Discrete tomography
for integer-valued functions

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties te verdedigen
op woensdag 15 juni 2011 klokke 16:15 uur

door

Arjen Pieter Stolk,

geboren te 's-Gravendeel in 1982.

Samenstelling van de promotiecommissie:

Promotor
prof. dr. S.J. Edixhoven

Co-promotor
prof. dr. K.J. Batenburg

Overige leden
prof. dr. L. Hajdu (University of Debrecen)
prof. dr. A. M. Cohen (Technische Universiteit Eindhoven)
prof. dr. R. Tijdeman
prof. dr. H. W. Lenstra
prof. dr. P. Stevenhagen

Discrete tomography
for integer-valued functions

Voor Loek

Discrete tomography for integer-valued functions
Arjen Stolk

THOMAS STIELTJES INSTITUTE
FOR MATHEMATICS

NWO
Nederlandse Organisatie voor Wetenschappelijk Onderzoek

The research leading to this thesis was supported by NWO.

Typeset in LaTeX
Printed by Ridderprint, Ridderkerk

Contents

Introduction

The material presented in this thesis belongs, as the title suggests, to the subject of discrete tomography. The word 'tomography' is derived from the Greek 'tomos' (slice) and 'graphein' (to write). It describes the study of reconstructing images from projection data. This is a very broad description and when working on tomography problems, one typically has in mind very specific types of images and projections.

One important application of tomography is in medical imaging, e.g. in CT-scans, where the images are density pictures of cross-sections of the body and the projections are the data obtained by transmitting X-rays through the desired cross-section from various angles. This idea of reconstructing a cross-section or slice of the body is the origin of 'tomos' in tomography.

In this thesis, the images are functions $f : A \to \mathbb{Z}$, where A is a subset of $\mathbb{Z}^r$, the integer lattice in r-space. These are the integer-valued functions from the title. The projection data are *line sums* of the function f: we have a set $\mathcal{L}$ of lines that meet A and for each $\ell \in \mathcal{L}$ we know

$$\sum_{x \in A \cap \ell} f(x).$$

In order to make these line sums well-defined in all cases, we restrict our attention to functions f that are 0 almost everywhere, i.e., such that the set $\{x \mid f(x) \neq 0\}$ is finite.

The first chapter presents a thread in the history of discrete tomography starting from the foundational work of Ryser [22] and leading to the work of Hajdu and Tijdeman [12] and van Dalen [7] that is the starting off point for the rest of this thesis.

Chapter 2 introduces the algebraic setup that will be used in the rest of the text. It contains many examples that connect the material from the first chapter to this setup. A fair amount of algebraic machinery is introduced and explained. The chapter contains no real new results, but provides a solid foundation for the theory.

Chapter 3 begins with an analysis of the case $A = \mathbb{Z}^r$, where the reconstruction problem has a very rich algebraic structure. Using the results from this 'global' case it then focuses on A that are finite and convex. A final section specialises to planar A, i.e., the case $r = 2$. This chapter extends the

results from the paper [2] by Joost Batenburg and the author, which only deals with the planar case.

Chapter 4 describes dependencies between line sums in the case $A = \mathbb{Z}^2$. These are linear relations that will always hold between the line sums of an image. It is shown in chapter 3 that the space of such dependencies has finite dimension and that satisfying them all is a sufficient condition for a set of line sums to come from an image. In this chapter we derive algorithms to produce bases for the space of dependencies, first over $\overline{\mathbb{Q}}$, then over $\mathbb{Q}$ and over $\mathbb{Z}$.

The fifth and final chapter studies the case where A is periodic, i.e. it has many translation symmetries. We describe the basic structure and give algorithms to compute important spaces associated to the line sum map. The theory in this chapter is not as complete as in the previous ones and the chapter concludes with some open questions and thoughts on future research.

1 – Discrete tomography

This chapter is a short overview of results in discrete tomography leading up to the problems that will be discussed in the rest of the thesis. Starting with Ryser's work on binary matrices [22], we present two generalisations or adaptations of the original setup. The same central questions are studied in each setup, but the answers and the mathematics that go into them vary from case to case.

The work of compiling this chapter was much alleviated by Herman and Kuba's excellent book [13]. The first part of this book deals with the foundations of discrete tomography. Though I shall usually cite more direct sources, I derived my understanding of the material in the current chapter through the expositions in this book.

As mentioned, the first section will deal with binary matrices and their row and column sums. In section 1.2 we generalise this in an obvious way to reconstructing more arbitrary subsets of a lattice from line sums in more general directions. From this point one can go in many different directions. We follow the work of Hajdu and Tijdeman in [12], which ties in with the material that will be studied in the rest of this thesis. The final section introduces a number of examples closely related to the material in 1.3. We will revisit these examples throughout the following chapters and answer some of the unresolved questions they raise at this point.

1.1 Binary matrices

Definition 1.1.1. Let m and n be positive integers. An $m \times n$ *binary matrix* is a matrix

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

where $a_{i,j} \in \{0, 1\}$ for all i and j.

Definition 1.1.2. Let $(a_{i,j})$ be an $m \times n$ binary matrix. For $i = 1, \ldots, m$ the i-th *row sum* is

$$a_{i,1} + \cdots + a_{i,n}.$$

For $j = 1, \ldots, n$ the j-th *column sum* is

$$a_{1,j} + \cdots + a_{m,j}.$$

The central theme of the questions we want study is the following.

Suppose we are given the row sums (r_i) and column sums (c_j) of some unknown $m \times n$ binary matrix. What can we say about the entries of this matrix from these row and column sums?

Specifically, three central questions are usually studied with regards to this and other problems in (discrete) tomography. Suppose that non-negative integers $r_1, \ldots, r_m$ and $c_1, \ldots, c_n$ are given.

1. **Consistency.** Is there an $m \times n$ binary matrix having these numbers as its row and column sums?

2. **Uniqueness.** Is there at most one $m \times n$ binary matrix having these numbers as its row and column sums?

3. **Reconstruction.** How can one construct an $m \times n$ binary matrix having these numbers as its row and column sums?

Certain variations to these questions can be considered as well. For example, the uniqueness problem may be modified by asking for a given binary matrix if there are other binary matrices having the same row and column sums. For a reconstruction algorithm, it may make a difference whether or not we know a priori that there exists a binary matrix with the given row and column sums.

We shall see these questions come back in slightly modified form throughout this chapter and they form the motivation for most if not all the work in the rest of this thesis.

In the context of binary matrices, the central questions have been completely answered. Ryser's 1957 paper [22] deals with all three of them. Independently in the same year, Gale in [9] derived the same consistency conditions and reconstruction algorithm by different means.

Uniqueness

Definition 1.1.3. Two $m \times n$ binary matrices are *tomographically equivalent* if they have the same row and column sums.

The simplest example of two distinct but tomographically equivalent binary matrices is the following pair.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The content of Ryser's theorem, which we are going to prove as theorem (1.1.9), is that this simple example essentially accounts for *all* tomographically equivalent matrices. The proof below is an adaptation of the proof in Chapter 3 of [13] by Herman and Kuba. Ryser's original paper [22] contains a different proof.

Definition 1.1.4. Let $(a_{i,j})$ and $(b_{i,j})$ be two $m \times n$ binary matrices. Then they differ by a *4-switch* if there are i_1, i_2, j_1 and j_2 such that

$$\{\begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} \\ a_{i_2,j_1} & a_{i_2,j_2} \end{pmatrix}, \begin{pmatrix} b_{i_1,j_1} & b_{i_1,j_2} \\ b_{i_2,j_1} & b_{i_2,j_2} \end{pmatrix}\} = \{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\}$$

and $a_{i,j} = b_{i,j}$ whenever i is not in $\{i_1, i_2\}$ or j is not in $\{j_1, j_2\}$.

The concept of 4-switch goes by many different names in the literature. Ryser in [22] used the term *interchange*. Several names come up in [13], *switching component* appears in the first chapter, while 4-switch is used in Chapter 3.

It is clear that binary matrices which differ by a 4-switch are tomographically equivalent.

Definition 1.1.5. Let $(a_{i,j})$ and $(b_{i,j})$ be two $m \times n$ binary matrices. A *switching chain* of length t for a and b is a sequence of distinct positions

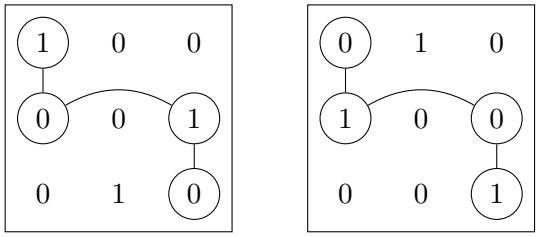
$$(i_1, j_1), \ldots, (i_t, j_t)$$

such that

- a_{i_1,j_1} and b_{i_1,j_1} are different;
- for $s = 1, \ldots, t-1$ one has $a_{i_s,j_s} - b_{i_s,j_s} = b_{i_{s+1},j_{s+1}} - a_{i_{s+1},j_{s+1}}$;
- for $s = 1, \ldots, t-1$ one has $i_s = i_{s+1}$ if s is even and $j_s = j_{s+1}$ if s is odd.

A switching chain is called *closed* if it has even length and $i_t = i_1$.

Example 1.1.6. In the picture below are two tomographically equivalent but distinct 3×3 binary matrices. The encircled positions form a switching chain of length 4. It is not closed.



Example 1.1.7. Let a and b be $m \times n$ binary matrices which differ by a 4-switch, then the sequence

$$(i_1, j_1), (i_2, j_1), (i_2, j_2), (i_1, j_2)$$

is a closed switching chain of length 4.

Lemma 1.1.8. *Let a and b be distinct but tomographically equivalent $m \times n$ binary matrices. Then there exists a closed switching chain for a and b of positive length.*

Proof. As a and b are distinct, there is some position (i, j) such that $a_{i,j}$ and $b_{i,j}$ are distinct. This position is a switching chain of positive length. As a chain consists of distinct positions, the length of a chain is bounded above by the number of positions. Hence there exist chains of maximal length and their length is positive. We will show such a chain is closed, by showing that a chain which is not closed can be extended to a longer chain.

Let

$$(i_1, j_1), \ldots, (i_t, j_t)$$

be a switching chain and suppose that t is odd. For $j = 1, \ldots, n$ define the sets

$$I_j^+ = \{i \mid a_{i,j} - b_{i,j} = 1\}$$

and

$$I_j^- = \{i \mid a_{i,j} - b_{i,j} = -1\}.$$

Note that these sets have the same cardinality, as the j-th column sum of a and b are the same. After possibly swapping a and b, we may assume that i_1 is in $I_{j_1}^+$. It follows that $i_{2s} \in I_{j_{2s}}^-$ and $i_{2s+1} \in I_{j_{2s+1}}^+$ for all s.

Let $S = \{s \mid j_{2s} = j_t\}$. Then the switching chain hits precisely $\#S$ distinct elements of $I^-_{j_t}$. Note that for every $s \in S$ we also have $j_{2s-1} = j_t$ and that additionally, t itself is also odd, so the switching chain hits $\#S + 1$ elements of $I^+_{j_t}$. Hence there is an element i_{t+1} of $I^-_{j_t}$ that is not hit by the switching chain. This means we can extend the switching chain with the position $(i_{t+1}, j_{t+1} = j_t)$.

If t is even but the chain is not closed, we can apply a similar counting argument to the rows of the matrices to find a point by which the chain can be extended. Thus every switching chain that is not closed can be extended into a longer chain and so a chain of maximal length must be closed. □

Theorem 1.1.9. (Ryser 1957) *Let a and b be tomographically equivalent $m \times n$ binary matrices. Then there is a sequence of $m \times n$ binary matrices $a_1, \ldots, a_t$ with $a_1 = a$ and $a_t = b$ such that a_i and a_{i+1} differ by a 4-switch for $i = 1, \ldots, t-1$.*

Proof. We proceed by induction on the number of positions in which a and b differ. If a and b are equal, then the theorem holds with $t = 1$ and $a = a_1 = b$.

If a and b are distinct, then by lemma (1.1.8) there is a closed switching chain of positive length for a and b. Let

$$(i_1, j_1), \ldots, (i_t, j_t)$$

be such a chain with minimal length. Note that $t \geq 4$, because if $t = 2$ we need to have $j_1 = j_2$ (as it is a switching chain) and $i_2 = i_1$ (as it is closed), but then $(i_1, j_1) = (i_2, j_2)$, violating the condition that a switching chain consists of distinct points. After possibly interchanging a and b, we may assume that $a_{i_2,j_2} - b_{i_2,j_2} = 1$. The points

$$(i_1, j_1), (i_2, j_2), (i_3, j_3)$$

are three corners of a rectangle, as $j_1 = j_2$ and $i_2 = i_3$. The fourth corner is (i_1, j_3). Note that we have

$$\begin{pmatrix} a_{i_2,j_2} & a_{i_3,j_3} \\ a_{i_1,j_1} & a_{i_1,j_3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$$

and

$$\begin{pmatrix} b_{i_2,j_2} & b_{i_3,j_3} \\ b_{i_1,j_1} & b_{i_1,j_3} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & y \end{pmatrix}.$$

If we have $x = 0$ and $y = 1$, then the sequence

$$(i_1, j_3), (i_4, j_4), \ldots, (i_t, j_t)$$

is a closed switching chain of smaller length than the original one. As such a chain does not exist, this cannot happen.

Hence we have either

$$\begin{pmatrix} a_{i_2,j_2} & a_{i_3,j_3} \\ a_{i_1,j_1} & a_{i_1,j_3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} b_{i_2,j_2} & b_{i_3,j_3} \\ b_{i_1,j_1} & b_{i_1,j_3} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This means we can do a 4-switch at these positions either in a or in b. Suppose this switch can be made in a, and call the resulting binary matrix c. Then c differs from a in 4 positions. In at least 3 of these, we changed the entry from a to the corresponding entry from b. It follows that the number of positions in which c and b differ is at least 2 fewer than the number of positions in which a and b differ. Hence by the induction hypothesis, there is a chain $c_1, \ldots c_t$ of binary matrices connecting c and b. We can then append a to this sequence, as a and c differ by a 4-switch, to obtain a sequence connecting a and b. The case where a switch can be made in b but not in a can be dealt with similarly. □

Corollary 1.1.10. *The $m \times n$ binary matrix a is uniquely determined by its row and column sums if and only if there are no $i_{1,2}$ and $j_{1,2}$ such that*

$$\begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} \\ a_{i_2,j_1} & a_{i_2,j_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We phrased the uniqueness question originally as taking the row and column sums as input, not a complete matrix as in the corollary above. However, this is not a problem as we shall see in corollary (1.1.14) that such a matrix can be reconstructed in polynomial time from the row and column sums.

Consistency and reconstruction

In 1957 Ryser [22] and Gale [9] independently gave necessary and sufficient conditions on numbers $r_1, \ldots, r_m$ and $c_1, \ldots c_n$ for them to be the row and column sums of an $m \times n$ binary matrix. We follow Ryser's proof here. To show that the conditions are sufficient one exhibits an algorithm which produces a matrix with the required row and column sums. Hence this result answers the reconstruction problem as well as the consistency problem.

Definition 1.1.11. Let $r = (r_1, \ldots, r_m)$ and $c = (c_1, \ldots, c_n)$ be sequences in $\mathbb{Z}_{\geq 0}$. The sequences r and c are *compatible* if

- for $i = 1, \ldots, m$ one has $r_i \leq n$;
- for $j = 1, \ldots, n$ one has $c_j \leq m$;
- $r_1 + \cdots + r_m = c_1 + \cdots + c_n$.

Lemma 1.1.12. *Let r and c be the row and column sums of a binary matrix, then r and c are compatible.*

Proof. Let $(a_{i,j})$ be an $m \times n$ binary matrix and let r and c be its row and column sums. Then we have

$$r_i = a_{i,1} + \cdots + a_{i,n},$$

which is bounded below by 0 and above by n as $a_{i,j} \in \{0, 1\}$ for all i and j. Likewise, c_j satisfies $0 \leq c_j \leq m$. Finally we have

$$\sum_{i=1}^{m} r_i = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{i,j} = \sum_{j=1}^{n} c_j. \qquad \square$$

Let a be an $m \times n$ binary matrix. Any permutation of the rows of a will result in a permutation of the row sums. It will not change the column sums. Conversely, any permutation of the columns will permute the column sums and leave the row sums invariant. It follows that we lose no generality if we impose an ordering on the row and column sums in what follows.

Theorem 1.1.13. *Let $r = (r_1, \ldots, r_m)$ and $c = (c_1, \ldots, c_n)$ be compatible sequences that are non-increasing (i.e. $r_1 \geq \cdots \geq r_m$ and $c_1 \geq \cdots \geq c_n$). Then there is a binary matrix whose row and column sums are r and c respectively if and only if*

$$\sum_{j=j_0}^{n} c_j \geq \sum_{j=j_0}^{n} r_j^* \qquad \text{for } 1 \leq j_0 \leq n,$$

where $r_j^* = \#\{i \mid r_i \geq j\}$.

Proof. Suppose that $(a_{i,j})$ is an $m \times n$ binary matrix whose row and column sums are the r_i and c_j respectively. We define another $m \times n$ binary matrix b by

$$b_{i,j} = \begin{cases} 0 & \text{if } j > r_i \\ 1 & \text{if } j \leq r_i \end{cases}.$$

Note that in row i of b, the first r_i entries are 1 and the rest is 0. In the corresponding row of a, some r_i entries are 1, but not necessarily the first ones. It follows that for every $1 \leq j_0 \leq n$ we have

$$a_{i,1} + \cdots + a_{i,j_0-1} \leq b_{i,1} + \cdots b_{i,j_0-1}$$

and so

$$a_{i,j_0} + \cdots + a_{i,n} \geq b_{i,j_0} + \cdots + b_{i,n}.$$

Let $1 \leq j_0 \leq n$. Summing the inequalities we have obtained for each row yields

$$\sum_{i=1}^{m} \sum_{j=j_0}^{n} a_{i,j} \geq \sum_{i=1}^{m} \sum_{j=j_0}^{n} b_{i,j}.$$

Note that for every j we have

$$c_j = \sum_{i=1}^{m} a_{i,j} \quad \text{and} \quad r_j^* = \sum_{i=1}^{m} b_{i,j},$$

hence by swapping the summation order in the previous inequality, we obtain

$$\sum_{j=j_0}^{n} c_j \geq \sum_{j=j_0}^{n} r_j^*.$$

This proves the 'only if' part.

We prove the 'if' part by induction on n, the number of columns. If $n = 1$ we have, by compatibility, that $0 \leq c_1 \leq n$ and $r_i \in \{0, 1\}$ for every i. Moreover $r_1 + \cdots + r_m = c_1$, so, as the sequence r is non-increasing, we must have $r_i = 1$ for $i = 1, \ldots, c_1$ and $r_i = 0$ for $i > c_1$. The $m \times 1$ binary matrix

$$a_{i,j} = \begin{cases} 0 & \text{if } i > c_1 \\ 1 & \text{otherwise} \end{cases}$$

yields the required row and column sums.

For $n > 1$ we will fill in the first column in such a way that the induction hypothesis produces the remaining $m \times (n-1)$ matrix. Let r and c be sequences that satisfy the hypotheses of the theorem. Note that

$$\sum_{j=1}^{n} r_j^* = \sum_{i=1}^{m} r_i$$

as both count the pairs (i, j) such that $r_i \geq j$. Hence

$$r_1^* + \sum_{j=2}^{n} r_j^* = \sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j = c_1 + \sum_{j=2}^{n} c_j$$

and so we have

$$r_1^* = c_1 + (\sum_{j=2}^{n} c_j - \sum_{j=2}^{n} r_j^*) \geq c_1.$$

As the sequence r is non-increasing, this implies that we have $r_1, \ldots, r_{c_1} \geq 1$.

Let $s = (s_1, \ldots, s_m)$ be the sequence given by $s_i = r_i - 1$ if $i \leq c_1$ and $s_i = r_i$ if $i > c_1$. From what we have just shown, s_i is non-negative for all i. I claim we also have $s_i \leq n - 1$ for all i. For suppose that $r_1 = \cdots = r_t = n$, then

$$nc_1 \geq c_1 + \cdots + c_n = r_1 + \cdots + r_m \geq tn$$

and so $c_1 \geq t$. Hence, if $r_i = n$, then $s_i = n - 1$. Also, if $r_i < n$, then $s_i \leq r_i \leq n - 1$.

Observe that we have

$$s_1 + \cdots + s_m = (r_1 + \cdots + r_m) - c_1 = c_2 + \cdots + c_n.$$

Let $d = (d_1, \ldots, d_{n-1})$ be the sequence given by $d_i = c_{i+1}$. Then we have just shown that the sequences s and d are compatible.

