though M need not be free?
- (b) Set

$$P := \begin{pmatrix} A & -C \\ 0 & I_q \end{pmatrix}.$$

Show that the columns of P lie in $\ker(\tau)$. If $(u, v) \in \ker(\tau)$, show that $u + Cv \in \ker(\pi)$, write it as Aw , and check that $(u, v) = P(w, v)$. Deduce that $\text{img}(P) = \ker(\tau)$.

- (c) Starting from the second presentation, add the generators of the first one, and reorder the coordinates of the target as $R^n \oplus R^q$. Show that the resulting presentation matrix is

$$Q := \begin{pmatrix} I_n & 0 \\ -C' & A' \end{pmatrix},$$

and prove that $\text{img}(Q) = \ker(\tau)$ by the corresponding calculation.

- (d) Deduce that each column of Q is a linear combination of the columns of P , and conversely. Compare both matrices with the concatenated matrix $[P \mid Q]$ by adding redundant relations, and prove that $I_k(P) = I_k(Q)$ for every k .
- (e) Use Problem 14 to compare the Fitting ideals computed from A , P , Q , and A' , keeping in mind that both P and Q present M with $n + q$ generators. Deduce that

$$I_{n-i}(A) = I_{n+q-i}(P) = I_{n+q-i}(Q) = I_{q-i}(A')$$

for every $i \geq 0$, completing the proof of the theorem. If the two cokernels are only isomorphic, explain how to use the isomorphism to regard both as presentations of the same module.

Return now to a PID R . By Problem 4(f), every finitely generated R -module has a finite presentation, and its presentation matrix has a Smith normal form. The cokernel of a diagonal matrix is easy to describe: a nonzero diagonal entry s contributes $R/\langle s \rangle$, a generator without a nonzero diagonal entry contributes a free summand R , and a unit diagonal entry contributes zero and can be omitted. This proves the existence part of the following theorem.

Theorem 6 (Structure Theorem over a PID). *Every finitely generated module M over a PID R is isomorphic to*

$$R^f \oplus R/\langle s_1 \rangle \oplus \cdots \oplus R/\langle s_r \rangle$$

for some $f, r \geq 0$ and some nonzero nonunits $s_1 \mid s_2 \mid \cdots \mid s_r$. The integer f and the list $s_1, \dots, s_r$, up to associates, are uniquely determined by M . When $r = 0$ the list is empty.

The elements s_j are called the *invariant factors* of M . Over $\mathbb{Z}$ we normalize them to be positive integers greater than 1, and over $\mathbb{K}[x]$ to be monic polynomials of positive degree. Notice that omitting the units is essential for the number of summands to be unique, since adding a zero summand does not change a direct sum.

Uniqueness of Smith normal form for a single matrix does not yet prove that the invariant factors are determined by the module, since two presentations may have different numbers of generators and relations. Fitting ideals resolve this. For the decomposition in the theorem there is a presentation with $f + r$ generators and r diagonal relations, and we will see that

$$\text{Fitt}_i(M) = \begin{cases} \langle 0 \rangle, & 0 \leq i < f, \\ \langle s_1 \cdots s_{f+r-i} \rangle, & f \leq i < f + r, \\ R, & i \geq f + r. \end{cases}$$

The first nonzero Fitting ideal locates the free rank, and the first Fitting ideal equal to R locates the total number of summands. The products in between recover the individual invariant factors. We prove the theorem in the following exercise.

16. Existence and Uniqueness of Invariant Factors. Throughout this exercise let R be a PID and let M be a finitely generated R -module.