For $j = 1, \ldots, n - 1$, let $s_j^* = \#\{i \mid s_i \geq j\}$. Note that $r_i \geq j + 1$ implies $s_i \geq j$, hence we have $r_{j+1}^* \geq s_j^*$. It follows that for $j_0 = 1, \ldots, n - 1$ we have

$$\sum_{j=j_0}^{n-1} d_j = \sum_{j=j_0+1}^{n} c_j \geq \sum_{j=j_0+1}^{n} r_j^* \geq \sum_{j=j_0}^{n-1} s_j^*.$$

Let π be some permutation of $\{1, \ldots, m\}$ such that $\pi(s) = (s_{\pi(1)}, \ldots, s_{\pi(m)})$ is a non-increasing sequence. The discussion above proves that the induction hypothesis can be applied to the sequences $\pi(s)$ and d. Hence there is some $m \times (n - 1)$ binary matrix $(b_{i,j})$ whose row sums are $\pi(s)$ and whose column sums are d. Now let $(a_{i,j})$ be the $m \times n$ binary matrix given by

$$a_{i,j} = \begin{cases} 0 & \text{if } i > c_1 \text{ and } j = 1 \\ 1 & \text{if } i \leq c_1 \text{ and } j = 1 \\ b_{\pi^{-1}(i),j-1} & \text{otherwise} \end{cases}.$$

We finish by showing this binary matrix has the desired row and column sums. For $i = 1, \ldots, m$ we have

$$\sum_{j=1}^{n} a_{i,j} = a_{i,1} + \sum_{j=1}^{n-1} b_{\pi^{-1}(i),j} = a_{i,1} + s_i = r_i,$$

as $a_{i,1} = 1$ precisely when $s_i = r_i - 1$. For the column sums, it is clear that $c_1 = \sum_{i=1}^{m} a_{i,1}$. For $j = 2, \ldots, n$, we have

$$\sum_{i=1}^{m} a_{i,j} = \sum_{i=1}^{m} b_{\pi^{-1}(i),j-1} = \sum_{i=1}^{m} b_{i,j-1} = d_{j-1} = c_j$$

as required. This completes the proof. □

Corollary 1.1.14. *There is a polynomial time algorithm that given sequences $r_1, \ldots, r_m$ and $c_1, \ldots, c_n$ of non-negative integers produces either a binary matrix having these as its row and column sums or an error message if no such matrix exists.*

Proof. Begin by checking the sequences are compatible, if they aren't there is no matrix by lemma (1.1.12). Rearrange the sequences to be non-increasing and check that they satisfy the condition from theorem (1.1.13). If they don't, there is no matrix. If they do, the proof of the theorem gives an algorithm to compute the matrix column by column. The work per column is clearly polynomial in the number of rows and columns, hence the whole procedure is polynomial. □

Example 1.1.15. While the proof of theorem (1.1.13) is somewhat involved, the reconstruction algorithm that comes out of it is actually quite easy. It proceeds one column at a time, starting with a column whose sum is the highest and working in order down to the lowest column sum. In each column, we place the 1's in those rows where with the largest number of 1-positions left, that is, such that the row sum minus the number of 1's already filled in is the largest.
Consider for example the following row and column sums

	2	2	2	1	1
3					
2					
2					
1					

According to the algorithm, the first column must be filled in as follows.

	2	2	2	1	1
3	1				
2	1				
2	0				
1	0				

For the second column, we look at the number of 1-positions left in each row. The first row has 3 in total, of which 1 has already been used, so there's 2 left. Similarly, the second row has 1 left. The third row has 2 and the fourth has 1. Thus the two 1's in the second column go into the first and third row. Continuing on we can find the following solution.

	2	2	2	1	1
3	1	1	1	0	0
2	1	0	1	0	0
2	0	1	0	1	0
1	0	0	0	0	1

1.2 Measuring subsets of $\mathbb{Z}^r$

A more geometric way to look at the binary matrices from the previous section is to consider them as representing subsets of

$$\{1, \dots, n\} \times \{1, \dots, m\} \subset \mathbb{Z}^2.$$

The row and column sums of a matrix then count the number of points of the corresponding subset on each horizontal and vertical line. In this section we will generalise these notions by looking at different classes of finite subsets of $\mathbb{Z}^r$ and different sets of lines in $\mathbb{Z}^r$ to count along.

Definition 1.2.1. Let $r \geq 2$. A *(lattice) direction* in $\mathbb{Z}^r$ is a vector $v \in \mathbb{Z}^r$ that is non-zero. A direction $v = (v_1, \dots, v_r)$ is called *primitive* if $\gcd(v_1, \dots, v_r) = 1$.

Definition 1.2.2. Let $v \in \mathbb{Z}^r$ be a lattice direction and $x \in \mathbb{Z}^r$ be a point. Then the *lattice line* (or simply *line*) in direction v through x is the set

$$\ell_{v,x} = \{x + \lambda v \mid \lambda \in \mathbb{Z}\}.$$

Write $\mathcal{L}_v$ for the set of all lattice lines in direction v.

Definition 1.2.3. Let ℓ be a line in $\mathbb{Z}^r$ and let $X \subset \mathbb{Z}^r$ be finite. Then the *line sum* of X along ℓ is

$$p_\ell(X) = \#(X \cap \ell).$$

The three central questions posed in section 1 about binary matrices can also be stated in this more general context. Let $\mathcal{E}$ be a collection of finite subsets of $\mathbb{Z}^r$ and let $\mathcal{L}$ be a collection of lines in $\mathbb{Z}^r$. Suppose a function $p : \mathcal{L} \longrightarrow \mathbb{Z}_{\geq 0}$ is given.

1. **Consistency**.
   Is there an $e \in \mathcal{E}$ such that $p_\ell(e) = p(\ell)$ for all $\ell \in \mathcal{L}$?

2. **Uniqueness**.
   Is there at most one $e \in \mathcal{E}$ such that $p_\ell(e) = p(\ell)$ for all $\ell \in \mathcal{L}$?

3. **Reconstruction**.
   How can one construct an $e \in \mathcal{E}$ such that $p_\ell(e) = p(\ell)$ for all $\ell \in \mathcal{L}$?

The results from the previous section can be generalised to the case where $\mathcal{E}$ consists of all finite subsets of $\mathbb{Z}^2$ and $\mathcal{L}$ consists of all the lines in two independent directions in $\mathbb{Z}^2$.

Theorem 1.2.4. *Let v, w be independent directions in $\mathbb{Z}^2$. Let $\mathcal{E}$ be the collection of all finite subsets of $\mathbb{Z}^2$ and $\mathcal{L}$ the collection of all lines in $\mathbb{Z}^2$ in directions v and w. Then, keeping v and w fixed, there are polynomial time algorithms that given a function $p : \mathcal{L} \to \mathbb{Z}_{\geq 0}$ decide whether there is an $e \in \mathcal{E}$ with $p_\ell(e) = p(\ell)$ for all $\ell \in \mathcal{L}$, find such an e, and decide if it is unique.*

Proof. First, suppose that v and w together generate $\mathbb{Z}^2$. Then there is a linear transformation $\mathbb{Z}^2 \to \mathbb{Z}^2$ sending v and w to $(0, 1)$ and $(1, 0)$. This transformation also maps $\mathcal{L}$ to the set of horizontal and vertical lines. This reduces the situation to one considered in the previous section, where we already saw how consistency, reconstruction and uniqueness can be decided in polynomial time.

If v and w do not span $\mathbb{Z}^2$, they span a subgroup L of $\mathbb{Z}^2$. As v and w are independent, L is of finite index in $\mathbb{Z}^2$. Note that any lattice line in direction v or w that has a point in common with L, lies entirely within L. This means

we can decompose the original problem into $[\mathbb{Z}^2 : L]$ sub-problems consisting of a translate of L and the lines going through that translate. Since v and w generate L, we can apply the trick we used before to reduce each of these sub-problems to a problem about binary matrices. □

When lines in three or more directions are considered, the computational complexity of the situation is a lot worse. In [10], Gardner, Gritzmann and Prangenberg give extensive complexity results in these cases. We will not go into these results here, except to state the following theorem as a counterpoint to theorem (1.2.4).

Theorem 1.2.5. *Let v_1, v_2 and v_3 be pairwise independent directions in $\mathbb{Z}^2$. Let $\mathcal{E}$ be the collection of all finite subsets of $\mathbb{Z}^2$ and $\mathcal{L}$ the collection of all lines in $\mathbb{Z}^2$ in directions v_1, v_2 and v_3. Then, for any fixed v_1, v_2 and v_3, the consistency and uniqueness problems are NP-complete and the reconstruction problem is NP-hard.*

1.3 Another generalisation

In [12], Hajdu and Tijdeman consider a generalisation of the problems discussed in the previous section.

A finite subset S of some fixed set $A \subset \mathbb{Z}^2$ can be represented as a function

$$f_S : A \longrightarrow \{0,1\} \qquad a \mapsto \begin{cases} 1 & \text{if } a \in S \\ 0 & \text{otherwise} \end{cases}.$$

In other words, there is a bijection between finite subsets of A and functions $f : A \to \{0,1\}$ that are 0 almost everywhere, i.e., such that $f(a) = 0$ for all but finitely many $a \in A$.

Line sums in this context can be represented by actual sums, as we have

$$p_\ell(S) = \#(S \cap \ell) = \sum_{a \in A \cap \ell} f_S(a)$$

for every finite subset S of A and every lattice line ℓ.

Hajdu and Tijdeman propose the following relaxation of the problem. Look at functions

$$f : A \longrightarrow \mathbb{Z}$$

that are 0 almost everywhere. For such functions we can extend the definition of line sums to

$$p_\ell(f) = \sum_{a \in A \cap \ell} f(a).$$

In this setup, we can ask the same central questions we considered in the previous sections. Let A be a subset of $\mathbb{Z}^r$ and let $\mathcal{L}$ be a collection of lines in $\mathbb{Z}^r$. Suppose a function $p : \mathcal{L} \longrightarrow \mathbb{Z}$ is given.

1. **Consistency**.
   Is there an $f : A \to \mathbb{Z}$ that is 0 almost everywhere such that $p_\ell(f) = p(\ell)$ for all $\ell \in \mathcal{L}$?

2. **Uniqueness**.
   Is there at most one $f : A \to \mathbb{Z}$ that is 0 almost everywhere such that $p_\ell(f) = p(\ell)$ for all $\ell \in \mathcal{L}$?

3. **Reconstruction**.
   How can one construct an $f : A \to \mathbb{Z}$ that is 0 almost everywhere such that $p_\ell(f) = p(\ell)$ for all $\ell \in \mathcal{L}$?

One may ask what results about this relaxed version of the problem tell us about the problems from the previous section, for functions $f : A \to \{0, 1\}$. There is one trivial implication for the consistency problem: if there are no solutions $f : A \to \mathbb{Z}$ then there are no solutions $f : A \to \{0, 1\}$. But there is an even stronger connection, as described in the next lemma.

Lemma 1.3.1. *Let A be a subset of $\mathbb{Z}^2$ and D a non-empty sequence of pairwise independent primitive directions in $\mathbb{Z}^2$. Let $\mathcal{L}$ be the set of lattice lines in directions from D that meet A. For a function $f : A \to \mathbb{Z}$ that is* 0 *almost everywhere, the* weight *of f is*

$$W(f) = \sum_{x \in A} f(x)^2.$$

Let $f : A \to \{0, 1\}$ be 0 *almost everywhere and suppose we have a function $g : A \to \mathbb{Z}$ that is* 0 *almost everywhere such that $p_\ell(f) = p_\ell(g)$ for all $\ell \in \mathcal{L}$. Then $W(g) \geq W(f)$ with equality if and only if* $\mathrm{im}(g) \subset \{0, 1\}$.

The lemma shows that if there are functions $A \to \{0, 1\}$ with given line sums, then they are those functions $A \to \mathbb{Z}$ which have minimal weight. We shall not prove this lemma here. Corollary (2.1.13) is a more general version of this statement.

What is gained by considering this relaxation of the problem? It has a lot of additional structure compared to the original. We can add functions

$f : A \to \mathbb{Z}$ and $g : A \to \mathbb{Z}$ together pointwise to obtain a new function

$$f + g : A \longrightarrow \mathbb{Z}.$$

The line sum maps respect this addition, i.e. they satisfy

$$p_\ell(f + g) = p_\ell(f) + p_\ell(g).$$

This shifts the emphasis of techniques for dealing with these problems away from combinatorics and towards algebra.

Uniqueness

Most of the work in proving Ryser's theorem (1.1.9) goes into finding a position where a 4-switch can be applied. This is also the primary thing that fails when trying to prove an analogue of this theorem for three or more directions.

When considering functions $f : A \to \mathbb{Z}$ things become a lot easier. First of all, note that we have

$$p_\ell(f) = p_\ell(g) \quad \Longleftrightarrow \quad p_\ell(f - g) = 0.$$

It follows that the line sums of f uniquely determine f if and only if the only function $g : A \to \mathbb{Z}$ such that $p_\ell(g) = 0$ for all lines $\ell \in \mathcal{L}$ is $g = 0$. In particular, this condition does not depend on f or its line sums, just on the choice of A and $\mathcal{L}$.

For rectangular A, i.e. $A = \{1, \ldots, n\} \times \{1, \ldots, m\}$, Hajdu and Tijdeman in [12] give a complete description of the maps $g : A \to \mathbb{Z}$ that have 0 line sums for all lines ℓ in a given set of primitive directions in $\mathbb{Z}^2$.

Before we state this theorem, we introduce one convenient notation. Given a function $f : \mathbb{Z}^2 \to \mathbb{Z}$ and $s \in \mathbb{Z}^2$ we define a function $f_s : \mathbb{Z}^2 \to \mathbb{Z}$ by $f_s(x) = f(x - s)$ for all $x \in \mathbb{Z}^2$.

Theorem 1.3.2. (Hajdu and Tijdeman, 2001) *Let m and n be positive integers and let $d_1, \ldots, d_k$ be a sequence of pairwise independent primitive directions in $\mathbb{Z}^2$. Let A be the set $\{1, \ldots, n\} \times \{1, \ldots, m\}$ and let $\mathcal{L}$ be the set of lattice lines in directions $d_1, \ldots, d_k$ that meet A. Then there are an explicit function $f : \mathbb{Z}^2 \to \mathbb{Z}$ and an explicit finite set $S \subset \mathbb{Z}^2$ such that for all $s \in S$ we have $f_s(x) = 0$ whenever $x \notin A$ and $p_\ell(f_s) = 0$ for all $\ell \in \mathcal{L}$. Moreover, any function $g : A \to \mathbb{Z}$ such that $p_\ell(g) = 0$ for all $\ell \in \mathcal{L}$ is an integer linear combination of the f_s with $s \in S$.*

A priori, it is not even clear that a function $f : \mathbb{Z}^2 \to \mathbb{Z}$ can be written down in a finite amount of time, but one easily sees that in the case of this theorem f takes non-zero values only in a finite number of points, contained in some translate of $\{1, \ldots, n\} \times \{1, \ldots, m\}$ (unless $S = \emptyset$, but in that case one should take $f = 0$).

We will not prove this theorem here, but merely remark it is a special case of theorem (3.2.8). The proof of Hajdu and Tijdeman is also very similar to the proof we shall give of that theorem. Their proof also shows that the same result holds for functions $g : A \to \mathbb{Q}$ of which all the line sums are 0, with the same f and the same S. This will be important for their reconstruction algorithm, which we shall describe next.

Consistency and reconstruction

Hajdu and Tijdeman in [12] deal with the consistency and reconstruction issues at the same time, by giving a reconstruction algorithm that will also detect if no reconstruction is possible. The central trick in their proof, of which we present an adaptation here, is to solve the problem with rational coefficients and then use theorem (1.3.2) to get rid of all non-integral coefficients.

Theorem 1.3.3. *For positive integers m, n and a non-empty sequence D of pairwise independent primitive directions in $\mathbb{Z}^2$, let A be the set*

$$\{1, \ldots, n\} \times \{1, \ldots, m\}$$

and let $\mathcal{L}$ be the set of lattice lines in directions from D that meet A. Then there is an algorithm that given m, n, D, and $c_\ell \in \mathbb{Z}$ for every $\ell \in \mathcal{L}$ produces, in polynomial time in the input length, a function $g : A \to \mathbb{Z}$ that satisfies $p_\ell(g) = c_\ell$ for all $\ell \in \mathcal{L}$ or outputs an error is no such function exists.

Proof. Note that when $\#D \geq 2$, the set $\mathcal{L}$ has at least $m + n$ elements: at least one direction is not horizontal and so every point in the first row lies in a different line in this direction and similarly, at least one direction is not vertical and so every point in the first column lies in a different line in this direction. It follows that the size of the input is at at least linear in $n + m$. When $\#D = 1$, the solution of the reconstruction problem is to give a point on each line the value of the line sum for the line. This solution can certainly be output in time polynomial in the input length.

The map sending a function $g : A \to \mathbb{Q}$ to the vector $(p_\ell(g))_{\ell \in \mathcal{L}}$ is a map between finite dimensional vector spaces over $\mathbb{Q}$. Finding a g that maps

to a particular vector thus comes down to solving a matrix equation of the from $Mx = b$, where the matrix M and vector b are known, or rather, can be computed in polynomial time from the input. There is much theory on computational linear algebra that tells one how to do this in time polynomial in the size of the matrix, and the size of the matrix is clearly polynomial in m, n and $k = \#D$ (see e.g. [6, Ch. 2]).

Note that any solution $g : A \to \mathbb{Z}$ is also a solution $A \to \mathbb{Q}$, so if the linear algebra problem we just described has no solutions over $\mathbb{Q}$, we know the reconstruction problem over $\mathbb{Z}$ also doesn't have a solution and we output that no solution exists. Otherwise, we may assume we have found a solution $g_0 : A \to \mathbb{Q}$. We will then use our description of the functions that map to the zero vector to try to modify our solution g_0 into one that takes integral values.

Let f and S be as in theorem (1.3.2). If $S = \emptyset$ or $f = 0$, that theorem states that the map sending a function g to the vector $(p_\ell(g))_{\ell \in \mathcal{L}}$ is injective. Thus g_0 is the only function $A \to \mathbb{Q}$ that has the given line sums. As remarked before, any integral solution would also be a rational solution, so if there is an integral solution it is g_0. Thus we check if g_0 takes only integral values (this can clearly be done in polynomial time). If so, we output g_0, if not, we output that no solution exists.

If f is non-zero and S is non-empty, the set of points (x, y) such that $f(x, y) \neq 0$ is finite and contains at most mn elements, as was remarked immediately after theorem (1.3.2). Therefore, one of these points, call it s_0, is lexicographically first among them, and such a point can be found in polynomial time, for example by sorting (see e.g. [15]).

Order the points in S lexicographically and label them $s_1, \ldots, s_t$ in order. Just as the number points such that $f(x, y) \neq 0$ is bounded by mn, so is the number of elements of S. Let $1 \leq i \leq j \leq t$ and consider

$$f_{s_j}(s_0 + s_i).$$

If $i = j$, this value is $f(s_0)$ which is non-zero by construction of s_0. If $i < j$, $s_0 + s_i$ comes lexicographically before $s_0 + s_j$ and so $s_0 + (s_i - s_j)$ comes lexicographically before s_0 and we have $f_{s_j}(s_0 + s_i) = f(s_0 + (s_i - s_j)) = 0$ as s_0 is the lexicographically first element (x, y) such that $f(x, y) \neq 0$.

Thus the map sending a linear combination

$$h = \lambda_1 f_{s_1} + \cdots + \lambda_t f_{s_t}$$

to the vector

$$(h(s_0 + s_1), \ldots, h(s_0 + s_t)) \in \mathbb{Q}^t$$

is a bijection (as the matrix describing it is upper-triangular). As theorem (1.3.2) tells us, these linear combinations correspond precisely to the functions whose line sums are all 0. So any solution $g : A \to \mathbb{Q}$ of our reconstruction problem is of the form $g_0 + h$. In this way, we see that there is a bijection between the set of solutions $g : A \to \mathbb{Q}$ and $\mathbb{Q}^t$ sending g to $(g(s_0 + s_i))_{i=0}^t$. In particular, there is a unique solution $g : A \to \mathbb{Q}$ such that $g(s_0 + s_i) = 0$ for $i = 1, \ldots, t$.

It is clear that we can construct g from g_0 in polynomial time: first make $g_1 = g_0 + \lambda_1 f_{s_1}$ such that $g_1(s_0 + s_1) = 0$, then $g_2 = g_1 + \lambda_2 f_{s_2}$ such that $g_2(s_0 + s_2) = 0$, etc. up to $g = g_t$.

Now suppose that there is a solution $g'_0 : A \to \mathbb{Z}$. The same argument as before shows that there is a bijection between integral solutions $g' : A \to \mathbb{Z}$ and $\mathbb{Z}^t$ sending g' to $(g'(s_0 + s_i))_{i=1}^t$. It follows that there is a unique solution $g' : A \to \mathbb{Z}$ such that $g'(s_0 + s_i) = 0$ for $i = 1, \ldots, t$. Note that g' is also a solution over $\mathbb{Q}$ satisfying this condition, and so by the uniqueness of g, we must have $g = g'$ and the solution g is indeed integral.

In other words, if there is an integral solution, the map g we have constructed is such an integral solution. Hence we simply check if g is integral (which can again clearly be done in polynomial time). If so we output g, if not we output that no solution exists. □

With some extra work, Hajdu and Tijdeman also show that a function can be found such that its values are in a precise sense 'not too big'. As we argued before in lemma (1.3.1), solutions $A \to \{0, 1\}$ are the smallest possible solutions. Theorem (1.2.5) shows there is very little hope for an efficient algorithm to find these smallest solutions. Thus an efficient algorithm that finds a small (but not necessarily smallest) solution is a noteworthy achievement.

The algorithm from theorem (1.3.3) has to do two tests for the existence of a solution. First there is the linear algebra problem over $\mathbb{Q}$ that may or may not have a solution. If there is a rational solution, we then try to make it into an integer solution, which may again fail.