- (a) Starting with a finite presentation of M , put its matrix in Smith normal form, and identify the cokernel one coordinate at a time. Remove the unit diagonal entries, and explain why the remaining nonzero entries still satisfy the divisibility conditions. Prove the existence part of the theorem.
- (b) For a module in the form of the theorem, write down its $(f + r) \times r$ diagonal presentation matrix. Compute its determinantal ideals, and derive the displayed formula for the Fitting ideals. Treat the cases $r = 0$, $f = 0$, and $M = 0$.
- (c) Prove that f is the least $i \geq 0$ with $\text{Fitt}_i(M) \neq \langle 0 \rangle$, and that $f + r$ is the least $i \geq 0$ with $\text{Fitt}_i(M) = R$. Use the facts that a product of nonzero elements in a domain is nonzero, and that a product with a nonunit factor is not a unit.
- (d) For $1 \leq k \leq r$, recover the ideal $\langle s_1 \cdots s_k \rangle$ as $\text{Fitt}_{f+r-k}(M)$. Compare generators of consecutive ideals of this form to recover s_k up to a unit. Conclude that f and the invariant factors are determined by M itself, and not by a choice of presentation.
- (e) The *torsion submodule* of M is

$$M_{\text{tor}} := \{m \in M \mid am = 0 \text{ for some nonzero } a \in R\}.$$

Prove that M_{tor} is a submodule, and use the decomposition to identify it with the sum of the cyclic summands $R/\langle s_j \rangle$. Show that $M/M_{\text{tor}} \cong R^f$. The submodule M_{tor} does not depend on any choices; is the same true of a complementary free summand?

- (f) The *annihilator* of M is $\text{Ann}_R(M) := \{a \in R \mid aM = 0\}$. Compute this ideal when $f > 0$, when $f = 0 < r$, and when $M = 0$. For a nonzero torsion module, explain why its largest invariant factor determines the annihilator, but need not determine all of the invariant factors.

There is a second useful way to arrange the torsion summands. By unique factorization, a nonzero nonunit of a PID is a unit times a product of powers of pairwise nonassociate irreducibles. We will see that the quotient by such an element splits into the quotients by the individual prime powers. The result underlying this works over any commutative ring.

Ideals $I, J \subset R$ are *comaximal* if $I + J = R$, or equivalently if there exist $a \in I$ and $b \in J$ with $a + b = 1$. The *product* IJ is the ideal consisting of finite sums of products uv with $u \in I$ and $v \in J$, and products of more than two ideals are defined in the same way. There is always an inclusion $IJ \subset I \cap J$, and we will see that comaximality makes it an equality.

Theorem 7 (Chinese Remainder Theorem). *If $I_1, \dots, I_t \subset R$ are pairwise comaximal ideals, then*

$$\bigcap_{j=1}^t I_j = \prod_{j=1}^t I_j \quad \text{and} \quad R / \prod_{j=1}^t I_j \cong \prod_{j=1}^t R / I_j,$$

where the isomorphism sends the class of $r \in R$ to its classes modulo each I_j .

The main point is surjectivity: can we prescribe arbitrary residue classes simultaneously? For two ideals, an identity $a + b = 1$ supplies coefficients which select one residue class at a time. For several ideals, pairwise comaximality guarantees that the product of all but one of the ideals is still comaximal with the remaining one. This allows an induction, and it also produces elements which isolate each coordinate of the product. We prove the theorem in the following exercise.

17. The Chinese Remainder Theorem. Throughout this exercise let R be a commutative ring.

- (a) Suppose $I + J = R$, and choose $a \in I$ and $b \in J$ with $a + b = 1$. For $z \in I \cap J$, use $z = za + zb$ to prove that $z \in IJ$. Deduce that $I \cap J = IJ$.
- (b) Show that the map $R \rightarrow R/I \times R/J$ has kernel $I \cap J$. Given $u, v \in R$, check that $bu + av$ maps to $(u + I, v + J)$. Deduce the theorem for two ideals.
- (c) Suppose $I + J = R$ and $I + K = R$. Choose $j \in J$ and $k \in K$ which are both congruent to 1 modulo I , and show that $jk \equiv 1 \pmod{I}$. Deduce that $I + JK = R$, and extend this to a finite product of ideals, each comaximal with I .
- (d) Use part (c) and induction on t to prove both assertions of the theorem for pairwise comaximal ideals $I_1, \dots, I_t$. Keep track of the maps, so that the resulting isomorphism is the simultaneous reduction map in the statement.
- (e) For each i , choose $e_i \in \prod_{j \neq i} I_j$ with $e_i \equiv 1 \pmod{I_i}$. Show that $\sum_i e_i r_i$ has the prescribed residues $r_i + I_i$. In $R / \prod_j I_j$, prove that the classes of the e_i are idempotent, that their pairwise products are zero, and that their sum is 1.
- (f) Apply the theorem to $\mathbb{Z}/60\mathbb{Z}$ using the ideals $\langle 4 \rangle$, $\langle 3 \rangle$, and $\langle 5 \rangle$. Construct an explicit inverse to the simultaneous reduction map by finding suitable elements e_i . Explain why the finite product on the right is also a direct sum of $\mathbb{Z}$ -modules.

In a PID, powers of nonassociate prime elements generate comaximal ideals. Indeed, a common nonunit divisor would have a prime divisor which divides both powers, contradicting unique factorization. Thus a greatest common divisor is a unit, and a Bézout identity shows that the ideals are comaximal. If $s = u \prod_p p^{e_p}$, where u is a unit and only finitely many exponents e_p are positive, the Chinese remainder

theorem gives

$$R/\langle s \rangle \cong \bigoplus_{p: e_p > 0} R/\langle p^{e_p} \rangle.$$

Applying this to each invariant factor gives the *elementary divisor form* of M : a free summand together with cyclic quotients by prime powers. These prime powers, considered up to associates, are the *elementary divisors* of M .

The two arrangements emphasize different information. Invariant factors form a single divisibility chain, while elementary divisors separate the contributions of different primes. To pass back from elementary divisors to invariant factors, list the exponents of each prime in increasing order, pad the shorter lists on the left with zeros, and multiply the prime powers in corresponding positions. The zeros contribute units. The resulting products form a divisibility chain, and the Chinese remainder theorem reassembles the same prime-power summands. This also shows that the elementary divisors are unique, using the uniqueness of invariant factors which we have already proved. We check these statements in the following exercise.

18. Elementary Divisors and Invariant Factors. Throughout this exercise let R be a PID.

- (a) For nonassociate prime elements $p, q \in R$ and positive integers a, b , prove that $\langle p^a \rangle + \langle q^b \rangle = R$. Apply the Chinese remainder theorem to prove the displayed decomposition of $R/\langle s \rangle$.
- (b) Given invariant factors $s_1 \mid \cdots \mid s_r$, factor each one and record the elementary divisors. For a fixed prime, show that its exponents in $s_1, \dots, s_r$ are nondecreasing, possibly beginning with some zero exponents.
- (c) Conversely, suppose we are given a finite multiset of prime powers. Let r be the largest number of these powers belonging to a single prime, with $r = 0$ for the empty multiset. Pad and align the lists of exponents as described above. Prove that the resulting elements $s_1, \dots, s_r$ are nonunits and satisfy $s_1 \mid \cdots \mid s_r$.
- (d) Use the Chinese remainder theorem to show that the module $\bigoplus_j R/\langle s_j \rangle$ constructed in part (c) is isomorphic to the original direct sum of quotients by prime powers. Deduce that the multiset of elementary divisors of a finitely generated module is unique, up to associates and order.
- (e) Over $\mathbb{Z}$, convert the elementary divisors 4, 2, 9, 3, 5 to invariant factors. Over $\mathbb{K}[x]$, convert the direct sum of $\mathbb{K}[x]/\langle x^3 \rangle$, $\mathbb{K}[x]/\langle x \rangle$, and $\mathbb{K}[x]/\langle (x-1)^2 \rangle$ to invariant factor form. Explain why the same procedure works in every characteristic.

Applying the structure theorem over $\mathbb{Z}$ classifies finitely generated abelian groups. Every such group is a direct sum of finitely many copies of $\mathbb{Z}$ and finitely many finite cyclic groups, whose orders can be arranged

in a divisibility chain. In particular, a finitely generated abelian group is finite if and only if its free rank is zero, and in this case its order is the product of its invariant factors.

The order alone does not classify finite abelian groups. Two groups can have the same order while their elements have different orders. For example, $\mathbb{Z}/p^2\mathbb{Z}$ and $\mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z}$ both have p^2 elements, but only the first has an element of order p^2 . For a fixed prime power p^a , the possible lists of elementary divisors correspond to the partitions of a , that is, to the ways of writing a as a sum of positive integers.