A consequence of the structure results we prove in chapter 3, specifically corollary (3.2.10), is that this second step will never fail. The only obstruction to the existence of a solution is the linear-algebraic one. This does not follow from the results contained in [12].

The linear-algebraic obstruction is non-trivial in most cases. We will describe it briefly here, following [12]. The linear-algebraic nature of this discussion means it is more convenient to consider functions with rational values.

Example 1.3.4. Suppose we consider a rectangular domain

$$\{1, \ldots, n\} \times \{1, \ldots, m\}$$

and projections in the directions $(1, 0)$ and $(0, 1)$. This is the analogue of the binary matrices we considered in the first section: we are looking at an $m \times n$ matrix of integers and its row and column sums.

Starting from such a matrix, it is clear that the sum of its row sums is the sum of all entries in the matrix, as is the sum of its column sums. Thus we find a linear-algebraic relation between the row and column sums of the matrix: the sum of former has to be equal to the sum of the latter. One checks readily that this condition is the only one. A problem of the form

	c_1	c_2	$\cdots$	c_n
r_1				
r_2				
$\vdots$				
r_m				

has a solution of the form

	c_1	c_2	$\cdots$	c_n
r_1	x	c_2	$\cdots$	c_n
r_2	r_2	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
r_m	r_m	0	$\cdots$	0

if and only if $r_1 + \cdots r_m = c_1 + \cdots + c_n$. Note that we saw this same condition already in section 1.1 in the definition of *compatible sequences*, (1.1.11).

Definition 1.3.5. Let m and n be positive integers and let $d_1, \ldots, d_k$ be a sequence of pairwise independent primitive directions in $\mathbb{Z}^2$. Let A be the

set $\{1,\ldots,n\}\times\{1,\ldots,m\}$ and let $\mathcal{L}$ be the set of lattice lines in directions $d_1,\ldots,d_k$ that meet A. A *dependency* is a vector $(a_\ell)_{\ell\in\mathcal{L}}$ of rational numbers such that for any function $f : A \to \mathbb{Q}$ we have

$$\sum_{\ell\in\mathcal{L}} a_\ell p_\ell(f) = 0.$$

It is clear that these dependencies form a vector space. In their paper, Hajdu and Tijdeman compute the dimension of this space, under some mild conditions. They also give the following example, where the space of dependencies has dimension 7.

Example 1.3.6. Consider a rectangular domain $A = \{1,\ldots,n\}\times\{1,\ldots,m\}$ and line sums in the directions $(1,0)$, $(0,1)$, $(1,1)$ and $(1,-1)$, i.e. horizontal, vertical, diagonal and anti-diagonal. For a function $f : A \to \mathbb{Q}$ these line sums are

$$\begin{aligned}
r_j &= \sum_{i=1}^{n} f(i,j) && 1\le j\le m && \text{the row sums,}\\
c_i &= \sum_{j=1}^{m} f(i,j) && 1\le i\le n && \text{the column sums,}\\
s_k &= \sum_{\substack{j=i+k\\(i,j)\in A}} f(i,j) && 1-n\le k\le m-1 && \text{the diagonal sums,}\\
t_k &= \sum_{\substack{i+j=k\\(i,j)\in A}} f(i,j) && 2\le k\le m+n && \text{the anti-diagonal sums.}
\end{aligned}$$

The following seven dependencies hold for these line sums.

$$\sum_{j=1}^{m} r_j = \sum_{i=1}^{n} c_i = \sum_{k=1-n}^{m-1} s_k = \sum_{k=2}^{m+n} t_k$$

$$\sum_{\substack{k=1-n\\2|k}}^{m-1} s_k = \sum_{\substack{k=2\\2|k}}^{m+n} t_k$$

$$\begin{aligned} -\sum_{j=1}^{m} jr_j + \sum_{i=1}^{n} ic_i &= \sum_{k=1-n}^{m-1} ks_k \\ \sum_{j=1}^{m} jr_j + \sum_{i=1}^{n} ic_i &= \sum_{k=2}^{m+n} kt_k \\ 2\sum_{j=1}^{m} j^2 r_j + 2\sum_{i=1}^{n} i^2 c_i &= \sum_{k=1-n}^{m-1} k^2 s_k + \sum_{k=2}^{m+1} k^2 t_k \end{aligned}$$

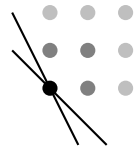
If m and n are sufficiently large, these dependencies are linearly independent. Counting dimensions, Hajdu and Tijdeman conclude in [12] that these then form a basis of the $\mathbb{Q}$-vector space of dependencies.

A striking feature of these dependencies is that the weights assigned to each line always seem to be polynomials in the numerical index of the line. We also see that higher degree polynomials appear in dependencies that involve lines in more directions. In one of the dependencies (on the second line) we see a congruence condition appear in the weights (this dependency only involves lines whose index is even).

Example 1.3.7. Once more, consider a rectangular domain

$$A = \{1, \ldots, n\} \times \{1, \ldots, m\}.$$

Taking line sums in the directions $(1, -1)$ and $(1, -2)$, something interesting happens in the lower-right and upper-left corners of the rectangle A. As can be seen in the following picture, the line through $(1, 1)$ in either direction does not go through any other points of A.



It follows that these two lines always have the same line sum, namely, the function value at $(1, 1)$. This is another example of a dependency. It is of a very different nature than the dependencies we saw in example (1.3.6) above.

Intuitively, there is a clear distinction between the dependencies that only involve a few points in the corner of A and dependencies that involve many lines going through points throughout A. We call the former 'local' and the latter 'global' dependencies. This distinction is of course not very rigorous. In chapter 3 we will make it more precise, leading up to corollary (3.5.10), which for convex sets A in the plane shows that the dependencies indeed decompose in a global part that doesn't depend on the shape of A and a local part that involves only lines in certain 'corners' of A.

A systematic discussion of dependencies for rectangular domains was attempted by van Dalen in her 2007 Master's thesis [7]. Her approach is to construct explicit dependencies and show that they are independent. The dimension of the space of dependencies is known from the work of Hajdu and Tijdeman [12], so one knows when one has found a maximal independent set.

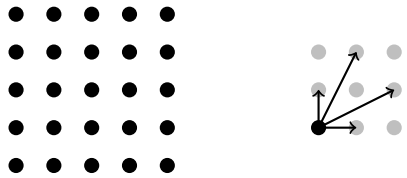
Van Dalen gives a conjecture for the dimensions of the spaces of global and local dependencies. For the local dependencies, she goes on to construct for every D and sufficiently large A an independent set of dependencies with the conjectured dimension. For the global dependencies, she constructs an independent set of the conjectured dimension in the cases where D has at most 4 elements.

Our work in chapter 3 is largely complementary to that of van Dalen. We prove the dimension of the space of global dependencies is as conjectured. Moreover, we show that the complement of the global dependencies involves only line sums in the 'corners' of the set A. The local dependencies constructed by van Dalen are precisely of this form. Our construction of the global dependencies in chapter 4 supersedes the work of van Dalen in that it gives a construction for all sequences D. Our approach is somewhat different in that we will construct a generating set of the space of dependencies, rather than an independent set.

1.4 Some examples

In this section we will introduce four examples of reconstruction problems similar to the ones described in the previous section. We will come back to these examples several times in the upcoming chapters to see how they tie into the theory we describe there. As studying these objects is the primary motivation for the work we will do, it is good to keep an eye on these concrete examples.

Example 1.4.1. The first example comes directly from section 1.3. We consider the square grid $A = \{1, \ldots, 5\} \times \{1, \ldots, 5\}$ and projection directions $(0, 1)$, $(1, 2)$, $(2, 1)$ and $(0, 1)$.



Using theorem (1.3.2), one shows that all functions $f : A \to \mathbb{Z}$ for which all the line sums in these directions are zero, are multiples of the following function.

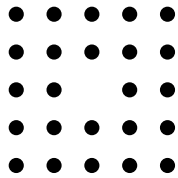
$$\begin{array}{rrrrr} 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 1 & -2 & 1 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{array}$$

We will revisit this example in chapter 3 to see how one can derive this (see example (3.2.9)).

To compute the number of independent dependencies we expect, we consider the reconstruction problem for functions $f : A \to \mathbb{Q}$ and count the dimensions of the vector spaces involved. The space of functions has dimension $\#A = 25$. The number of line sums is 5 in each of the directions $(0, 1)$ and $(1, 0)$ and 13 in each of the directions $(1, 2)$ and $(2, 1)$. This leads to a total number of 36 line sums. As remarked before, the map sending a function to its line sums is linear. Its kernel, the functions that have all line sums zero, has dimension 1 as we have just seen. That means the image of the map will have dimension $25 - 1 = 24$ inside a 36-dimensional space. Hence there will be 12 independent dependencies.

Two of these dependencies are of the 'local' type described in example (1.3.7). The upper left point of A is the only point in A on a particular line in direction $(1, 2)$ and a particular line in direction $(2, 1)$. The same is true for the lower right point of A. These two dependencies are clearly independent, thus we are left with a 10-dimensional space of dependencies that have not yet been accounted for. We shall see in chapter 3 how these relate to the dependencies in example (1.4.3) below.

Example 1.4.2. A slight variation of the previous example is obtained by removing the central point from the set A, while keeping the same line sum directions. Thus we have $B = \{1, \ldots, 5\} \times \{1, \ldots, 5\} \setminus \{(3, 3)\}$, or in a picture:



What makes this example behave significantly different from the previous one, is that this set is not a convex grid set, because there is a 'hole' in it.

Keeping the previous example in mind, one sees easily that there are no non-zero functions $f : B \to \mathbb{Q}$ that have all line sums 0. The dimension count for the number of dependencies again yields 12 independent dependencies, as the rank of the domain and the kernel each decrease by 1. What is noteworthy about this example is that there is a difference between the consistency problem for rational solutions (which is entirely governed by the dependencies) and the consistency problem for integral solutions.

Consider the following function $B \to \mathbb{Z}$:

$$\begin{array}{rrrrr} 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 1 & & 1 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 \end{array}$$

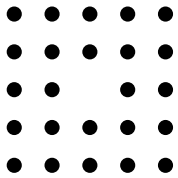
Its line sums are all 0, with the exception of the four lines that pass through the (absent) middle point, which each have line sum 2. As the line sum map is injective, this is the unique map $B \to \mathbb{Z}$ having these line sums. But all these line sums are even, so we can divide them all by 2. This yields a collection of integral line sums. There is clearly a map $B \to \mathbb{Q}$ yielding these line sums, which one obtains by dividing all the entries in the table above by 2. Again, as the line sum map is injective, this is the unique map $B \to \mathbb{Q}$ having these line sums. But clearly, this function does not take integral values. Thus we have a set of integral line sums for which there is a rational solution, but not an integral solution.

Example 1.4.3. For dependencies such as the ones described in example (1.3.6), it seems the shape of the set A is more or less irrelevant. If we want to study such dependencies, it makes sense to simply remove all restrictions on the shape and study general functions $f : \mathbb{Z}^2 \to \mathbb{Z}$. In order to make sense of the line sums, we need to impose the restriction that $f(x, y) = 0$ for almost all $(x, y) \in \mathbb{Z}^2$. That brings us to our third example. We consider functions $\mathbb{Z}^2 \to \mathbb{Z}$ and their line sums in the same four directions as the previous examples: $(0, 1)$, $(1, 2)$, $(2, 1)$ and $(0, 1)$.

We shall consider reconstruction problems of this nature extensively in chapters 3 and 4. It turns out that the functions whose line sums are all zero have a description very similar to that of theorem (1.3.2). There is a single function $\mathbb{Z}^2 \to \mathbb{Z}$ such that the functions with all line sums zero are precisely the linear combinations of translates of this function. Not surprisingly, this function turns out to be the one we encountered already in example (1.4.1). See corollary (3.1.3) and example (3.1.4) for the precise results.

When we consider the corresponding problem with rational coefficients, we now encounter infinite dimensional vector spaces and cannot do a simple dimension computation to figure out the number of dependencies. Indeed, there is no reason to expect this number is finite. Yet, this turns out to be the case, as we shall see in chapter 3 (see example (3.5.2)). In chapter 4, we will explore the structure of these dependencies in great detail. As a result of this, we can give explicit generators of the space of dependencies. We will also show that in this case, the dependencies once again form the only obstruction to the reconstruction problem for $\mathbb{Z}$-valued functions.

Example 1.4.4. The fourth and last example in this section can be considered as a combination of the previous two. We begin with the full grid $\mathbb{Z}^2$ as in the previous example and then introduce some 'holes' in it. We could just make a single hole (in the origin, say), but rather we chose to make infinitely many holes, in a periodic fashion. We consider the subset $A = \mathbb{Z}^2 \setminus 3\mathbb{Z}^2$. In this set a neighbourhood of any of the holes looks like



We remember this from example (1.4.2). However, unlike in that example, there is in this case no difference between the reconstruction problem over $\mathbb{Q}$ and that over $\mathbb{Z}$. We will look at this example in detail in chapter 5.

2 – Reconstruction systems

In this chapter we develop from the ground up an algebraic context in which we can study problems similar to the ones considered in sections 1.3 and 1.4. The purpose is to provide a solid algebraic basis that can be used in the upcoming chapters. Most of the results in the current chapter are therefore fairly straightforward, many are special cases of well-known results from the algebraic literature.

The first section of this chapter introduces the objects of interest and describes their relation to the material from the previous chapter. This is very important to understand the relation of the work in upcoming chapters to the problems described in the first chapter.

Sections 2.2 and beyond go into the algebraic structure of the objects we consider. Some basic familiarity with modern algebra is assumed. Care has been taken to minimise this assumed knowledge. Many notions that will be used shall also be explained, albeit briefly, in the text. We advise the reader not to spend too much time on these sections, but to treat them mostly as reference material. Each section begins with an introduction that highlights the most important results it contains and explains briefly where they shall be used in the remainder of the work.

2.1 Definitions and examples

Definition 2.1.1. Let k be a commutative ring. A *reconstruction system* over k is a triple (T, P, p) where T and P are k-modules and $p : T \to P$ is a k-linear map.

The elements of T and P are referred to as 'tables' (or 'images') and 'projections' respectively. These names of course are meant to emphasise the connection with discrete tomograpy. By a slight abuse of notation we will often refer to the reconstruction system (T, P, p) as $p : T \to P$ or even just as p.

The notion of a reconstruction system is very general. While the current chapter deals with them in this full generality, later chapters shall focus on particular types of systems. To give the reader some motivation for studying reconstruction systems and some concrete examples to keep in mind, we will now look at how the objects studied in section 1.3 can be viewed as reconstruction systems.

Remark 2.1.2. Recall the definitions of (lattice) directions and (lattice) lines, (1.2.1) and (1.2.2) in the previous chapter. Let $r \geq 2$, let $p, q \in \mathbb{Z}^r$ be two points and d a direction (i.e. non-zero vector) in $\mathbb{Z}^r$. Then the following are equivalent

1. p and q lie on the same line in direction d;
2. $p - q = \lambda d$ for some $\lambda \in \mathbb{Z}$;
3. $\bar{p} = \bar{q}$ holds in $\mathbb{Z}^r/\mathbb{Z} \cdot d$.

Thus the set of lines in direction d is parametrised by the quotient $\mathbb{Z}^r/\mathbb{Z} \cdot d$. For convenience of notation we shall usually write $\mathbb{Z}^r/d$ for $\mathbb{Z}^r/\mathbb{Z}d$.

Definition 2.1.3. Let n be a non-negative integer and let $D = (d_1, \ldots, d_n)$ be an n-tuple of directions in $\mathbb{Z}^r$. We call D *primitive* if each d_i is primitive. Let t be an integer such that $2 \leq t \leq r$. We call D *t-regular* if any t elements of D are independent in $\mathbb{Z}^r$.

Remark 2.1.4. Let k be a commutative ring. For a set X, write $k[X]$ for the set

$$k[X] = \{f : X \to k \mid f(x) = 0 \text{ for almost all } x \in X\},$$

endowed with the structure of a k-module in the obvious way. By slight abuse of notation we often identify an element $x \in X$ with the function $e_x : X \to k$ given by $e_x(y) = 1$ if $y = x$ and $e_x(y) = 0$ otherwise. It is clear that any element of $k[X]$ can be written in a unique way as a k-linear combination of functions e_x with $x \in X$. Thus the e_x (or simply the elements of X) form a basis for the free k-module $k[X]$.

Example 2.1.5. We have seen these objects several times already in sections 1.3 and 1.4. For various $A \subset \mathbb{Z}^2$ we considered functions $A \to \mathbb{Z}$ and $A \to \mathbb{Q}$ that were 0 almost everywhere. Using the notation above, we write $\mathbb{Z}[A]$ for the $\mathbb{Z}$-module (i.e. abelian group) of functions $A \to \mathbb{Z}$ that are 0 almost everywhere and $\mathbb{Q}[A]$ for the $\mathbb{Q}$-module (i.e. $\mathbb{Q}$-vector space) of functions $A \to \mathbb{Q}$ that are 0 almost everywhere.

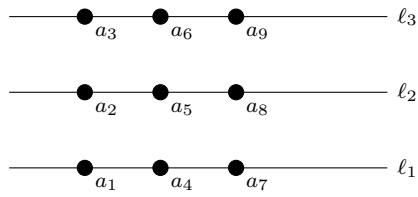
Remark 2.1.6. Let A be a subset of $\mathbb{Z}^r$ and d a direction in $\mathbb{Z}^r$. The image of A inside $\mathbb{Z}^r/d$ parametrises the set of all lines in direction d that pass

through some point of A. Write A_d for this image. Let k be a commutative ring. There is an obvious k-linear map

$$k[A] \longrightarrow k[A_d]$$

that sends an element of A (viewed as a basis element of $k[A]$) to the corresponding element of A_d (viewed as a basis element of $k[A_d]$). This map identifies elements of A that are on the same line in direction d and thus adds together the coefficients at these elements. Hence this map takes simultaneously the line sums for all lines in direction d that pass through A.

Example 2.1.7. Consider the set $A = \{1, 2, 3\} \times \{1, 2, 3\}$ and $d = (1, 0)$.



The quotient group $\mathbb{Z}^2/d$ is isomorphic to $\mathbb{Z}$; an explicit isomorphism is given by sending the pair (x, y) to y. We see this identifies points that lie on the same horizontal line. The set A_d under this isomorphism is $\{1, 2, 3\}$. Labeling the points and lines as in the picture above, we see that a_1, a_4 and a_7 get sent to ℓ_1. Likewise, a_2, a_5 and a_8 are sent to ℓ_2 and a_3, a_6 and a_9 to ℓ_3. The map $k[A] \to K[A_d]$ is therefore given by

$$c_1a_1 + \cdots + c_9a_9 \longmapsto (c_1 + c_4 + c_7)\ell_1 + (c_2 + c_5 + c_8)\ell_2 + (c_3 + c_6 + c_9)\ell_3.$$

This is indeed the map that associates to a function $A \to k$ its row sums.

Definition 2.1.8. Let $r \geq 2$, n a non-negative integer, $D = (d_1, \ldots, d_n)$ a sequence directions in $\mathbb{Z}^r$ and $A \subset \mathbb{Z}^r$. Let k be a commutative ring. Then the *grid reconstruction system* over k associated to (A, D) is the natural map

$$p : k[A] \longrightarrow \bigoplus_{i=1}^{n} k[A_i],$$

where A_i is the image of A inside $\mathbb{Z}^r/d_i$.

Example 2.1.9. In section 1.3, we looked a functions $A \to \mathbb{Z}$ where

$$A = \{1, \ldots, n\} \times \{1, \ldots, m\}$$

is a rectangular subset of $\mathbb{Z}^2$. The set of all these functions, with pointwise addition and scalar multiplication is the $\mathbb{Z}$-module $\mathbb{Z}[A]$. As remarked in (2.1.6), taking all the line sums in a particular direction $d \in \mathbb{Z}^2$, corresponds to looking at the map $\mathbb{Z}[A] \to \mathbb{Z}[A_d]$.

If D is a sequence of pairwise independent, primitive directions $d_1, \ldots, d_n$, then the grid reconstruction system over $\mathbb{Z}$ associated to (A, D) is the map that associates to a function $A \to \mathbb{Z}$ its line sums in all the lines in directions in D that meet A.

Example 2.1.10. All of the examples we described in section 1.4 are grid reconstruction systems. In every case, D is the 4-tuple

$$((0,1),(1,2),(2,1),(1,0)).$$

The sets A are given in the following table.

example	A
(1.4.1)	$\{1, \ldots, 5\} \times \{1, \ldots, 5\}$
(1.4.2)	$\{1, \ldots, 5\} \times \{1, \ldots, 5\} \setminus \{(3,3)\}$
(1.4.3)	$\mathbb{Z}^2$
(1.4.4)	$\mathbb{Z}^2 \setminus 3\mathbb{Z}^2$

The questions we looked at in section 1.3 concerning uniqueness, consistency and reconstruction can be related in a direct way to properties of the corresponding grid reconstruction system. This suggests that these properties may also be of interest for more general reconstruction systems.

Let $A \subset \mathbb{Z}^r$ and D a sequence of directions in $\mathbb{Z}^r$. Let p be the grid reconstruction system over $\mathbb{Z}$ associated with (A, D). The *uniqueness* problem

asks when two functions $f, g : A \to \mathbb{Z}$ have the same line sums, i.e., when $p(f) = p(g)$. The linearity of p implies that

$$p(f) = p(g) \Longleftrightarrow p(f - g) = 0.$$

So to answer the uniqueness problem, or study which functions have the same line sums, we have to look at the kernel

$$\ker(p) = \{f : A \to \mathbb{Z} \mid p(f) = 0\}.$$

For the *consistency* problem, an integer is given for each line in a direction from D that goes through A, and we are asked when there is a function $f : A \to \mathbb{Z}$ such that its line sums are precisely these numbers. Phrased in terms of the reconstruction system, we are given an element $c \in \bigoplus_{i=1}^{n} \mathbb{Z}[A_i]$ and the question is if there is an $f \in \mathbb{Z}[A]$ such that $p(f) = c$. In other words, the question is to identify the image

$$\mathrm{im}(p) = \{p(f) \mid f \in \mathbb{Z}[A]\}$$

of the projection map.

In this algebraic setup, the image of the projection map is not a convenient space to work with. Instead, we will look at the so-called cokernel

$$\mathrm{cok}(p) = \Big(\bigoplus_{i=1}^{n} \mathbb{Z}[A_i]\Big) \Big/ \mathrm{im}(p).$$

Note that an element of $\bigoplus_{i=1}^{n} \mathbb{Z}[A_i]$ is in the image of p if and only if it maps to 0 in $\mathrm{cok}(p)$, thus the cokernel serves just as well to answer questions about consistency.

As was remarked in section 1.3, there is a nice connection between the reconstruction of functions $A \to \mathbb{Z}$ from their line sums and the reconstruction of subsets of A from their line sums as discussed in section 1.2.

Definition 2.1.11. Let $p : T \to P$ be a reconstruction system over $\mathbb{Z}$ and suppose that $T = \mathbb{Z}[A]$ for some set A. Let B be a finite subset of A. Then the *subset function* corresponding to B is the element $t = (t_a)_{a \in A}$ of T given by $t_a = 1$ if $a \in B$ and $t_a = 0$ otherwise.

Lemma 2.1.12. *Let $p : T \to P$ be a reconstruction system over $\mathbb{Z}$ with $T = \mathbb{Z}[A]$. Let w be the function*

$$w : \mathbb{Z}[A] \longrightarrow \mathbb{Z} \quad (t_a)_{a \in A} \mapsto \sum_{a \in A} t_a$$

and that we have a function $\bar{w} : P \to \mathbb{Z}$ such that $w = \bar{w} \circ p$. Let $x \in P$. Let W be the function

$$W : \mathbb{Z}[A] \longrightarrow \mathbb{Z} \quad (t_a)_{a \in A} \mapsto \sum_{a \in A} t_a^2.$$

Then for any $t \in T$ with $p(t) = x$ we have $W(t) \geq \bar{w}(x)$ with equality if and only if t is a subset function.

Proof. As the t_a are integers, we have $t_a^2 \geq t_a$ for all $a \in A$. It follows that $W(t) \geq w(t)$ with equality if and only if $t_a^2 = t_a$ for all $a \in A$, i.e. if and only if $t_a \in \{0, 1\}$ for all $a \in A$. As $w(t) = \bar{w}(x)$, the result follows. □

Corollary 2.1.13. *Let $p : T \to P$ be a grid reconstruction system with $P \neq 0$ and let $x \in P$. If there are subset functions t satisfying $p(t) = x$, they are the elements of minimal Euclidean norm in $p^{-1}\{x\}$.*

In section 1.3 we also looked at *dependencies* between line sums. For general reconstruction systems, we give a definition below.

Definition 2.1.14. Let p be a reconstruction system over k. A *dependency* for p is a k-linear map $d : \mathrm{cok}(p) \to k$. The set of all such dependencies is the k-module $\mathrm{Hom}_k(\mathrm{cok}(p), k)$. We will write $\mathrm{Dep}(p)$ for this module.

At first sight, the relation between this definition of dependencies and the one described in section 1.3 is maybe somewhat non-obvious. The following example shows how one can connect the two. After this example there are two lemmas which describe this connection in full generality.

Example 2.1.15. Consider the reconstruction problem described in example (1.3.6). Just as the examples from section 1.4, this one can be viewed as a grid reconstruction system $p : T \to P$. The set A is $\{1, \ldots, n\} \times \{1, \ldots, m\}$ and the projection directions are $d_1 = (1, 0)$, $d_2 = (0, 1)$, $d_3 = (1, 1)$ and $d_4 = (1, -1)$. We have

$$T = \mathbb{Q}[A]$$

and

$$P = \mathbb{Q}[A_1] \oplus \cdots \oplus \mathbb{Q}[A_4],$$

where A_i is the image of A inside $\mathbb{Z}^2/d_i$.

For every i, we choose an isomorphism of $\mathbb{Z}^2/d_i$ with $\mathbb{Z}$ as in the table below. The table also shows the image A_i under the isomorphism

i	$(x,y) \mapsto$	A_i
1	y	$\{1,\ldots,m\}$
2	x	$\{1,\ldots,n\}$
3	$y-x$	$\{1-n\ldots,m-1\}$
4	$x+y$	$\{2,\ldots,m+n\}$

To distinguish them more clearly from each other and from any numerical constant that may come up, we will write $a_{i,j}$ for the element of $\mathbb{Q}[A_i]$ corresponding to $j \in A_i$. For a function $f : A \to Q$, that is, an element $f \in T$, we see that its image in $\mathbb{Q}[A_1]$ is the function that sends $i \in \{1,\ldots,n\}$ to r_i, where r_i is the i-th row sum as defined in example (1.3.6). For the images in the other $\mathbb{Q}[A_i]$ we have similar interpretations, as collected in the following table.

$$\begin{aligned}
\mathbb{Q}[A_1]: &\quad a_{1,1}r_1 + \cdots + a_{1,m}r_m \\
\mathbb{Q}[A_2]: &\quad a_{2,1}c_1 + \cdots + a_{2,n}c_n \\
\mathbb{Q}[A_3]: &\quad a_{3,1-n}s_{1-n} + \cdots + a_{3,m-1}s_{m-1} \\
\mathbb{Q}[A_4]: &\quad a_{4,2}t_2 + \cdots + a_{4,m+n}t_{m+n}
\end{aligned}$$

Let's now look at one of the dependencies from example (1.3.6). The relation

$$\sum_{j=1}^{m} jr_j + \sum_{i=1}^{n} ic_i = \sum_{k=2}^{m+n} kt_k$$

can also be written as

$$\sum_{j=1}^{m} jr_j + \sum_{i=1}^{n} ic_i - \sum_{k=2}^{m+n} kt_k = 0.$$

Inspired by the formula on the left-hand side of this equation, together with the relations described above, we define the following linear maps

$$\delta_i : \mathbb{Q}[A_i] \to \mathbb{Q},$$

defined by giving the image of each basis element $a_{i,j} \in A_i$.

i			
1	$a_{1,j}$	$\mapsto$	j
2	$a_{2,j}$	$\mapsto$	j
3	$a_{3,j}$	$\mapsto$	0
4	$a_{4,j}$	$\mapsto$	$-j$

Together these define a linear map

$$\begin{array}{rcl} \delta : \mathbb{Q}[A_1] \oplus \cdots \oplus \mathbb{Q}[A_4] & \longrightarrow & \mathbb{Q} \\ (x_1, \ldots, x_4) & \longmapsto & \delta_1(x_1) + \cdots + \delta_4(x_4). \end{array}$$

The dependency between the line sums that we have just mentioned means that for any $f : A \to \mathbb{Q}$, the element $p(f) \in P$ satisfies $\delta(p(f)) = 0$. Hence $\mathrm{im}(p)$ is sent to 0 by δ and thus δ induces a linear map on the quotient vector space $\mathrm{cok}(p) = P/\mathrm{im}(p)$. This linear map is a dependency for the reconstruction system p, as per definition (2.1.14).

Lemma 2.1.16. *Let $p : T \to P$ be a reconstruction system over k. Let $q : P \to \mathrm{cok}(p)$ be the quotient map. Then there is an isomorphism of k-modules*

$$\begin{array}{ccc} \mathrm{Dep}(p) & \longrightarrow & \{f \in \mathrm{Hom}_k(P, k) \mid f \circ p = 0\} \\ d & \longmapsto & d \circ q. \end{array}$$

Proof. Note that for any $d \in \mathrm{Dep}(p)$, $d \circ q$ is a k-linear map $P \to k$ and that

$$(d \circ q) \circ p = d \circ (q \circ p) = d \circ 0 = 0.$$

Thus the map from the lemma is well-defined. One sees readily that it is k-linear. As q is surjective, we have $d \circ q = d' \circ q$ if and only if $d = d'$, so the map is injective.

Suppose that $f : P \to k$ is k-linear and $f \circ p = 0$. For every $x \in \mathrm{im}(p)$ we then have $f(x) = 0$. Therefore, if we have $y \in \mathrm{cok}(p) = P/\mathrm{im}(p)$ and let $\tilde{y} \in P$ be any lift of y, the value $f(\tilde{y})$ does not depend on the chosen lift, only on y. In this way f induces a map $d : \mathrm{cok}(p) \to k$. One checks that d is k-linear and satisfies $f = d \circ q$. Thus the morphism from the lemma is surjective, and therefore is an isomorphism. □

Lemma 2.1.17. *Let $A \subset \mathbb{Z}^r$ and let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^r$. Let*

$$\mathcal{L} = A_1 \sqcup \cdots \sqcup A_n$$

be the set of all lines in directions d_1 up to d_n passing though the set A. Let p be the grid reconstruction system over k corresponding to (A, D). Then there

is a bijection between $\mathrm{Dep}(p)$ *and the set of sequences* $(c_\ell)_{\ell\in\mathcal{L}}$ *with* $c_\ell \in k$ *for all* $\ell \in \mathcal{L}$ *such that for every* $t \in k[A]$ *we have*

$$\sum_{\ell\in\mathcal{L}} c_\ell p_\ell(t) = 0.$$

Proof. From the previous lemma we know that there is a bijection between $\mathrm{Dep}(p)$ and the set

$$\{f \in \mathrm{Hom}_k(P, k) \mid f \circ p = 0\}.$$

Note that P is a free k-module with basis $\mathcal{L}$, as it is the direct sum of free k-modules with bases $A_1, \ldots, A_n$. A k-linear map $P \to k$ is uniquely determined by where it sends the elements of a basis of P. Moreover, any map $\mathcal{L} \to k$ can be extended linearly to a $P \to k$.

For $t \in k[A]$, the element $p(t)$ in P can be written as

$$\sum_{\ell\in\mathcal{L}} p_\ell(t)\ell,$$

that is to say, the coefficient for the line ℓ is the line sum corresponding to that line, as we have established before in examples (2.1.7) and (2.1.15). It follows that the k-linear map $f : P \to k$ corresponding to the sequence $(c_\ell)_{\ell\in\mathcal{L}}$ sends $p(t)$ to

$$\sum_{\ell\in\mathcal{L}} c_\ell p_\ell(t)$$

and we conclude such a map is a dependency if and only if all these are sent to 0. □

The lemma tells us that for grid reconstruction systems, dependencies in the sense of definition (2.1.14) correspond with those described in definition (1.3.5). Using the algebraic description of $\mathrm{Dep}(p)$ we can now be very precise about the relation between dependencies and the consistency problem. The lemma below describes a very general case in which the dependencies suffice to describe $\mathrm{cok}(p)$ completely. Corollary (2.1.19) is a special case that will come up several times in chapter 3.

Lemma 2.1.18. *Let* k *be commutative ring, let* M *be a* k*-module and suppose that there is a free* k*-module* F *such that* $M \subset F$*. For* $m \in M$ *one then has*

$$m = 0 \iff \forall d \in \mathrm{Hom}_k(M, k) : d(m) = 0.$$

If F has finite rank, then there is a finite set of maps $d_1, \ldots, d_n$ such that one has

$$m = 0 \Longleftrightarrow d_1(m) = 0, \ldots, d_n(m) = 0.$$

If M itself is free of finite rank, then $\mathrm{Hom}_k(M, k)$ is free of the same rank.

Proof. Let $(f_i)_{i \in I}$ be a basis of F. Then any $m \in M$ can be uniquely written as

$$m = \sum_{i \in I} m_i f_i$$

with all but finitely many of the $m_i = 0$. For $i \in I$ let c_i be the map sending m to the coefficient m_i in the expansion above. The map $c_i : F \to k$ is k-linear. Denote by d_i the restriction of c_i to M. Then the d_i are in $\mathrm{Hom}_k(M, k)$. Thus if we have $d(m) = 0$ for all $f \in \mathrm{Hom}_k(M, k)$, we have in particular that $m_i = d_i(m) = 0$ for all $i \in I$, and therefore $m = \sum_{i \in I} m_i f_i = 0$. The other implication is trivial.

Suppose F has finite rank. Let $f_1, \ldots, f_n$ be a basis of F and let $c_1, \ldots, c_n$ be the corresponding coefficient maps $F \to k$ as defined in the previous paragraph. Let g be any homomorphism $F \to k$ and let $g_i = g(f_i)$ for $i = 1, \ldots, n$. Then we have

$$g = g_1 c_1 + \cdots + g_n c_n,$$

as these are linear maps $F \to k$ which take the same values on all d_i. Thus the c_i generate $\mathrm{Hom}_k(F, k)$. Moreover, as the coefficients f_i can be recovered from f, they are unique. Hence the c_i are a basis of $\mathrm{Hom}_k(F, k)$. Let $d_1, \ldots, d_n$ be the restrictions of the c_i to M. As remarked before, we have $m = 0$ if and only if $d_i(m) = 0$ for $i = 1, \ldots, n$.

Lastly, if M is free of finite rank, then we can take $F = M$ and hence by the previous paragraph, $\mathrm{Hom}_k(M, k)$ is free of the same rank, as the c_i form a basis. □

Corollary 2.1.19. (Consistency conditions) *Let $p : T \to P$ be a reconstruction system such that $\mathrm{cok}(p)$ is a free k-module of finite rank. Then $\mathrm{Dep}(p)$ is also free, of the same rank. Moreover, for $x \in P$ we have*

$$x \in \mathrm{im}(p) \iff \forall d \in \mathrm{Dep}(p) : d(x) = 0.$$

That is, a vector of projections comes from a table if and only if it satisfies all dependencies.

Example 2.1.20. Looking at the examples from section 1.4, there are three where the corollary does apply, and one (example (1.4.2)) where it does not. We will see in chapter 3 that for full grids (such as in example (1.4.3)) and finite convex grids (such as in example (1.4.1)), the corollary above applies, i.e. the cokernel is free of finite rank and therefore the dependencies are the only obstruction to the consistency problem. In chapter 5 we will show that the cokernel for the periodic grid from example (1.4.4) is the same as that for the full grid, so it is also free.

Example (1.4.2) describes an element v of P with the property that it is not in the image of p, but $2v$ is. In other words, v represents an element of order 2 in $\mathrm{cok}(p)$. We conclude that in this case $\mathrm{cok}(p)$ is not free, and so the corollary above does not apply.

2.2 Morphisms and functors

In this section, we will describe the maps that we want to consider between reconstruction systems. These will allow us to study reconstruction systems in relation to each other and to try to take what we know about one system and apply it to others. The first half of the section contains the required definitions, some very basic results and a few examples. In the second half of the section, we show that morphisms between reconstruction systems also give rise to maps between their kernels, cokernels and dependencies in a nice way.

A reconstruction system, as we've defined in the previous section, is a homomorphism between k-modules. These are very basic objects in modern algebra, about which a great deal has been written. Very many elementary results hold for these objects, far more than can reasonably be collected into this text. The results presented here are those that will be of interest to us in the upcoming chapters. Proofs are included mostly for the convenience of the reader.

As a general reference for all basic results about modern algebra, Lang's book [17] is indispensable. It has a chapter on modules and also a brief introduction on categories. Another brief introduction on categories can be found in [11, Ch. 2]. Northcott in [19, Ch. 4] discusses diagrams and morphisms between them, of which our reconstruction systems are a simple case.

The category of reconstruction systems

Definition 2.2.1. Let k be a commutative ring and let $p : T \to P$ and $p' : T' \to P'$ be reconstruction systems over k. A *morphism* of reconstruction systems is a pair $f = (f_T, f_P)$ of k-linear maps $f_T : T \to T'$ and $f_P : P \to P'$ such that $p' \circ f_T = f_P \circ p$ holds.

We write $f : p \to p'$ to say f is a morphism from p to p'.

We often find it convenient to explain the relationships between various modules and maps pictorially, through the use of *diagrams*. Formally, a diagram is a directed graph, whose vertices are labeled by modules and whose edges are labeled by k-linear maps. The map corresponding to an edge of course has as its domain and codomain the modules corresponding respectively to the starting and ending vertex of that edge.

For example, the spaces and maps in definition (2.2.1) above can be captured in a diagram as shown here.

$$\begin{array}{ccc} T & \xrightarrow{p} & P \\ {\scriptstyle f_T}\downarrow & & \downarrow{\scriptstyle f_P} \\ T' & \xrightarrow[p']{} & P' \end{array}$$

A path in a diagram represents a sequence of maps that can be composed to form another map, whose domain is the module corresponding to the starting vertex of the path, and codomain is the module corresponding to the ending vertex of the path. We say a diagram is *commutative* (or that the diagram *commutes*) if whenever there are multiple paths between two vertices of the graph, the corresponding maps are equal.

For example, the required condition in definition (2.2.1) is equal to saying that the diagram above is commutative. In this case the formula describing the relation is arguably simpler than the diagram, however when more spaces are involved, commutative diagrams very quickly become much easier to think about than the relations between maps they represent.

Example 2.2.2. Let $A \subset \mathbb{Z}^r$ and $B \subset A$. Let D be a sequence of directions in $\mathbb{Z}^r$. Let k be a commutative ring. Note that any line in direction d_i that has points in common with B, also has points in common with A. Hence there

are maps $\iota_i : B_i \to A_i$ for all i, in addition to the inclusion map $\iota : B \to A$. These maps give rise to maps $\iota : k[B] \to k[A]$ and $\iota_i : k[B_i] \to k[A_i]$. The map $k[B] \to k[A]$ takes a function $f : B \to k$ and extends it to a function $f : A \to k$ by putting $f(x) = 0$ for all $x \in A \setminus B$. Likewise, a function $g : B_i \to k$ is extended to $A_i \to k$ by putting $g(x) = 0$ for all $x \in A_i \setminus B_i$.

Let $x \in k[B]$ and consider the line sums of $\iota(x)$ in direction d_i. For any line that doesn't meet B, the line sum is 0, as it's only summing coefficients of points in $A \setminus B$, which are all 0. Likewise, for any line that does meet B, the line sum of $\iota(x)$ as a function on A is the same as the line sum of x as a function on B, as the coefficient at any additional point in the sum is 0.

The upshot of the discussion above is the following. Let p and p' be the grid reconstruction systems over k associated to (B, D) and (A, D) respectively. Let f_T be the map $\iota : k[B] \to k[A]$ and $f_P = (\iota_1, \ldots, \iota_n)$. Then (f_T, f_P) is a morphism of reconstruction systems $p \to p'$.

Example 2.2.3. We continue with the same setup as in the previous example: $A \subset \mathbb{Z}^r$, $B \subset A$ and D a sequence of directions in $\mathbb{Z}^r$. Note that any map $f : A \to k$ can be restricted to a map $f|_B : B \to k$, and likewise, any map $g : A_i \to k$ can be restricted to a map $g|_{B_i} : B_i \to k$.

Thus if we let p and p' be the grid reconstruction systems over k associated to (B, D) and (A, D) respectively, there are restriction maps $f_T : k[A] \to k[B]$ and $f_P : \bigoplus_{i=1}^n k[A_i] \to \bigoplus_{i=1}^n k[B_i]$.

However, the pair (f_T, f_P) is, in general, **not** a morphism of reconstruction systems. Consider for example

$$A = \{1, \ldots, 5\} \times \{1, \ldots, 5\}$$

and the directions as in example (1.4.1) and

$$B = A \setminus \{(3, 3)\}$$

as in example (1.4.2). Let p and p' be the corresponding grid reconstruction systems and f_T and f_P as above.

Let $t : A \to \mathbb{Z}$ be the map sending $(3, 3)$ to 1 and every other point of 0. Then $f_T(t)$ is the restriction of t to B, which is simply the zero map on B. Note that every line in a direction we consider that passes through A, also passes through B, and vice versa. Thus the map f_P is simply the identity. Note that t has line sum 1 in each of the lines passing through $(3, 3)$, thus

$p(t)$ is non-zero. On the other hand, as $f_T(t) = 0$, we have $p'(f_T(t)) = 0$. We conclude that $f_P(p(t))$ and $p'(f_T(t))$ are distinct, so the pair (f_T, f_P) is not a morphism of reconstruction systems $p \to p'$.

Lemma 2.2.4. *Let p, p' and p'' be reconstruction systems over k, and let $f : p \to p'$ and $g : p' \to p''$ be morphisms between them. Then*

$$g \circ f = (g_T \circ f_T, g_P \circ f_P)$$

is a morphism of reconstruction systems $p \to p''$.