Fitting ideals give another way to distinguish such groups. Over $\mathbb{Z}$, each Fitting ideal has a unique nonnegative generator, which we can compare directly with the determinantal divisors of a presentation, keeping in mind that the two are indexed differently. We compute some examples in the following exercise.

19. Classification of Finitely Generated Abelian Groups.

- (a) Write $\mathbb{Z}^2 \oplus \mathbb{Z}/12\mathbb{Z} \oplus \mathbb{Z}/18\mathbb{Z}$ in invariant factor form and in elementary divisor form. Identify its torsion subgroup, and compute each of its Fitting ideals.
- (b) Classify all abelian groups of order $72 = 2^3 \cdot 3^2$. First list the partitions of 3 and of 2, and then form the corresponding direct sums of cyclic groups of prime-power order. Give the invariant factors of each group, and use uniqueness to show that your list has no repetitions.
- (c) Compare $\mathbb{Z}/4\mathbb{Z} \oplus \mathbb{Z}/18\mathbb{Z}$ and $\mathbb{Z}/6\mathbb{Z} \oplus \mathbb{Z}/12\mathbb{Z}$. Compute their invariant factors and their annihilators. They have the same order; which invariant shows that they are not isomorphic?
- (d) Let $M = \text{coker}(B)$, for the matrix B of Problem 10. Compute $\text{Fitt}_0(M), \dots, \text{Fitt}_4(M)$ directly from the determinantal divisors of its presentation with four generators. Check your answers using the invariant factor form of M .
- (e) For a nonzero finite abelian group in invariant factor form, show that its largest invariant factor is the least positive integer which annihilates the whole group. Show that a finite abelian group is cyclic if and only if it has at most one invariant factor, including the case of the zero group.

The integer f in the structure theorem also has a direct interpretation using the fraction field $K := \text{Frac}(R)$. Extending scalars to K makes every nonzero element of R invertible. It therefore kills each torsion summand, while turning each free summand R into a copy of K :

$$K \otimes_R M \cong K^f.$$

For this reason $f = \dim_K(K \otimes_R M)$ is called the *rank* of M . This gives a second proof that the free rank is determined by M , without using Fitting ideals.

The exterior powers of Worksheet 3.2 give a test for this dimension which does not use a basis. For a finite-dimensional K -vector space V , the last nonzero exterior power is $\Lambda_K^{\dim(V)}(V)$. Applied to $V = K \otimes_R M$,

this detects the free rank of M and ignores its torsion. Notice that reducing modulo a maximal ideal $\mathfrak{m}$ behaves differently: a torsion summand survives in $M/\mathfrak{m}M$ when its invariant factor lies in $\mathfrak{m}$. Comparing these two changes of scalars explains why the dimension of $M/\mathfrak{m}M$ can be larger than the free rank. We compare them in the following exercise.

20. Rank, Residue Fields, and Exterior Powers. Throughout this exercise let R be a PID with fraction field K , and let M be a finitely generated R -module in the form of the structure theorem.

(a) For nonzero $s \in R$, use $K \otimes_R R/\langle s \rangle \cong K/sK$ to prove that this tensor product is zero. Use the fact that tensor products preserve finite direct sums to show that $K \otimes_R M \cong K^f$, and deduce the uniqueness of f by comparing dimensions.

(b) If $\dim_K(V) = f$, use the basis of increasing wedges from Worksheet 3.2 to prove that $\Lambda_K^j(V) \neq 0$ if and only if $0 \leq j \leq f$. Include the case $V = 0$, for which $\Lambda_K^0(V) = K$. Deduce that the rank f of M is characterized by

$$\Lambda_K^f(K \otimes_R M) \neq 0 \quad \text{and} \quad \Lambda_K^{f+1}(K \otimes_R M) = 0.$$

(c) Let $\mathfrak{m} \subset R$ be a maximal ideal and put $\kappa := R/\mathfrak{m}$. Compute $\kappa \otimes_R -$ applied to each summand in the decomposition of M , and prove that

$$\dim_\kappa(M/\mathfrak{m}M) = f + \#\{j \mid s_j \in \mathfrak{m}\}.$$

(d) For $M = \text{coker}(B)$, compute the rank of M and the dimension of M/pM over $\mathbb{F}_p$ for $p = 2, 3, 5, 7$. Determine which exterior powers of $\mathbb{Q} \otimes_{\mathbb{Z}} M$ are nonzero. Explain why these answers describe different features of the same module.