Proof. Fixing notation, we have $p : T \to P$ and similarly for p' and p''. It is clear that $g_T \circ f_T$ is a k-linear map $T \to T''$ and $g_P \circ f_P$ is a k-linear map $P \to P''$. To show that $g \circ f$ is a morphism of reconstruction systems, it remains to be shown that $(g \circ f)_P \circ p = p'' \circ (g \circ f)_T$. Plugging in the definitions we see that

$$\begin{aligned}(g \circ f)_P &= (g_P \circ f_P) \circ p = g_P \circ (f_P \circ p) = g_P \circ (p' \circ f_T) \\ &= (g_P \circ p') \circ f_T = (p'' \circ g_T) \circ f_T = p'' \circ (g_T \circ f_T) \\ &= p'' \circ (g \circ f)_T\end{aligned}$$

holds as required. □

Corollary 2.2.5. *Let k be a commutative ring. Then the reconstructions systems over k together with their morphisms form a category.*

Proof. A *category* consists of a collection (the reader will forgive me for not going into the set-theoretical minutiae here) of *objects*, sets of *morphisms* between them and a *composition law* for morphisms, satisfying certain conditions.

In our case, the objects are reconstruction systems over k, the morphisms are the ones defined in (2.2.1) and the composition law is the one described in lemma (2.2.4).

The conditions come down to three things: that the 'right' morphisms be composable, that the composition is associative, and the existence of identity morphisms.

The first of these is dealt with in lemma (2.2.4), which defines composition precisely for the right morphisms. Composiition is possible whenever the codomain of the first map is the domain of the second.

The associativity of composition follows immediately from the fact it holds for the composition of k-linear maps.

This leaves the identity morphisms to be dealt with now. Let $p : T \to P$ be a reconstruction system over k. The pair $\mathrm{id}_p = (\mathrm{id}_T, \mathrm{id}_P)$ clearly defines a morphism of reconstruction systems $p \to p$. Moreover, it is clear that if $f : p' \to p$ and $g : p \to p''$ are morphisms of reconstruction systems, we have $\mathrm{id}_p \circ f = f$ and $g \circ \mathrm{id}_p = g$. These are the properties the identity morphisms in a category must satisfy. □

The fact that we have these morphisms and the notion of composition means that it also makes sense to make diagrams where the vertices represent reconstruction systems over k and the edges morphisms between them. Indeed one can sensibly talk about diagrams and commutative diagrams in any category.

Categories, like diagrams, are fairly pervasive in modern algebra. They reflect the philosophy that one should not just study objects, but also the relations between them, i.e. the morphisms. The language of categories makes it easy to say with few words and treat as single notions things that have fairly complicated definitions.

Definition 2.2.6. An *isomorphism* of reconstruction systems over k is a morphism $f : p \to p'$ such that there is a morphism $g : p' \to p$ satisfying $g \circ f = \mathrm{id}_p$ and $f \circ g = \mathrm{id}_{p'}$.

Note that this definition of isomorphism makes sense in any category. A basic fact about k-linear maps between modules is that the notion of isomorphism described here is equivalent to the k-linear map being bijective. Bijectivity doesn't make much sense in the category of reconstruction systems, where the objects are not 'sets with structure'.

Lemma 2.2.7. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then f is an isomorphism if and only if f_T and f_P are bijective.*

Proof. Fixing notation, let $p : T \to P$, $p' : T' \to P'$ and $f = (f_T, f_P)$. Suppose that f is an isomorphism. Then, by definition, there is a morphism $g : p' \to p$ such that $g \circ f = \mathrm{id}_p$ and $f \circ g = \mathrm{id}_{p'}$. Write $g = (g_T, g_P)$. Plugging in the definition of the composition law, we see that we have $g_T \circ f_T = \mathrm{id}_T$ and $f_T \circ g_T = \mathrm{id}_{T'}$, so f_T is an isomorphism of k-modules and is therefore bijective. Similarly, f_P is a isomorphism of k-modules $P \to P'$, so it too is bijective.

Conversely, suppose that f_T and f_P are bijective. Write g_T and g_P for their inverses. These are k-linear maps. To show that $g = (g_T, g_P)$ is a morphism of reconstruction systems, we have to prove that $p \circ g_T = g_P \circ p'$.

Note that

$$f_P \circ p \circ g_T = p' \circ f_T \circ g_T = p' = f_P \circ g_P \circ p'.$$

As f_P is injective, it follows that $p \circ g_T = g_P \circ p'$ holds as required. It is clear that f and g satisfy $g \circ f = \mathrm{id}_p$ and $f \circ g = \mathrm{id}_{p'}$. □

Some functorial constructions

We now direct our attention to proving some basic properties of the kernels and cokernels of reconstruction systems. As remarked before, these have a lot to do with questions of uniqueness and consistency.

Lemma 2.2.8. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then one has $f_P(\ker(p)) \subset \ker(p')$.*

Proof. This follows at once from the compatibility condition morphisms must satisfy. If $x \in \ker(p)$ we have

$$p'(f_P(x)) = f_T(p(x)) = f_T(0) = 0,$$

so $f_P(x) \in \ker(p')$ as required. □

Corollary 2.2.9. *There is a covariant functor*

$$\text{KER} : \textit{reconstruction systems over } k \longrightarrow k\textit{-modules}$$

that sends a reconstruction system p over k to the k-module $\ker(p)$.

Proof. If, in a somewhat recursive fashion, we think of categories as objects, *functors* would be the morphisms that connect them. Specifically, a functor F between categories C and D associates to every object X of C an object $F(X)$ of D and to every morphism $f : X \to Y$ of C a morphism $F(f)$ satisfying

$$\begin{array}{ll} F(f) : F(X) \to F(Y) & \text{if } F \text{ is } \textit{covariant} \\ F(f) : F(Y) \to F(X) & \text{if } F \text{ is } \textit{contravariant.} \end{array}$$

These should satisfy $F(\mathrm{id}_X) = \mathrm{id}_{F(X)}$ for all objects X of C and for morphisms $f : Y \to Z$ and $g : X \to Y$ in C:

$$\begin{array}{ll} F(f \circ g) = F(f) \circ F(g) & \text{if } F \text{ is } \textit{covariant} \\ F(f \circ g) = F(g) \circ F(f) & \text{if } F \text{ is } \textit{contravariant.} \end{array}$$

In the case of this lemma, the functor KER associates to the reconstruction system $p : T \to P$ the k-module $\ker(p)$. Lemma (2.2.8) shows that if f is

a morphism of reconstruction systems $p \to p'$, then $\textsc{ker}(f) = f_p|_{\ker(p)}$ is a k-linear map $\ker(p) \to \ker(p')$. A straightforward verification shows that this satisfies the properties mentioned above. □

Constructions that take an object of one type and associate to it another object of another type are very common. To say such a construction is a functor (or 'functorial') is to say that it behaves well with respect to maps between the objects.

Example 2.2.10. Consider A and D as in example (1.4.1) and let p be the corresponding grid reconstruction system. The inclusion $A \subset \mathbb{Z}^2$ gives rise to a morphism of reconstruction systems $f : p \to q$ as described in example (2.2.2), where q is the grid reconstruction system corresponding to $(\mathbb{Z}^2, D)$, as in example (1.4.3).

For $t \in \ker(p)$, the function $f_T(t)$ is the map $\mathbb{Z}^2 \to k$ that sends x to $t(x)$ if $x \in A$ and 0 if $x \notin A$. For any line not going through A, the line sum of $f_T(t)$ is therefore 0, and for any line that does pass through A, it is equal to the corresponding line sum of t, which is also 0. Thus $f_T(t)$ is indeed in the kernel of q and we have a corresponding map $\ker(p) \to \ker(q)$.

This map is clearly injective and so we can identify $\ker(p)$ with a subspace of $\ker(q)$. Looking at the definition it is clear which subspace it is:

$$\ker(p) = \{t \in \ker(q) \mid \forall x \notin A : t(x) = 0\},$$

i.e. it is the space of those functions in $\ker(q)$ whose *support* is contained in A. In chapter 3, we will use algebraic methods to derive a description of $\ker(q)$. We can then use the relation between the kernels that we have just shown to arrive at a description for $\ker(p)$.

Lemma 2.2.11. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then one has $f_P(\mathrm{im}(p)) \subset \mathrm{im}(p')$.*

Proof. Let $y \in \mathrm{im}(p)$, then there is some $x \in T$ such that $y = p(x)$ and therefore we have

$$f_P(y) = f_P(p(x)) = p'(f_T(x))$$

and so $f_P(y)$ is in $\mathrm{im}(p')$. □

Corollary 2.2.12. *There are covariant functors* IM *and* COK *from reconstruction systems over k to k-modules sending a reconstruction system p to* $\mathrm{im}(p)$ *and* $\mathrm{cok}(p)$ *respectively.*

Proof. Let $f : p \to p'$ be a morphism of reconstruction systems over k. By lemma (2.2.11), $\textsc{im}(f) = f_P|_{\mathrm{im}(p)}$ is a k-linear map $\mathrm{im}(p) \to \mathrm{im}(p')$. A straightforward verification shows that IM satisfies the conditions required of a functor.

Let $y \in \mathrm{cok}(p)$ and let x and x' be lifts of y to P. Then $x - x'$ is in $\mathrm{im}(p)$, and so $f_P(x - x') \in \mathrm{im}(p')$ by lemma (2.2.11). It follows that the images in $\mathrm{cok}(p')$ of $f_P(x)$ and $f_P(x')$ are the same. Hence there is a well defined map $\textsc{cok}(f) : \mathrm{cok}(p) \to \mathrm{cok}(p')$ sending $y \in \mathrm{cok}(p)$ to the class of $f_P(x)$ for some lift x of y to P. Straightforward computations show that this map is k-linear and that COK satisfies the conditions required for being a functor. □

Lemma 2.2.13. *There is contravariant functor* DEP *from reconstruction systems over k to k-modules, sending a reconstruction system p to*

$$\mathrm{Dep}(p) = \mathrm{Hom}_k(\mathrm{cok}(p), k).$$

Proof. Let $f : p \to p'$ be a morphism of reconstruction systems. By corollary (2.2.12) there is a k-linear map $\textsc{cok}(f) : \mathrm{cok}(p) \to \mathrm{cok}(p')$. Write ϕ for this map. When $d : \mathrm{cok}(p') \to k$ is an element of $\mathrm{Dep}(p')$, the composition $d \circ \phi$ is a k-linear map $\mathrm{cok}(p) \to k$, i.e. an element of $\mathrm{Dep}(p)$. A straightforward computation shows that ϕ in this way induces a k-linear map $\textsc{Dep}(f) : \mathrm{Dep}(p') \to \mathrm{Dep}(p)$. More simple computations show that these maps satisfy the conditions for a contravariant functor as given in the proof of (2.2.9). □

Example 2.2.14. Continuing from example (2.2.10), the morphism of reconstruction systems $f : p \to q$ gives rise to a morphism

$$g = \textsc{Dep}(f) : \mathrm{Dep}(q) \longrightarrow \mathrm{Dep}(p).$$

We can think of this map as follows. Any element

$$c \in k[A_1] \oplus \cdots \oplus k[A_4],$$

can be extended by zeros to the element

$$f_P(c) \in k[\mathbb{Z}^2/d_1] \oplus \cdots \oplus k[\mathbb{Z}^2/d_4].$$

By lemma (2.1.16), any dependency for q gives rise to a k-linear map

$$d : \bigoplus_{i=1}^{4} k[\mathbb{Z}^2/d_i] \longrightarrow k$$

such that $d \circ q = 0$. Thus sending c to $d(f_P(c))$ is a k-linear map

$$d' : \bigoplus_{i=1}^{4} k[A_i] \longrightarrow k$$

which satisfies $d' \circ p = 0$, as by lemma (2.2.11), $f_P(c)$ is in $\mathrm{im}(q)$ whenever c is in $\mathrm{im}(p)$. By lemma (2.1.16) this therefore is a dependency for p.

Looking at the coefficients for each line as in (2.1.17), we see that the coefficients for d' are simply the coefficients for d that correspond to lines going through A. We will study the dependencies that can be obtained in this way extensively in chapter 3.

2.3 Categorical constructions

There are very strong similarities between the category of reconstruction systems over k and the category of k-modules. In this section, we will take concepts that are well-known in the context of modules, such as direct sums, and kernels and cokernels of morphisms, and construct analogues of these in the category of reconstruction systems. The entire analogy can be expressed succinctly by saying the category of reconstruction systems is an Abelian category, see theorem (2.3.25).

Abelian categories are often studied in the context of homological algebra. There, one begins with objects from some Abelian category and studies *complexes*

$$\cdots \to K^{n-1} \underset{d^{n-1}}{\to} K^n \underset{d^n}{\to} K^{n+1} \to \cdots$$

where $d^n \circ d^{n-1} = 0$ for all $n \in \mathbb{Z}$. Associated to such a complex are *homology groups* $H^n = \ker(d^n)/\mathrm{im}(d^{n-1})$. We can view our reconstruction systems as very simple complexes

$$\cdots \to 0 \to 0 \to T \underset{p}{\to} P \to 0 \to 0 \to \cdots.$$

The kernel and cokernel of p then arise as the only non-trivial homology groups. Books on homological algebra rarely mention that the category of

complexes is again Abelian, as it often suffices simply to define what an exact sequence of complexes is (or trust their readers to figure this out by themselves). Some references for homological algebra include [5], [11], [14] and [19], as well as chapter 20 from [17] and appendix 3 from [8].

Lemma 2.3.1. *Let p, p' be reconstruction systems over k. Then the set*

$$\mathrm{Hom}(p, p') = \{f : p \to p' \mid f \text{ is a morphism of reconstruction systems}\}$$

has a natural k-module structure defined by

- $f + g = (f_T + g_T, f_P + g_P)$
  for any morphisms $f = (f_T, f_P)$ and $g = (g_T, g_P)$ in $\mathrm{Hom}(p, p')$*;*
- $\lambda f = (\lambda f_T, \lambda f_P)$
  for any morphism $f = (f_T, f_P)$ in $\mathrm{Hom}(p, p')$ *and any $\lambda \in k$.*

Proof. For k-modules, M and M', the set $\mathrm{Hom}(M, M')$ is a k-module with point-wise addition and scalar multiplication. Moreover, the composition law

$$\mathrm{Hom}(M, M') \times \mathrm{Hom}(M', M'') \longrightarrow \mathrm{Hom}(M, M'')$$

is k-bilinear. From this one sees immediately that $f + g$ and λf defined above are indeed morphisms of reconstruction systems. That they satisfy the required axioms to be a k-module follows from the fact that we can check this for the T- and P-maps separately, and we already know that for k-modules they do give a module structure on the Hom-set. □

Lemma 2.3.2. *Let p, p' and p'' be reconstruction systems over k. Then the composition of morphisms defined in lemma* (2.2.4) *is a k-bilinear map*

$$\mathrm{Hom}(p, p') \times \mathrm{Hom}(p', p'') \longrightarrow \mathrm{Hom}(p, p'').$$

Proof. As we saw in the previous lemma, the composition of k-linear maps is bilinear. This carries over to the case of reconstruction systems as we can verify the required axioms on the T- and P-maps separately. □

Remark 2.3.3. Let Z be the trivial k-module, $Z = \{0\}$. The unique k-linear map $0 : Z \to Z$ is a reconstruction system over k. We call it the zero reconstruction system and write 0 for it. For every k-module M, there is a unique k-linear map $Z \to M$. This implies that for every reconstruction

system p, there is a unique morphism of reconstruction systems $0 \to p$ and a unique morphism $p \to 0$. If p and p' are two reconstruction systems, the zero element of $\mathrm{Hom}(p, p')$ is the composition of $p \to 0$ and $0 \to p'$.

The reconstruction system 0 is a so-called *initial object* in the category of reconstruction systems over k. In a category C, an object I is called initial, if for every object X in C, there is a unique morphism $I \to X$ in C. Other examples of initial objects include the trivial group in the category of groups and $\mathbb{Z}$ in the category of rings.

In addition to being an initial object, the reconstruction system 0 is also a *final object* in the category of reconstruction systems over k. In a category C, an object F is called final, if for every object X in C there is a unique morphism $X \to F$ in C. Another example of a final object is the trivial group in the category of groups, or the zero k-module in the category of k-modules.

Not all categories have initial and final objects. Though the previous paragraphs always speak of *an* initial or final object, they are what is called 'uniquely unique' or 'unique up to unique isomorphism'. We state this result in its categorical generality here, because we shall use it several times later on to show that other objects are likewise uniquely unique.

Lemma 2.3.4. *Let C be a category and suppose that U and U' both are initial or both are final objects of C. Then the unique morphism $U \to U'$ is an isomorphism.*

Proof. As U is an initial object (or U' is a final object), there is a unique morphism $f : U \to U'$. Likewise, as U' is an initial object (or U a final object), there is a unique morphism $g : U' \to U$. Consider the composition $g \circ f$. This is a morphism $U \to U$. The identity morphism id_U is also a morphism, $U \to U$. As U is an initial (or final) object, there is a unique morphism $U \to U$, so we must have $g \circ f = \mathrm{id}_U$. In the same way, we must have $f \circ g = \mathrm{id}_{U'}$ as U' is an initial (or final) object. Thus the unique morphism $f : U \to U'$ is an isomorphism. □

Sums and products

Definition 2.3.5. Let $p_1 : T_1 \to P_1$ and $p_2 : T_2 \to P_2$ be reconstruction systems over k. The *direct sum* of p_1 and p_2 is the reconstruction system

$$\begin{array}{rcl} p_1 \oplus p_2 : T_1 \oplus T_2 & \longrightarrow & P_1 \oplus P_2 \\ (x_1, x_2) & \longmapsto & (p_1(x_1), p_2(x_2)). \end{array}$$

Remark 2.3.6. Let $p_1 : T_1 \to P_1$ and $p_2 : T_2 \to P_2$ be reconstruction systems over k and let $p_1 \oplus p_2$ be their direct sum. Let $\iota_{1,T}$ be the map

$$\begin{array}{ccc} T_1 & \longrightarrow & T_1 \oplus T_2 \\ x & \longmapsto & (x, 0) \end{array}$$

and define $\iota_{1,P}$ and $\iota_{2,T}$, $\iota_{2,P}$ similarly. These maps give rise to morphisms of reconstruction systems

$$\iota_i : p_i \longrightarrow p_1 \oplus p_2 \qquad i \in \{1, 2\}.$$

Let $\pi_{1,T}$ be the map

$$\begin{array}{ccc} T_1 \oplus T_2 & \longrightarrow & T_1 \\ (x, y) & \longmapsto & x \end{array}$$

and define $\pi_{1,P}$ and $\pi_{2,T}$, $\pi_{2,P}$ similarly. These maps give rise to morphisms of reconstruction systems

$$\pi_i : p_1 \oplus p_2 \longrightarrow p_i \qquad i \in \{1, 2\}.$$

For $i \in \{1, 2\}$, the morphism $\pi_i \circ \iota_i$ is the identity on p_i.

Lemma 2.3.7. *Let p_1 and p_2 be reconstruction systems over k. Let q be another reconstruction system over k and let $f_1 : q \to p_1$ and $f_2 : q \to p_2$ be morphisms. Then there is a unique morphism of reconstruction systems $F : q \to p_1 \oplus p_2$ such that $f_i = \pi_i \circ F$ for $i \in \{1, 2\}$.*

Proof. Fixing notation, we write $p_i : T_i \to P_i$ for $i \in \{1, 2\}$ and $q : T \to P$. Let $F_T : T \to T_1 \oplus T_2$ be the k-linear map sending $x \in T$ to $(f_{1,T}(x), f_{2,T}(x))$ and let $F_P : P \to P_1 \oplus P_2$ the map sending $x \in P$ to $(f_{1,P}(x), f_{2,P}(x))$. Note that for $x \in T$, we have

$$\begin{array}{rcl} [(p_1 \oplus p_2) \circ F_T](x) & = & [p_1 \oplus p_2](f_{1,T}(x), f_{2,T}(x)) \\ & = & ([p_1 \circ f_{1,T}](x), [p_2 \circ f_{2,T}](x)) \\ & = & ([f_{1,P} \circ q](x), [f_{2,P} \circ q](x)) \\ & = & F_P(q(x)), \end{array}$$

so $F = (F_T, F_P)$ is a morphism of reconstruction systems $q \to p_1 \oplus p_2$.

For $x \in T$ and $i \in \{1, 2\}$ we have

$$(\pi_i \circ F)_T(x) = \pi_{i,T}(f_{1,T}(x), f_{2,T}(x)) = f_{i,T}(x),$$

so $(\pi_i \circ F)_T = f_{i,T}$. Similarly, one shows $(\pi_i \circ F)_P = f_{i,P}$, thus we have $f_i = \pi_i \circ F$ as required.

Now suppose that $G : q \to p_1 \oplus p_2$ is a morphism of reconstruction systems that satisfies $f_i = \pi_i \circ G$ for $i \in \{1, 2\}$. For $x \in T$ and $i \in \{1, 2\}$ we then have

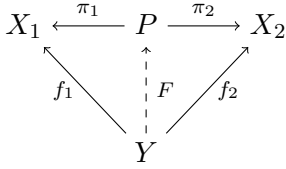
$$\pi_{i,T}(G_T(x)) = f_{i,T}(x),$$

that is, $G_T(x) = (f_{1,T}(x), f_{2,T}(x)) = F_T(x)$. Similarly, one shows that G_P and F_P are equal. We conclude that F is unique. □

The lemma shows that $p_1 \oplus p_2$ is a *product* of p_1 and p_2 in the category of reconstruction systems over k. If C is a category and X_1 and X_2 are two objects in C, then a product of X_1 and X_2 is an object P of C together with maps $\pi_1 : P \to X_1$ and $\pi_2 : P \to X_2$.

$$X_1 \xleftarrow{\pi_1} P \xrightarrow{\pi_2} X_2$$

They satisfy the *universal property* described in the lemma: whenever we have another object Y in C together with morphisms $f_1 : Y \to X_1$ and $f_2 : Y \to X_2$, there is a unique morphism $F : Y \to P$ such that $f_1 = \pi_1 \circ F$ and $f_2 = \pi_2 \circ F$.



Not in all categories do products always exist. In the category of reconstruction systems over k they do, this is what we have just shown. In the category of k-modules, they also exist, one verifies easily that the direct sum of two modules, together with the projections to the summands, is a product in that category.

Products, if they exist, are uniquely unique, in the same way initial or final objects are. One derives this formally from lemma (2.3.4) by noting that the triple (P, π_1, π_2) is a final object in the category whose objects are triples (Y, f_1, f_2), where Y is an object in C and $f_i : Y \to X_i$ for $i \in \{1, 2\}$. What the morphisms in this category are, is left as an exercise to the reader.

Example 2.3.8. Let $A = \{1, \ldots, n\} \times \{1, \ldots, m\}$ be a rectangular set in $\mathbb{Z}^2$. Let p_1 be the reconstruction system

$$\begin{array}{ccc} \mathbb{Z}[A] & \longrightarrow & \mathbb{Z}^m \\ f & \longmapsto & (r_1, \ldots, r_m), \end{array}$$

sending a function $f : A \to \mathbb{Z}$ to its row sums and p_2 the system

$$\begin{array}{ccc} \mathbb{Z}[A] & \longrightarrow & \mathbb{Z}^n \\ f & \longmapsto & (c_1, \ldots, c_n) \end{array}$$

sending a function $f : A \to \mathbb{Z}$ to its column sums. Let $p : \mathbb{Z}[A] \to \mathbb{Z}^m \oplus \mathbb{Z}^n$ be the grid reconstruction system corresponding to A and $D = ((0, 1), (1, 0))$.

The direct sum $p_1 \oplus p_2$ is the reconstruction system

$$\begin{array}{ccc} \mathbb{Z}[A] \oplus \mathbb{Z}[A] & \longrightarrow & \mathbb{Z}^m \oplus \mathbb{Z}^n \\ (f, g) & \longmapsto & (p_1(f), p_2(g)). \end{array}$$

It sends two functions $A \to \mathbb{Z}$ to the row sums of the first and the column sums of the second. This map is clearly not the same as the grid reconstruction system p, but there is a morphism between them, which we will construct now.

Let d be the diagonal map

$$\begin{array}{ccc} \mathbb{Z}[A] & \longrightarrow & \mathbb{Z}[A] \oplus \mathbb{Z}[A] \\ f & \longmapsto & (f, f). \end{array}$$

One sees that d and the identity map on $\mathbb{Z}^m \oplus \mathbb{Z}^n$ form a morphism of reconstruction systems $p \to p_1 \oplus p_2$. We will see later in lemma (2.4.14) how to relate the kernel and cokernel of p to those of $p_1 \oplus p_2$. This gives us a way to 'build up' more complicated reconstruction systems out of simpler ones.

We can understand the morphism $p \to p_1 \oplus p_2$ we have just constructed in terms of the universal property from lemma (2.3.7) as well. There are obvious maps $p \to p_1$ and $p \to p_2$ that one obtains by simply forgetting the sums one is not interested in. Corresponding to these two morphisms, there should be a unique morphism $p \to p_1 \oplus p_2$. One checks it is the morphism that we have just described.

Lemma 2.3.9. *Let p_1 and p_2 be reconstruction systems over k. Let q be another reconstruction system over k and let $f_1 : p_1 \to q$ and $f_2 : p_2 \to q$ be morphisms. Then there is a unique morphism of reconstruction systems $F : p_1 \oplus p_2 \to q$ such that $f_i = F \circ \iota_i$ for $i \in \{1, 2\}$.*

Proof. Fixing notation, we write $p_i : T_i \to P_i$ for $i \in \{1, 2\}$ and $q : T \to P$.

Let $F_T : T_1 \oplus T_2 \to T$ be the k-linear map sending $(x, y) \in T_1 \oplus T_2$ to $f_{1,T}(x) + f_{2,T}(y)$ and let $F_P : P_1 \oplus P_2 \to P$ the map sending $(x, y) \in P_1 \oplus P_2$ to $f_{1,P}(x) + f_{2,P}(y)$.

Note that for $(x, y) \in T_1 \oplus T_2$ we have

$$\begin{aligned} [q \circ F_T](x, y) &= q(f_{1,T}(x) + f_{2,T}(y)) \\ &= [q \circ f_{1,T}](x) + [q \circ f_{2,T}](y) \\ &= [f_{1,P} \circ p_1](x) + [f_{2,P} \circ p_2](y) \\ &= F_P(p_1(x), p_2(y)) \\ &= [F_P \circ (p_1 \oplus p_2)](x, y), \end{aligned}$$

so $F = (F_T, F_P)$ is indeed a morphism of reconstruction systems $p_1 \oplus p_2 \to q$.

For $x \in T_1$ we have

$$F_T(\iota_{1,T}(x)) = F_T((x, 0)) = f_{1,T}(x)$$

and similarly $F_T \circ \iota_{2,T} = f_{2,T}$. In much the same way one also shows

$$F_P \circ \iota_{i,P} = f_{i,P} \qquad \text{for } i \in \{1, 2\}.$$

It follows that $F \circ \iota_i = f_i$ for $i \in \{1, 2\}$ as required.

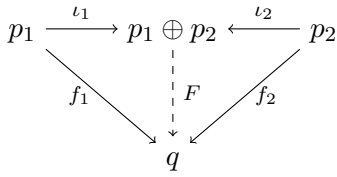
Conversely, suppose that G is a morphism $p_1 \oplus p_2 \to q$. Let $(x, y) \in T_1 \oplus T_2$. Then we have

$$\begin{aligned} G_T(x, y) &= G_T(\iota_1(x) + \iota_2(y)) \\ &= [G_T \circ \iota_1](x) + [G_T \circ \iota_2](y) \\ &= f_{1,T}(x) + f_{2,T}(y) \\ &= F_T(x, y) \end{aligned}$$

and similarly $G_P = F_P$. We conclude that F is unique as required. □

This lemma shows that the direct sum $p_1 \oplus p_2$ is a *co-product* in the category of reconstruction systems over k. The notion of co-product is dual to that of product. In the diagram for a co-product all the arrows point in the opposite

direction than they do in the diagram for a product.



Like products, co-products do not always exist in every category. When they do, they are essentially unique. This can be shown in much the same way as for products, the universal property from the lemma translates to the co-product being an initial object in a suitable category.

Even if products and co-products exist in a category, there is in general no reason to expect they are the same, unless the categories have additional structure. In the category of k-modules, one checks that the direct sum of two modules also is both product and co-product. In the category of groups, the product is given by the usual direct product of groups, but this is not the co-product. The co-product does always exist, but showing this is quite complicated.

Given any finite collection of reconstruction systems $p_1, \ldots, p_n$ we can construct the direct sum $p_1 \oplus \cdots \oplus p_n$ by repeatedly summing pairs, or give a direct definition similar to (2.3.5). One shows it satisfies universal properties similar to those in lemmas (2.3.7) and (2.3.9).

Given an arbitrary index set I and reconstruction systems p_i for $i \in I$, we can still define a product and co-product, but when I is infinite these will no longer be the same. The co-product

$$\bigoplus_{i \in I} T_i \longrightarrow \bigoplus_{i \in I} P_i$$

is referred to as the *direct sum* of the p_i's. One application of such potentially infinite direct sums is the following lemma about the decomposition of certain grid reconstruction systems.

Lemma 2.3.10. *Let $r \geq 2$, let D an n-tuple of directions in $\mathbb{Z}^r$ and let $A \subset \mathbb{Z}^r$. Let $p : T \to P$ be the grid reconstruction system corresponding to (A, D). Let $L \subset \mathbb{Z}^r$ be a subgroup and suppose that $d_1, \ldots, d_n \in L$. Let s be the rank of L. Then there is an n-tuple D' of directions in $\mathbb{Z}^s$ and there are*

subsets $A_x \subset \mathbb{Z}^s$ for all $x \in \mathbb{Z}^r/L$ such that

$$p \cong \bigoplus_{x \in \mathbb{Z}^r/L} p_x,$$

where p_x is the grid reconstruction system corresponding to (A_x, D').

Proof. As $L \subset \mathbb{Z}^r$ is a subgroup of a free abelian group of finite rank, it is itself a free abelian group of finite rank s. Fix an isomorphism $\phi : L \to \mathbb{Z}^s$. For $i = 1, \ldots, n$ we have $d_i \in L$ and so we can define $d_i' = \phi(d_i)$. Note that any line in direction d_i that is contained in L is transformed by ϕ into a line in direction d_i' in $\mathbb{Z}^s$.

For $x \in \mathbb{Z}^r/L$, let B_x be the subset of A given by

$$B_x = \{a \in A \mid a \equiv x \bmod L\}.$$

It is clear that the B_x are pairwise disjoint and that we have $A = \bigcup_{x \in \mathbb{Z}^r/L} B_x$. Furthermore, note that for every i, any line in direction d_i that has some points in common with B_x does not have any points in common with $A \setminus B_x$. Let q_x be the grid reconstruction system corresponding to (B_x, D). From the above discussion it follows that p decomposes as a direct sum

$$p \cong \bigoplus_{x \in \mathbb{Z}^r/L} q_x.$$

Let $\tilde{x} \in \mathbb{Z}^r$ be a lift of $x \in \mathbb{Z}^r/L$ and let $t : \mathbb{Z}^r \to \mathbb{Z}^r$ be the translation $y \mapsto y - \tilde{x}$. Let

$$A_x = \phi(t(B_x))$$

and let p_x be the reconstruction system corresponding to (A_x, D'). Then $\phi \circ t$ induces an isomorphism $q_x \cong p_x$. This completes the proof. □

Example 2.3.11. As an example of the previous lemma, consider the rectangular set

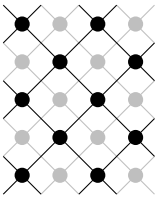
$$A = \{1, \ldots, n\} \times \{1, \ldots, m\}$$

and directions $d_1 = (1, 1)$ and $d_2 = (1, -1)$. Note that d_1 and d_2 span the subgroup

$$L = \{(x, y) \in \mathbb{Z}^2 \mid x + y \text{ is even}\}$$

of index 2 inside $\mathbb{Z}^2$.

As in the proof of the lemma above, A is split into 2 pieces, corresponding to the 2 residue classes in $\mathbb{Z}^2/L$. In the picture below, the dark points correspond to the trivial class (i.e. the elements of L itself), which we shall call B_0, and the light points to the other class, which we shall call B_1. The lines in directions d_1 and d_2 are also drawn into the picture and we see they likewise split up into two classes, each going only through points of one colour.



We pick the isomorphism $\phi : L \to \mathbb{Z}^2$ given by

$$\begin{array}{rccc} \phi: & L & \longrightarrow & \mathbb{Z}^2 \\ & (x,y) & \longmapsto & \left(\frac{x+y}{2}, \frac{x-y}{2}\right). \end{array}$$

Following the proof, we then find $d_1' = (1,0)$ and $d_2' = (0,1)$. Up to translation, the sets A_0 and A_1 are depicted in the following picture.

Kernels and cokernels

In our investigation of the similarities between reconstruction systems and modules, we now turn our attention to certain objects associated with morphisms in the category of k-modules: kernels and cokernels.

Lemma 2.3.12. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then one has $p(\ker(f_P)) \subset \ker(f_T)$.*

Proof. Let $x \in \ker(f_P)$. Then we have

$$f_T(p(x)) = p'(f_P(x)) = p'(0) = 0,$$

so $p(x) \in \ker(f_T)$ as required. □

Definition 2.3.13. Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then a *kernel* of f is a reconstruction system $\ker(f)$ together with a morphism $\iota : \ker(f) \to p$ such that $f \circ \iota = 0$ and such that for every reconstruction system q and morphism $g : q \to p$ satisfying $f \circ g = 0$, there is a unique morphism $h : q \to \ker(f)$ such that $g = \iota \circ h$.

For the direct sum, we gave a construction of the object and then showed it satisfies certain universal properties. For kernels and cokernels we take another approach, they are defined as object satisfying a certain universal property, and we then have a lemma that shows they exist and are unique up to unique isomorphism.

Lemma 2.3.14. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then a kernel of this morphism exists and is given by*

$$\ker(f) = p|_{\ker(f_P)} : \ker(f_P) \longrightarrow \ker(f_T).$$

Moreover, if (q_1, ι_1) and (q_2, ι_2) are two kernels of f, the unique morphism $\phi : q_1 \to q_2$ such that $\iota_1 = \iota_2 \circ \phi$ is an isomorphism.

Proof. By lemma (2.3.12), $p|_{\ker(f_P)}$ is a k-linear map $\ker(f_P) \to \ker(f_T)$ as claimed. The inclusion maps $\iota_T : \ker(f_T) \to T$ and $\iota_P : \ker(f_P) \to P$ give rise to a morphism of reconstruction systems $\iota : \ker(f) \to p$.

It is clear that $f \circ \iota = 0$. We will now show that the pair $(\ker(f), \iota)$ satisfies the universal property in definition (2.3.13). Let $q : T_q \to P_q$ be another reconstruction system over p together with a morphism $g : q \to p$ satisfying $f \circ g = 0$. Then $f_T \circ g_T = 0$, so the image of g_T is contained in $\ker(f_T)$. It follows that g_T can be written as the composition of a map $h_T : T_q \to \ker(f_T)$ and the inclusion $\iota_T : \ker(f_T) \to T$.

Similarly, g_P can be written as a composition $h_P : P_q \to \ker(f_P)$ followed by $\iota_P : \ker(f_P) \to P$. Since g is a morphism of reconstruction systems, $h = (h_T, h_P)$ is also one. It satisfies $g = h \circ \iota$ by construction Moreover, if h' is another morphism satisfying $g = h' \circ \iota$, then we have $h_T = h'_T$ as ι_T is injective, and similarly for the P-maps. We conclude that the pair $(\ker(f), \iota)$ satisfies the universal property of the kernel.

The universal property shows that the pair $(\ker(f), \iota)$ is a final object in a suitable category. The objects of this category are pairs (q, g) where q is a reconstruction system over k and $g : q \to p$ a morphism such that $f \circ g = 0$. A morphism $(q_1, g_1) \to (q_2, g_2)$ in this category is a morphism of reconstruction systems $\phi : q_1 \to q_2$ such that $g_1 = g_2 \circ \phi$. Applying lemma (2.3.4) to the final object of this category yields the uniqueness of the kernel as required. $\square$

Example 2.3.15. Consider the sets A and B as in examples (1.4.1) and (1.4.2), i.e.

$$A = \{1, \ldots, 5\} \times \{1, \ldots, 5\}$$

and $B = A \setminus \{(3,3)\}$ and let D be the sequence of directions from these examples. As we saw before in example (2.2.2), the inclusion $B \subset A$ gives rise to a morphism of reconstruction systems $f : p \to q$, where p and q are the grid reconstruction systems associated to (B, D) and (A, D) respectively.

The maps f_T and f_P are both related to extending a function by 0 to a larger domain. As such, it is clear that they are injective maps. From the computation in the previous lemma, it follows that $\ker(f)$ is the zero reconstruction system.

Example 2.3.16. An interesting example of a non-trivial kernel comes up in the following situation. Consider a rectangular grid

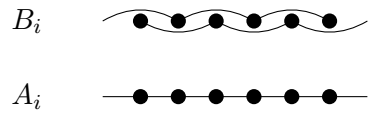
$$A = \{1, \ldots, n\} \times \{1, \ldots, m\}$$

and the two sequences of directions

$$\begin{aligned} D_1 = &\quad ((0,1),(1,0)) \\ D_2 = &\quad ((0,2),(2,0)). \end{aligned}$$

For convenience of notation, let $B = A$ and write $A_{1,2}$ for the sets of lines through A in the directions from D_1 and $B_{1,2}$ for the sets of lines through B in the directions from D_2.

For $i \in \{1, 2\}$ there is a 2-to-1 map from B_i to A_i, expressing that a line in A_i is split up in two in B_i, as shown in the following picture.



These maps induce k-linear maps

$$f_i : k[B_i] \longrightarrow k[A_i],$$

which fit into the commutative diagram below.

$$\begin{array}{ccc} k[B] & \longrightarrow & k[B_1] \oplus k[B_2] \\ \| & & \downarrow{\scriptstyle f_1 \oplus f_2} \\ k[A] & \longrightarrow & k[A_1] \oplus k[A_2] \end{array}$$

In other words, there is a morphism of reconstruction systems $f : p \to q$, where p and q are the grid reconstruction systems associated to (B, D_2) and (A, D_1) respectively.

The kernel of f, following the previous lemma, is the reconstruction system

$$0 \longrightarrow \ker(f_1) \oplus \ker(f_2).$$

We will make use of a pair of reconstruction systems very similar to these in chapter 5. In that case p can be understood quite well and $\ker(f)$ has a very explicit structure as we have described just now. These together can then be used to understand the system q, using the results we will describe at the end of this section.

Lemma 2.3.17. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then $p'(\mathrm{im}(f_T)) \subset \mathrm{im}(f_P)$. Hence p' induces a k-linear map $\mathrm{cok}(f_T) \to \mathrm{cok}(f_P)$.*

Proof. Let $x \in T$. Then we have $p'(f_T(x)) = f_P(p(x))$, so $p'(\mathrm{im}(f_T))$ is indeed contained in $\mathrm{im}(f_P)$. Let $y \in \mathrm{cok}(f_T)$ and let $z, z' \in T'$ be two lifts of y. Then $z' = z + f_T(x)$ for some $x \in T$ and so we have

$$p'(z') = p'(z + f_T(x)) = p'(z) + p'(f_T(x)) = p'(z) + f_P(p(x)).$$

Therefore, the class of $p'(z)$ in $\mathrm{cok}(f_P)$ does not depend on the chosen list z, only on y. One checks readily that this map induced by p' is indeed k-linear. □

Definition 2.3.18. Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then a *cokernel* of f is a reconstruction system $\mathrm{cok}(f)$ together with a morphism $\pi : p' \to \mathrm{cok}(f)$ such that $\pi \circ f = 0$ and such that for every reconstruction system q and morphism $g : p' \to q$ satisfying $g \circ f = 0$, there is a unique morphism $h : \mathrm{cok}(f) \to q$ such that $g = h \circ \pi$.

Like kernels, we define cokernels as objects satisfying a universal property. Again, we have to show that they exist and are unique.

Lemma 2.3.19. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then a cokernel of this morphism exists and is given by the map*

$$\mathrm{cok}(f) : \mathrm{cok}(f_P) \longrightarrow \mathrm{cok}(f_T)$$

induced by p'. Moreover, if (q_1, π_1) and (q_2, π_2) are two cokernels of f, the unique morphism $\phi : q_1 \to q_2$ such that $\pi_2 = \phi \circ \pi_1$ is an isomorphism.

Proof. By lemma (2.3.17), p' induces a k-linear map $\mathrm{cok}(f_T) \to \mathrm{cok}(f_P)$ as claimed. Let π_T be the quotient map $T' \to \mathrm{cok}(f_T)$ and π_P the quotient map $P' \to \mathrm{cok}(f_P)$. One sees at once that $\pi = (\pi_T, \pi_P)$ is a morphism of reconstruction systems and that it satisfies $\pi \circ f = 0$.

We will now show that the pair $(\mathrm{cok}(f), \pi)$ satisfies the universal property of the cokernel as in definition (2.3.18). Let $q : T_q \to P_q$ be some other reconstruction system with a morphism $g : p' \to q$ satisfying $g \circ f = 0$. Then we have $g_T \circ f_T = 0$, and so g_T induces a well-defined map $h_T : \mathrm{cok}(f_T) \to T_q$. Similarly, g_P induces a well-defined map $h_P : \mathrm{cok}(f_P) \to P_q$. One sees easily that $h = (h_T, h_P)$ is again a morphism of reconstruction systems and that it satisfies $g = h \circ \pi$.

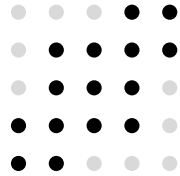
Suppose h' is another morphism $\mathrm{cok}(f) \to q$ such that $g = h' \circ \pi$. Then $h'_T = h_T$ as π_T is surjective, and similarly for h'_P. Thus h is the unique morphism so that $g = h \circ \pi$ and therefore the pair $(\mathrm{cok}(f), \pi)$ satisfies the universal property of the cokernel as claimed. The uniqueness of the cokernel is another application of lemma (2.3.4). □

Example 2.3.20. Continuing from example (2.3.15), the structure of $\mathrm{cok}(f)$ is relatively simple in this case too. The cokernel of the inclusion map

$$k[B] \longrightarrow k[A]$$

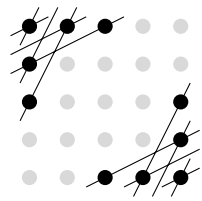
is isomorphic to k, the isomorphism given by taking a function $t \in k[A]$ and evaluating it at the point $(3, 3)$. In the directions we consider in this example, the sets of lines going through A and B are the same, so f_P is the identity morphism, whose cokernel is 0. It follows that $\mathrm{cok}(f)$ is the reconstruction system $k \to 0$.