We finish by classifying nilpotent operators. Fix a positive integer N and let $A := \mathbb{K}[x]/\langle x^N \rangle$. An A -module is the same thing as a $\mathbb{K}[x]$ -module on which x^N acts as zero. In terms of the dictionary between $\mathbb{K}[x]$ -modules and operators from Worksheet 1.2, it is a $\mathbb{K}$ -vector space together with a linear map T satisfying $T^N = 0$. A linear map with some power equal to zero is called *nilpotent*. We will classify the A -modules which are finite dimensional over $\mathbb{K}$.

The structure theorem applies to their underlying $\mathbb{K}[x]$ -modules. The relation $x^N M = 0$ rules out a nonzero free summand, and it forces every invariant factor to divide x^N , so that after normalizing each invariant factor is a power of x . Consequently every such module is isomorphic to

$$\bigoplus_{j=1}^r \mathbb{K}[x]/\langle x^{a_j} \rangle \quad \text{for some} \quad 1 \leq a_1 \leq \cdots \leq a_r \leq N.$$

This is also an isomorphism of A -modules, since on both sides the action of $\mathbb{K}[x]$ factors through the same quotient ring. The uniqueness of invariant factors determines the list of exponents, and each exponent contributes that many dimensions over $\mathbb{K}$.

The summand $\mathbb{K}[x]/\langle x^a \rangle$ has basis $1, x, \dots, x^{a-1}$. Multiplication by x sends each basis vector to the next, and sends the last one to zero. Thus the decomposition describes a nilpotent operator using finitely many strings of successive images. In the next worksheet we will study these summands as indecomposable modules, and compare this decomposition with composition series. For now, the structure theorem already gives their existence and uniqueness, as we check in the following exercise.

21. Modules over a Truncated Polynomial Ring. Throughout this exercise let $A = \mathbb{K}[x]/\langle x^N \rangle$.

- (a) Show that an A -module which is finite dimensional over $\mathbb{K}$ is finitely generated as a $\mathbb{K}[x]$ -module. Apply the structure theorem, and use the fact that x^N annihilates the entire module to prove that its free rank is zero.
- (b) Suppose $\mathbb{K}[x]/\langle s \rangle$ is an invariant factor summand of such a module. Prove that $s \mid x^N$, and use unique factorization in $\mathbb{K}[x]$ to show that the monic normalization of s is x^a for some $1 \leq a \leq N$.
- (c) Prove that a map between $\mathbb{K}[x]$ -modules annihilated by x^N is $\mathbb{K}[x]$ -linear if and only if it is A -linear. Deduce that the displayed decomposition classifies finite-dimensional A -modules up to isomorphism, with a unique list of exponents. Include the zero module.
- (d) Prove that $\dim_{\mathbb{K}}(M) = \sum_j a_j$. Classify all 5-dimensional modules over $\mathbb{K}[x]/\langle x^3 \rangle$ by listing the partitions of 5 with every part at most 3.
- (e) On $\mathbb{K}[x]/\langle x^a \rangle$, compute the kernel and image of multiplication by x^l for $l \geq 0$. Deduce that

$$\dim_{\mathbb{K}}(\ker(T^l)) = \sum_j \min(l, a_j).$$

For $l \geq 1$, show that the difference between these dimensions for l and for $l - 1$ is the number of summands with $a_j \geq l$. Deduce that the dimensions of these kernels also determine the list of exponents.

- (f) Write down the matrices of multiplication by x on $\mathbb{K}[x]/\langle x^3 \rangle \oplus \mathbb{K}[x]/\langle x \rangle$ and on $(\mathbb{K}[x]/\langle x^2 \rangle)^2$, using the bases of powers of x in each summand. Both modules have dimension 4. Use either the dimensions of kernels or the invariant factors to prove that they are not isomorphic as $\mathbb{K}[x]/\langle x^3 \rangle$ -modules.
-