Example 2.3.21. A more involved cokernel comes up in a similar situation to the previous example. Consider again the situation as in example (1.4.1) and let A' be the following subset of A.



As explained in example (2.2.2) there is a morphism of reconstruction systems $f : p \to q$, where p and q are the grid reconstruction systems associated to (A', D) and (A, D) respectively.

The cokernel of the map $f_T : k[A'] \to k[A]$ is the free k-module on the set $A \setminus A'$, as the image of $k[A']$ inside $k[A]$ consists precisely of the maps that are 0 on those points. For $\mathrm{cok}(f_P)$ things are a little more complicated. We are looking now for which lines in the directions we consider go through points of A, but not through any points of A'. It turns out there are some, but not many. They are shown in the picture below.



The cokernel of f is the reconstruction system $r : \mathrm{cok}(f_T) \to \mathrm{cok}(f_P)$. It is similar to a grid reconstruction system, in that $\mathrm{cok}(f_T)$ is the space of functions $A \setminus A' \to k$ and the map r takes line sums of these functions. What is different is that we do not consider all lines in the projection directions. Most notably there are no horizontal or vertical lines at all, but there are also lines in the directions $(1, 2)$ and $(2, 1)$ which intersect $A \setminus A'$ but then also go through points in A', so they are not considered.

Nevertheless, the reconstruction system r is rather simple and straightforward computations show that $\ker(r)$ is trivial and $\mathrm{cok}(r)$ is a free k-module of rank 2. The two independent dependencies for r express the fact that the line sums for lines going through the upper left point are equal, as are those going through the lower right point.

For a morphism of k-modules, the cases where the kernel or cokernel is the zero module correspond to the morphism being injective or surjective. For reconstruction systems, these notions do not make sense directly, as the morphisms are not simply maps between sets. The following lemmas show how certain other properties of injective or surjective maps also hold for those morphisms of reconstruction systems whose kernel or cokernel are 0.

Lemma 2.3.22. *Let $f : p \to p'$ be a morphism of reconstruction systems over k, let $(\ker(f), \iota)$ be its kernel and $(\operatorname{cok}(f), \pi)$ be its cokernel. Then the following are equivalent:*

1. *$\iota = 0$;*
2. *$\ker(f) = 0$;*
3. *f_T and f_P are injective;*
4. *(p, f) is the kernel of π;*
5. *(p, f) is the kernel of some morphism $g : p' \to p''$;*
6. left-cancellation*: for all reconstruction systems q and morphisms $g_1 : q \to p$ and $g_2 : q \to p$, we have $g_1 = g_2$ if and only if $f \circ g_1 = f \circ g_2$.*

We call a morphism that satisfies these equivalent criteria injective.

Proof. We will prove that each item implies the next, and that the last implies the first.

$1 \Rightarrow 2$ The zero morphism $0 : \ker(f) \to p$ satisfies $f \circ 0 = 0$, and so there must be a unique morphism $\phi : \ker(f) \to \ker(f)$ such that $\phi \circ \iota = 0$. As $\iota = 0$, both $\phi = 0$ and $\phi = \mathrm{id}_{\ker(f)}$ would satisfy this, so they must be the same and therefore $\ker(f) = 0$.

$2 \Rightarrow 3$ As $\ker(f) = 0$, we have $\ker(f_T) = 0$ and $\ker(f_P) = 0$. For k-linear maps, the kernel is 0 if and only if the map is injective.

$3 \Rightarrow 4$ The kernel of the quotient map $\pi_T : T' \to \operatorname{cok}(f_T)$ is $\operatorname{im}(f_T)$. As f_T is injective, it gives rise to an isomorphism

$$T \to \operatorname{im}(f_T) = \ker(\pi_T).$$

Similarly, f_P gives rise to an isomorphisms

$$P \to \operatorname{im}(g_T) = \ker(\pi_P).$$

Together, these identify (p, f) with the kernel of π.

$4 \Rightarrow 5$ Trivial.

$5 \Rightarrow 6$ Suppose we have a reconstruction system q and morphisms $g_1 : q \to p$ and $g_2 : q \to p$ such that $f \circ g_1 = f \circ g_2$. Write h for this morphism $q \to p'$. Note that it satisfies

$$g \circ h = g \circ f \circ g_1 = 0 \circ g_1 = 0,$$

so by the universal property of the kernel, there is a unique ϕ such that $h = f \circ \phi$. By construction, both $\phi = g_1$ and $\phi = g_2$ satisfy this requirement, so $g_1 = g_2$ as required.

$6 \Rightarrow 1$ Note that $f \circ \iota$ is the zero map $\ker(f) \to p'$ by definition of the kernel. So we have $f \circ \iota = f \circ 0$, where 0 is the zero map $\ker(f) \to p$. By left-cancellation, it follows that $\iota = 0$. □

Lemma 2.3.23. *Let $f : p \to p'$ be a morphism of reconstruction systems over k, let $(\ker(f), \iota)$ be its kernel and $(\operatorname{cok}(f), \pi)$ be its cokernel. Then the following are equivalent:*

1. *$\pi = 0$;*
2. *$\operatorname{cok}(f) = 0$;*
3. *f_T and f_P are surjective;*
4. *(p', f) is the cokernel of ι;*
5. *(p', f) is the cokernel of some morphism $g : p'' \to p$;*
6. right-cancellation*: for all reconstruction systems q and morphisms $g_1 : p' \to q$ and $g_2 : p' \to q$, we have $g_1 = g_2$ if and only if $g_1 \circ f = g_2 \circ f$.*

We call a morphism that satisfies these equivalent criteria surjective.

Proof. We will prove that each item implies the next, and that the last implies the first. The astute reader may observe a strong similarity between the proofs of this lemma and the previous one.

$1 \Rightarrow 2$ The zero morphism $p' \to \operatorname{cok}(f)$ can be written both as the $0 \circ \iota$ and as $\mathrm{id}_{\operatorname{cok}(f)} \circ \iota$. By the universal property of the cokernel, there is a unique ϕ such that $0 = \phi \circ \iota$, so it follows that $\mathrm{id}_{\operatorname{cok}(f)} = 0$ and therefore, $\operatorname{cok}(f) = 0$.

$2 \Rightarrow 3$ As $\mathrm{cok}(f) = 0$, we have $\mathrm{cok}(f_T) = 0$ and $\mathrm{cok}(f_P) = 0$. For k-linear maps, the cokernel is 0 if and only if the map is surjective.

$3 \Rightarrow 4$ The map f_T induces an isomorphism $T/\ker(f_T) \to T'$. Note that $T/\ker(f_T)$ is $\mathrm{cok}(\iota_T)$. Similarly, f_P induces an isomorphism

$$\mathrm{cok}(\iota_P) \to P'.$$

Together these give an isomorphism $\mathrm{cok}(\iota) \to p'$, from which the required result follows.

$4 \Rightarrow 5$ Trivial.

$5 \Rightarrow 6$ Suppose we have a reconstruction system q and morphisms $g_1 : p' \to q$ and $g_2 : p' \to q$ such that $g_1 \circ f = g_2 \circ f$. Write h for this morphism $p \to q$. Note that it satisfies

$$h \circ g = g_1 \circ f \circ g = g_1 \circ 0 = 0,$$

so by the universal property of the cokernel, there is a unique ϕ such that $h = \phi \circ f$. By construction, both $\phi = g_1$ and $\phi = g_2$ satisfy this requirement, so $g_1 = g_2$ as required.

$6 \Rightarrow 1$ Note that $\pi \circ f$ is the zero map $p' \to \mathrm{cok}(f)$ by definition of the cokernel. So we have $\pi \circ f = 0 \circ f$, where 0 is the zero map $p' \to \mathrm{cok}(f)$. By right-cancellation, it follows that $\pi = 0$. □

Corollary 2.3.24. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then f is an isomorphism if and only if* $\ker(f) = 0$ *and* $\mathrm{cok}(f) = 0$.

Proof. By lemma (2.2.7), f is an isomorphism if and only if f_T and f_P are both isomorphisms. A k-linear map is an isomorphism if and only if it is surjective and injective. By the preceding lemmas, f_T and f_P are both surjective and injective if and only if $\ker(f) = 0$ and $\mathrm{cok}(f) = 0$. □

Theorem 2.3.25. *Let k be a commutative ring. Then the category of reconstruction systems over k is an Abelian category.*

Proof. There are very many equivalent definitions of what an Abelian category is. Here we will follow Lang's Algebra, [17, III.3]. It states that the requirements are as follows.

1. The set of morphisms $\mathrm{Hom}(p, p')$ is an abelian group for every pair of reconstruction systems p and p'. This is a consequence of lemma (2.3.1).

2. The composition law is bilinear. This is lemma (2.3.2).

3. There exists a zero object that is an initial and final object in the category. This is remark (2.3.3).

4. Finite products and coproducts exist. This follows from the construction of the direct sum and lemmas (2.3.7) and (2.3.9).

5. Kernels and cokernels (as defined by their universal properties) exist. This was shown in lemmas (2.3.14) and (2.3.19).

6. If $f : p \to p'$ is a morphism whose kernel is 0, then f is the kernel of its cokernel. This is point 4 from lemma (2.3.22).

7. If $f : p \to p'$ is a morphism whose cokernel is 0, then f is the cokernel of its kernel. This is point 4 from lemma (2.3.23).

8. If $f : p \to p'$ is a morphism whose kernel and cokernel are 0, the f is an isomorphism. This is corollary (2.3.24).

As the category of reconstruction systems over k satisfies all these properties, it is an Abelian category. □

2.4 Exactness properties

The notion of an *exact sequence* is a very powerful extension to the language of (commutative) diagrams. It can be defined in any Abelian category. We will first explain what exact sequences of k-modules are and then expand this notion to sequences of reconstruction systems over k. After that, we will explore to what extent the functors KER, COK and DEP we defined before respect exactness.

This section, just like the previous one, discusses results related to the basics of homological algebra. In particular, lemma (2.4.10) can be seen as a special case of the connection homomorphisms that are associated to a short exact sequence of complexes. Some references relating to homological algebra were given at the start of the previous section, we will not re-iterate them here.

A sequence of k-modules

$$M' \longrightarrow M \longrightarrow M''$$

is called *exact* at M if the image of the first map is the kernel of the second. A longer sequence

$$M_1 \longrightarrow M_2 \longrightarrow \cdots \longrightarrow M_n$$

is called *exact* if it is exact at $M_2, \ldots, M_{n-1}$.

A morphism of k-modules $M \to N$ is injective if and only if the sequence

$$0 \longrightarrow M \longrightarrow N$$

is exact. Similarly, it is surjective if and only if

$$M \longrightarrow N \longrightarrow 0$$

is exact. Every k-linear map $f : M \to N$ fits into an exact sequence

$$0 \to \ker(f) \to M \to N \to \operatorname{cok}(f) \to 0.$$

Perhaps the most useful type of exact sequence is a so called short exact sequence. It is one of the form

$$0 \to M' \to M \to M'' \to 0.$$

Such a sequence expresses the fact that M' can be seen as a submodule of M and that the quotient by this submodule is M''.

To carry the definition of an exact sequence over to reconstruction systems, we need to define what the image of a morphism of reconstruction systems is. The following lemma relates the naive definition to a more categorical construction.

Lemma 2.4.1. *Let $f : p \to p'$ be a morphism of reconstruction systems over k. Then p' induces a reconstruction system $\operatorname{im}(f) : \operatorname{im}(f_T) \to \operatorname{im}(f_P)$. The inclusion maps $\operatorname{im}(f_T) \to T'$ and $\operatorname{im}(f_P) \to P'$ give rise to a morphism*

$$\iota : \operatorname{im}(f) \to p'.$$

The pair $(\operatorname{im}(f), \iota)$ is the kernel of the morphism $\pi : p' \to \operatorname{cok}(f)$.

Proof. It was shown in lemma (2.3.17) that $p'(\operatorname{im}(f_T))$ is contained in $\operatorname{im}(f_T)$, so p' indeed induces a reconstruction system $\operatorname{im}(f_T) \to \operatorname{im}(f_P)$. By construction, ι is a morphism of reconstruction systems.

Let $\pi : p' \to \mathrm{cok}(f)$ be the cokernel of f. Then π_T is the map

$$T' \to \mathrm{cok}(f_T) = T'/\mathrm{im}(f_T),$$

whose kernel clearly is $\mathrm{im}(f_T)$. Likewise $\ker(\pi_P) = \mathrm{im}(f_P)$. From the proof of lemma (2.3.14) we conclude that $\mathrm{im}(f)$ satisfies the universal property of the kernel of π. □

We can now define exact sequences of reconstruction systems in much the same way as we did for k-modules. A sequence of morphisms

$$p' \longrightarrow p \longrightarrow p''$$

is called exact at p if the image of the first is (i.e. satisfies the universal property of) the kernel of the second. A longer sequence is called exact if it is exact at every intermediate step.

Lemma 2.4.2. *Let $p_i : T_i \to P_i$ for $i \in \{1, 2, 3\}$ be reconstruction systems over k and let $f : p_1 \to p_2$ and $g : p_2 \to p_3$ be morphisms of reconstruction systems. Then*

$$p_1 \underset{f}{\longrightarrow} p_2 \underset{g}{\longrightarrow} p_3$$

is exact at p_2 if and only if

$$T_1 \underset{f_T}{\longrightarrow} T_2 \underset{g_T}{\longrightarrow} T_3$$

and

$$P_1 \underset{f_P}{\longrightarrow} P_2 \underset{g_P}{\longrightarrow} P_3$$

are exact as sequences of k-modules.

Proof. This follows at once from the description of the kernel given in (2.3.14) and the description of the image given in lemma (2.4.1). □

Lemma 2.4.3. *Let $p_i : T_i \to P_i$ for $i \in \{1, 2, 3\}$ be reconstruction systems over k and let $f : p_1 \to p_2$ and $g : p_2 \to p_3$ be morphisms of reconstruction systems. Then*

$$0 \longrightarrow p_1 \underset{f}{\longrightarrow} p_2 \underset{g}{\longrightarrow} p_3$$

is exact if and only if (p_1, f) is the kernel of g. Likewise,

$$p_1 \underset{f}{\longrightarrow} p_2 \underset{g}{\longrightarrow} p_3 \longrightarrow 0$$

is exact if and only if (p_3, g) *is the cokernel of* f*.*

Proof. For the first claim, note that exactness at p_1 is equivalent to saying f is injective, which is in turn equivalent to saying f induces an isomorphism $p_1 \to \mathrm{im}(f)$. Exactness at p_2 is equivalent to saying $\mathrm{im}(f) = \ker(g)$. The result follows.

For the second claim, note that exactness at p_3 is equivalent to saying g is surjective. Let ι be the inclusion map $\ker(g) \to p_2$. By lemma (2.3.23), surjectivity of g is equivalent to having $p_3 = \mathrm{cok}(\iota)$. Exactness at p_2 is equivalent to having $\mathrm{im}(f) = \ker(g)$ and the result follows. □

The following lemma describes a short exact sequence for grid reconstruction systems. We have already seen this construction in examples (2.3.20) and (2.3.21).

Lemma 2.4.4. *Let* $r \geq 2$ *and* D *an* n*-tuple of directions in* $\mathbb{Z}^r$*. Let* $A \subset \mathbb{Z}^r$ *and* $B \subset A$*. Let* k *be a commutative ring and let* p_A *and* p_B *be the grid reconstruction systems over* k *corresponding to* (A, D) *and* (B, D) *respectively. Let* $p_{A/B}$ *be the natural map*

$$p_{A/B} : k[A \setminus B] \longrightarrow \bigoplus_{i=1}^{n} k[A_i \setminus B_i],$$

where A_i *and* B_i *are the images of* A *and* B *in* $\mathbb{Z}^r/d_i$*. Then there is a short exact sequence of reconstruction systems*

$$0 \to p_B \to p_A \to p_{A/B} \to 0.$$

Proof. The sequences

$$0 \to k[B] \to k[A] \to k[A \setminus B] \to 0$$

and

$$0 \to k[B_i] \to k[A_i] \to k[A_i \setminus B_i] \to 0$$

are clearly exact. The diagram

$$\begin{array}{ccc} k[B] & \longrightarrow & k[A] \\ \Big\downarrow{\scriptstyle p_B} & & \Big\downarrow{\scriptstyle p_A} \\ \bigoplus_{i=1}^{n} k[B_i] & \longrightarrow & \bigoplus_{i=1}^{n} k[A_i] \end{array}$$

commutes, as was shown in example (2.2.2). The diagram

$$\begin{array}{ccc} k[A] & \longrightarrow & k[A \setminus B] \\ \Big\downarrow{\scriptstyle p_A} & & \Big\downarrow{\scriptstyle p_{A/B}} \\ \bigoplus_{i=1}^{n} k[A_i] & \longrightarrow & \bigoplus_{i=1}^{n} k[A_i \setminus B_i] \end{array}$$

commutes, as any line in $A_i \setminus B_i$ meets only points in $A \setminus B$. □

In the remainder of this section, we look at what the functors KER, COK and DEP do with exact sequences of reconstruction systems. A surprising relation between the kernels and cokernels of a short exact sequence, known as the *snake lemma*, is proved (see corollary (2.4.12)) and some consequences are derived from it.

Lemma 2.4.5. *Let*

$$0 \to p_1 \to p_2 \to p_3$$

be an exact sequence of reconstruction systems over k. Then the induced sequence of k-modules

$$0 \to \ker(p_1) \to \ker(p_2) \to \ker(p_3)$$

is also exact.

Proof. Fixing notation, we let p_i be the reconstruction system $T_i \to P_i$ for $i \in \{1, 2, 3\}$ and write f for the morphism $p_1 \to p_2$ and g for the morphism $p_2 \to p_3$.

Let ι_i be the inclusion map $\ker(p_i) \subset T_i$ for $i \in \{1, 2, 3\}$. Then the definition KER of a morphism implies that the following diagram is commutative.

$$\begin{array}{ccccccc} 0 & \longrightarrow & \ker(p_1) & \longrightarrow & \ker(p_2) & \longrightarrow & \ker(p_3) \\ & & \Big\downarrow & & \Big\downarrow & & \Big\downarrow \\ 0 & \longrightarrow & T_1 & \longrightarrow & T_2 & \longrightarrow & T_3 \end{array}$$

It is given that the bottom row of the diagram is exact, and we have to prove that the top row is as well. Note that $f_T \circ \iota_1$ is injective, as it is the composition of two injective maps. As the left square in the diagram

commutes, we have $f_T \circ \iota_1 = \iota_2 \circ \mathrm{KER}(f)$ and so $\mathrm{KER}(f)$ must also be injective. This shows exactness at $\ker(p_1)$.

Let $x \in \ker(p_2)$. As ι_2 is injective and the right square in the diagram commutes, we see that $[\mathrm{KER}(g)](x) = 0$ if and only if $g_T(\iota_2(x)) = 0$. By the exactness of the bottom row, this happens if and only if $\iota_2(x)$ is in the image of f_T. What remains to be shown is that if $\iota_2(x) = f_T(y)$, then $y \in \ker(p_1)$. Note that $p_2(\iota_2(x)) = 0$, so we have $p_2(f_T(y)) = 0$. As f is a morphism, it follows that $f_P(p_1(y)) = 0$. As f_P is injective, it follows that $y \in \ker(p_1)$. □

Example 2.4.6. Following examples (1.4.1) and (1.4.2), let

$$A = \{1, \ldots, 5\} \times \{1, \ldots, 5\}$$

and $B = A \setminus \{(3,3)\}$. We computed in example (2.3.20) the cokernel of the morphism of reconstruction systems $p_B \to p_A$ where p_A and p_B are the grid reconstruction systems over $\mathbb{Z}$ corresponding to (A, D) and (B, D) respectively. Following the names from lemma (2.4.4), we write $p_{A/B}$ for this cokernel.

Applying lemma (2.4.5) to the short exact sequence

$$0 \to p_B \to p_A \to p_{A/B} \to 0$$

we find that there is an exact sequence

$$0 \to \ker(p_B) \to \ker(p_A) \to \ker(p_{A/B}).$$

In example (1.4.1) we saw that $\ker(p_A)$ is isomorphic to $\mathbb{Z}$, and in example (1.4.2) we concluded that $\ker(p_B)$ is trivial. We computed in example (2.3.20) that $p_{A/B}$ is the reconstruction system $\mathbb{Z} \to 0$, whose kernel is clearly $\mathbb{Z}$. Thus the only information we still need about the sequence of kernels is which map $\mathbb{Z} \to \mathbb{Z}$ is the one between $\ker(p_A)$ and $\ker(p_{A/B})$.

On the level of T-spaces, the morphism $p_A \to p_{A/B}$ gives rise to the linear map $\mathbb{Z}[A] \to \mathbb{Z}[A \setminus B]$ sending a function $f : A \to \mathbb{Z}$ to its restriction to $A \setminus B$. In this case $A \setminus B$ contains only the single point $(3,3)$, so the map is in fact evaluation of a function $f : A \to \mathbb{Z}$ in the point B. We saw the generator of $\ker(p_A)$ in example (1.4.1). Its value in $(3,3)$ is -2. Hence the exact sequence of kernels is the sequence

$$0 \longrightarrow 0 \longrightarrow \mathbb{Z} \underset{\cdot -2}{\longrightarrow} \mathbb{Z}.$$

This sequence is indeed exact, and we see it would not be exact if we added a zero at the end, as multiplication by -2 is not surjective on $\mathbb{Z}$.

We invite the reader to revisit example (2.3.21) and see for themselves that in that case one obtains an exact sequence of kernels

$$0 \longrightarrow \mathbb{Z} \xrightarrow[\text{id}]{} \mathbb{Z} \longrightarrow 0.$$

Lemma 2.4.7. *Let*

$$p_1 \to p_2 \to p_3 \to 0$$

be an exact sequence of reconstruction systems over k. Then the induced sequence of k-modules

$$\operatorname{cok}(p_1) \to \operatorname{cok}(p_2) \to \operatorname{cok}(p_3) \to 0$$

is also exact.

Proof. Fixing notation, we let p_i be the reconstruction system $T_i \to P_i$ for $i \in \{1, 2, 3\}$ and write f for the morphism $p_1 \to p_2$ and g for the morphism $p_2 \to p_3$.

For $i \in \{1, 2, 3\}$, let π_i be the quotient map $P_i \to \operatorname{cok}(p_i)$. From the definition of COK of a morphism, it follows that the diagram below is commutative.

$$\begin{array}{ccccccc} P_1 & \longrightarrow & P_2 & \longrightarrow & P_3 & \longrightarrow & 0 \\ \downarrow & & \downarrow & & \downarrow & & \\ \operatorname{cok}(p_1) & \longrightarrow & \operatorname{cok}(p_2) & \longrightarrow & \operatorname{cok}(p_3) & \longrightarrow & 0 \end{array}$$

It is given that the top row of this diagram is exact and we want to prove that the bottom row is as well. Note that $\pi_3 \circ g_P$ is surjective, as it is the composition of two surjective maps. As the right square in the diagram commutes, this means $\text{COK}(g) \circ \pi_2$ is also surjective, so $\text{COK}(g)$ is surjective. This shows the exactness at $\operatorname{cok}(p_3)$.

We prove exactness at $\operatorname{cok}(p_2)$ in two steps. First, if $x \in \operatorname{cok}(p_1)$, then there is a $y \in P_1$ such that $\pi_1(y) = x$. By the commutativity of the diagram, it follows that

$$[\text{COK}(g) \circ \text{COK}(f)](x) = \pi_3([g_P \circ f_P](y)) = \pi_3(0) = 0.$$

Secondly, if $x \in \mathrm{cok}(p_2)$ satisfies $[\textsc{cok}(g)](x) = 0$, we want to show that x is in the image of $\textsc{cok}(f)$. Let $y \in P_2$ be such that $\pi_2(y) = x$. Note that $\pi_3(g_P(y)) = 0$ by the commutativity of the right square of the diagram. Hence there is some $z \in T_3$ such that $p_3(z) = g_P(y)$. As g_T is surjective, there is a $u \in T_2$ such that $g_T(u) = z$. Let $v = y - p_2(u)$. This element satisfies $\pi_2(v) = \pi_2(y) = x$ and

$$g_P(v) = g_P(y) - g_P(p_2(u)) = g_P(y) - p_3(g_T(u)) = g_P(y) - p_3(z) = 0.$$

By the exactness of the top row at P_2, it follows that there is a $w \in P_1$ such that $f_P(w) = v$. Now we have

$$[\textsc{cok}(f)](\pi_1(w)) = \pi_2(f_P(w)) = \pi_2(v) = x$$

as required. □

Example 2.4.8. This is a continuation of example (2.3.16). One sees easily that there is a short exact sequence of reconstruction systems

$$0 \longrightarrow \ker(f) \longrightarrow p \xrightarrow[f]{} q \longrightarrow 0$$

where $\ker(f)$ is the kernel computed in that earlier example. Applying lemma (2.4.7) we obtain an exact sequence of k-modules

$$\ker(f_1) \oplus \ker(f_2) \longrightarrow \mathrm{cok}(p) \longrightarrow \mathrm{cok}(q) \longrightarrow 0.$$

In other words, $\mathrm{cok}(q)$ is the quotient of $\mathrm{cok}(p)$ by the image inside that module of $\ker(f_1) \oplus \ker(f_2)$.

The content of lemma (2.4.5) can be described by saying that the functor KER is *left-exact*. Likewise, lemma (2.4.7) shows that COK is *right-exact*. For DEP the nomenclature is inevitably a bit confusing, as contravariant functors invert the direction of arrows. The convention is to say that the following lemma shows DEP is left-exact.

Lemma 2.4.9. *Let*

$$p_1 \to p_2 \to p_3 \to 0$$

be an exact sequence of reconstruction systems over k. Then the induced sequence of k-modules

$$0 \to \mathrm{Dep}(p_3) \to \mathrm{Dep}(p_2) \to \mathrm{Dep}(p_1)$$

is also exact.

Proof. The map

$$\textsc{Dep}(g) : \mathrm{Dep}(p_3) \longrightarrow \mathrm{Dep}(p_2)$$

is given by sending $d : \mathrm{cok}(p_3) \to k$ to $d \circ \textsc{cok}(g)$. Lemma (2.4.7) shows that $\textsc{cok}(g)$ is surjective, which implies that $\textsc{Dep}(g)$ is surjective, showing the exactness at $\mathrm{Dep}(p_3)$.

For the exactness at $\mathrm{Dep}(p_2)$, note first that

$$\textsc{Dep}(f) \circ \textsc{Dep}(g) = \textsc{Dep}(g \circ f) = \textsc{Dep}(0) = 0$$

holds by the (contravariant) functoriality of Dep. Conversely, suppose that $d : \mathrm{cok}(p_2) \to k$ is a dependency satisfying $d \circ \textsc{cok}(f) = 0$. Let $x \in \mathrm{cok}(p_3)$ and let $y \in \mathrm{cok}(p_2)$ be any element so that $[\textsc{cok}(g)](y) = x$. These exist by functoriality. If we have another such element y', then $y - y'$ is in the kernel of $\textsc{cok}(g)$ and therefore in the image of $\textsc{cok}(f)$, by lemma (2.4.7). Hence $y' = y + [\textsc{cok}(f)](z)$ for some $z \in \mathrm{cok}(p_1)$ and we have

$$d(y') = d(y) + d([\textsc{cok}(f)](z)) = d(y) + 0 = d(y).$$

It follows that $d(y)$ only depends on x, not on the chosen lift y, i.e., there is a well-defined map $d' : \mathrm{cok}(p_3) \to k$ given by x goes to $d(y)$. More or less by construction d' satisfies $d = d' \circ \textsc{cok}(g)$. One sees easily that d' is k-linear, and so we conclude that $d = [\textsc{Dep}(g)](d')$ is in the image of $\textsc{Dep}(g)$, which confirms the exactness as $\mathrm{Dep}(p_2)$. □

Lemma 2.4.10. *Let*

$$0 \to p_1 \to p_2 \to p_3 \to 0$$

be an exact sequence of reconstruction systems over k. Then there is a k-linear map

$$\delta : \ker(p_3) \to \mathrm{cok}(p_1)$$

so that

$$\ker(p_2) \to \ker(p_3) \to \mathrm{cok}(p_1) \to \mathrm{cok}(p_2)$$

is exact.

Proof. Fixing notation, we let p_i be the reconstruction system $T_i \to P_i$ for $i \in \{1,2,3\}$ and write f for the morphism $p_1 \to p_2$ and g for the morphism $p_2 \to p_3$. For $i \in \{1,2,3\}$, let ι_i be the inclusion map $\ker(p_i) \subset T_i$ and let π_i be the quotient map $P_i \to \operatorname{cok}(p_i)$.

We first define the map δ. Let $x \in \ker(p_3)$. From the surjectivity of g_T, it follows that there is a $y \in T_2$ such that $g_T(y) = x$. Note that

$$g_P(p_2(y)) = p_3(g_T(y)) = p_3(x) = 0,$$

so $p_2(y) \in \ker(g_P) = \operatorname{im}(f_P)$. As f_P is injective, it follows that there is a unique $z \in P_1$ (depending on y) such that $f_P(z) = p_2(y)$. We put

$$\delta(x) = \pi_1(z).$$

To show that this is well-defined, we have to check that this value does not depend on the choice of y we made along the way. For any other $y' \in T_2$ with $g_T(y') = x$, we have $y - y' \in \ker(g_T) = \operatorname{im}(f_T)$, so there is a $w \in T_1$ such that $y' = y + f_T(w)$. We then have

$$p_2(y') = p_2(y) + p_2(f_T(w)) = p_2(y) + f_P(p_1(w)).$$

It follows that $z' = z + p_1(w)$ satisfies $f_P(z') = p_2(y')$. As f_P is injective, this is the unique z' corresponding to y'. It follows that

$$\pi_1(z') = \pi_1(z) + \pi_1(p_1(w)) = \pi_1(z),$$

so $\delta(x)$ is indeed well-defined.

We first prove exactness at $\ker(p_3)$. Suppose x is in the image of $\textsc{ker}(g)$, then per definition we have $x = g_T(y)$ for some $y \in \ker(p_2)$. It follows that in the construction above $p_2(y) = 0$, so $z = 0$ and $\delta(x) = 0$. Hence the image of $\textsc{ker}(g)$ is contained in the kernel of δ.

Suppose x satisfies $\delta(x) = 0$. Then $\pi_1(z) = 0$, so there is a $w \in T_1$ such that $z = p_1(w)$. Put $y' = y - f_T(w)$. It satisfies $g_T(y') = x$ and

$$p_2(y') = p_2(y) - p_2(f_T(w)) = p_2(y) - f_P(p_1(w)) = p_2(y) - f_P(z) = 0,$$

so $y' \in \ker(p_2)$. Hence x is in the image of $\textsc{ker}(g)$. Hence the sequence from the lemma is exact at $\ker(p_3)$.

We now prove exactness at $\operatorname{cok}(p_1)$. For every $x \in \ker(p_3)$, we see that

$$[\textsc{cok}(f)](\delta(x)) = [\textsc{cok}(f)](\pi_1(z)) = \pi_2(f_p(z)) = \pi_2(p_2(y)) = 0,$$

so the image of δ is contained in the kernel of $\textsc{cok}(f)$.

Now let $a \in \text{cok}(p_1)$ and suppose that $[\textsc{cok}(f)](a) = 0$. Let $b \in P_1$ be such that $\pi_1(b) = a$. Then

$$\pi_2(f_P(b)) = [\textsc{cok}(f)](\pi_1(b)) = [\textsc{cok}(f)](a) = 0,$$

so there is a $c \in T_2$ such that $p_2(c) = f_P(b)$. Let $x = g_T(c)$. Then

$$p_3(x) = p_3(g_T(c)) = g_P(p_2(c)) = g_P(f_P(b)) = 0,$$

so $x \in \ker(p_3)$, moreover $y = c$ satisfies $g_T(y) = x$ and $z = b$ satisfies $f_P(z) = p_2(y)$, so $\delta(x) = \pi_1(z) = a$. We conclude that the sequence from the lemma is exact at $\text{cok}(p_1)$. □

Example 2.4.11. Continuing with an example we have seen many times now, let A and B be as in examples (1.4.1) and (1.4.2). As we remarked in example (2.4.6), there is a short exact sequence of reconstruction systems over $\mathbb{Z}$

$$0 \longrightarrow p_B \longrightarrow p_A \longrightarrow p_{A/B} \longrightarrow 0.$$

We computed in that example that $\ker(p_A)$ and $\ker(p_{A/B})$ are both isomorphic to $\mathbb{Z}$ and that the map between them is the multiplication by -2.

We follow now the construction from lemma (2.4.10) of the map

$$\delta : \ker(p_{A/B}) \to \text{cok}(p_B).$$

It is enough to check what happens to the element $1 \in \mathbb{Z} = \ker(p_{A/B})$. Its lift to $\mathbb{Z}[A]$ is the function $f : A \to \mathbb{Z}$ that sends $(3, 3)$ to 1 and all other points to 0. Its vector of line sums $p_A(f)$ is

$$v \in \mathbb{Z}[A_1] \oplus \cdots \oplus \mathbb{Z}[A_4]$$

assigning 1 to the four lines going through $(3, 3)$ and zero to all other lines. As we have $B_i = A_i$ for all i, the lift of v to $\bigoplus_{i=1}^{4} \mathbb{Z}[B_i]$ is again just v. We saw already in example (1.4.2) that this vector of line sums represents a non-trivial element of the cokernel: there is no integral valued function $g : B \to \mathbb{Z}$ that has $p_B(g) = v$.

On the other hand, when we compute $\delta(2)$, we find $2v$ and as remarked in example (1.4.2), this for these line sums we can find a function $B \to \mathbb{Z}$,

and so $\delta(2) = 0$ in $\mathrm{cok}(p_B)$. This is in accordance with the exactness of the sequence

$$\mathbb{Z} \xrightarrow[\cdot -2]{} \mathbb{Z} \xrightarrow[\delta]{} \mathrm{cok}(p_B) \longrightarrow \mathrm{cok}(p_A)$$

from lemma (2.4.10).

Corollary 2.4.12. *Let*

$$0 \to p_1 \to p_2 \to p_3 \to 0$$

be an exact sequence of reconstruction systems over k. Then there is a long exact sequence

$$\begin{array}{ccccccccc} 0 & \to & \ker(p_1) & \to & \ker(p_2) & \to & \ker(p_3) & & \\ & \to & \mathrm{cok}(p_1) & \to & \mathrm{cok}(p_2) & \to & \mathrm{cok}(p_3) & \to & 0. \end{array}$$

The preceding corollary is known as the *snake lemma*. The motivation for this is that if we capture it in a large diagram, the map δ 'snakes' through the diagram from the top right to the bottom left. We conclude the section with some consequences of this lemma.

Corollary 2.4.13. *Let*

$$0 \to p_1 \to p_2 \to p_3 \to 0$$

be an exact sequence of reconstruction systems over k. Then

$$\ker(p_2) \to \ker(p_3) \text{ is surjective}$$

if and only if

$$\mathrm{cok}(p_1) \to \mathrm{cok}(p_2) \text{ is injective.}$$

Lemma 2.4.14. *Let $p : T \to P_1$ and $p : T \to P_2$ be reconstruction systems over k. Let $p : T \to P_1 \oplus P_2$ be the reconstruction system sending $t \in T$ to $(p_1(t), p_2(t))$. Then*

$$\ker(p) = \ker(p_1) \cap \ker(p_2)$$

and there is a short exact sequence of k-modules

$$0 \to T/(\ker(p_1) + \ker(p_2)) \to \mathrm{cok}(p) \to \mathrm{cok}(p_1) \oplus \mathrm{cok}(p_2) \to 0.$$

Proof. It is clear that $(p_1(t), p_2(t)) = (0,0)$ if and only if $t \in \ker(p_1) \cap \ker(p_2)$. Consider the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & T & \longrightarrow & T \oplus T & \longrightarrow & T & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle p_1 \oplus p_2} & & \downarrow & & \\ 0 & \longrightarrow & P_1 \oplus P_2 & = & P_1 \oplus P_2 & \longrightarrow & 0 & \longrightarrow & 0 \end{array}$$

where the maps on the top row are $t \mapsto (t,t)$ and $(s,t) \mapsto s - t$. It is commutative and has exact rows. This means we can also see it as a short exact sequence of reconstruction systems. Applying the snake lemma (2.4.12) we obtain an exact sequence

$$0 \to \ker(p) \to \ker(p_1) \oplus \ker(p_2) \to T \to \operatorname{cok}(p) \to \operatorname{cok}(p_1) \oplus \operatorname{cok}(p_2) \to 0$$

and the result follows. □

2.5 Change of rings

Sometimes it is useful to consider (grid) reconstruction systems that are 'the same' except the coefficients come from different rings k. We saw this, for example, in the proof Hajdu and Tijdeman's reconstruction theorem for rectangular domains (1.3.3). Base change is the mathematical concept that formalises these ideas.

The construction that lies at the heart of base change is the tensor product of modules. It is debatable whether or not knowledge of tensor products falls within the 'basic understanding of modern algebra' that was claimed to be the prerequisite for this chapter. Nevertheless, we will not explain the tensor product here, but refer the reader to any good textbook on rings and modules, for example [17, Ch. 16]. We will also refer to this text whenever we need some elementary result on tensor products.

We begin by defining base change for reconstruction systems and explain how it can be used to detect torsion in the cokernel of reconstruction systems over $\mathbb{Z}$. This torsion is of interest as it captures an obstruction to consistency that is not seen by the dependencies of the reconstruction system.

The later part of the section is devoted to Galois descent. When we use base change to 'change coefficients' in a reconstruction system, it is also convenient to have a way to go back. When the two rings involved form a

Galois extension of fields, this can be done with the descent technique we will describe.

Definition 2.5.1. Let k and K be commutative rings and let $f : k \to K$ be a ring homomorphism. Then there is a covariant functor $\mathrm{BC}_{K/k}$ from reconstruction systems over k to reconstruction systems over K sending a reconstruction system $p : T \to P$ to the system $\mathrm{BC}_{K/k}(p) : K \otimes_k T \to K \otimes_k P$. We call this functor the *base change* functor from k to K.

Example 2.5.2. Let k and K be commutative rings and let $f : k \to K$ be a ring homomorphism. Tensor products of free modules are again free (see e.g. [17, Ch. 16, Prop. 4.1]), i.e. for any set X one has $K \otimes_k k[X] = K[X]$. The upshot of this is that the base change of a grid reconstruction system is again a grid reconstruction system. More explicitly, we have the following.

Let $A \subset \mathbb{Z}^r$ and D a sequence of directions in $\mathbb{Z}^r$. Let p_k be the grid reconstruction system over k associated to (A, D) and let p_K the one over K associated to (A, D). Then the isomorphisms

$$K \otimes_k k[A] \longrightarrow K[A]$$

and

$$K \otimes_k k[A_i] \longrightarrow K[A_i]$$

give rise to an isomorphism of reconstruction systems over K

$$\mathrm{BC}_{K/k}(p_k) \longrightarrow p_K.$$

In other words, for grid reconstruction systems, base change really comes down to changing the base ring over which we define the reconstruction system.

Lemma 2.5.3. *Let k and K be commutative rings and let $f : k \to K$ be a ring homomorphism. Let p be a reconstruction system over k. Then there is a natural isomorphism*

$$K \otimes_k \mathrm{cok}(p) = \mathrm{cok}(\mathrm{BC}_{K/k}(p)).$$

Proof. The function $K \otimes_k -$ is a right-exact functor from k-modules to K-modules, see for example [17, Ch. 16, prop. 2.6]. Hence the exact sequence of k-modules

$$T \underset{p}{\longrightarrow} P \longrightarrow \operatorname{cok}(p) \longrightarrow 0$$

gives rise to an exact sequence

$$K \otimes_k T \underset{\mathrm{BC}_{K/k}(p)}{\longrightarrow} K \otimes_k P \longrightarrow K \otimes_k \operatorname{cok}(p) \longrightarrow 0.$$

We conclude that $K \otimes_k \operatorname{cok}(p)$ is naturally isomorphic to the cokernel of $\mathrm{BC}_{K/k}(p)$. □

When we go from cokernels to dependencies, the relation between $\operatorname{Dep}(p)$ and $\operatorname{Dep}(\mathrm{BC}_{K/k}(p))$ is not so straightforward. However in the cases where $\operatorname{cok}(p)$ is a free k-module of finite rank, the situation is nicer. This is also when we are most interested in dependencies, as they solve the consistency problem in these cases, cf. corollary (2.1.19).

Lemma 2.5.4. *Let k and K be commutative rings and let $f : k \to K$ be a ring homomorphism. Let p be a reconstruction system over k. Then there is a well-defined K-linear map*

$$\begin{array}{ccc} K \otimes_k \operatorname{Dep}(p) & \longrightarrow & \operatorname{Dep}(\mathrm{BC}_{K/k}(p)) \\ \lambda \otimes d & \longmapsto & [\mu \otimes x \mapsto \lambda\mu d(x)]. \end{array}$$

If $\operatorname{cok}(p)$ is a free k-module of finite rank, this map is an isomorphism.

Proof. Let $d \in \operatorname{Dep}(p)$ and $\lambda \in K$. The map

$$\begin{array}{ccc} K \times \operatorname{cok}(p) & \longrightarrow & K \\ (\mu, x) & \longmapsto & \lambda\mu d(x) \end{array}$$

is well-defined and k-bilinear and therefore gives rise to a map

$$\phi_{\lambda,d} : K \otimes_k \operatorname{cok}(p) \longrightarrow K.$$

One checks that this map is K-linear. Moreover, we have

$$K \otimes_k \operatorname{cok}(p) = \operatorname{cok}(\mathrm{BC}_{K/k}(p))$$

by lemma (2.5.3). It follows that there is a well-defined map

$$\begin{array}{ccc} K \times \operatorname{Dep}(p) & \longrightarrow & \operatorname{Dep}(\mathrm{BC}_{K/k}(p)) \\ (\lambda, d) & \longmapsto & \phi_{\lambda,d}. \end{array}$$

One checks easily that this map is k-bilinear and therefore induces a map $K \otimes_k \mathrm{Dep}(p) \to \mathrm{Dep}(\mathrm{BC}_{K/k}(p))$ satisfying the formula claimed in the lemma. It is easily seen to be K-linear.

Suppose now that $c_1, \ldots, c_n$ is a k-basis of $\mathrm{cok}(p)$. Then $1 \otimes c_1, \ldots, 1 \otimes c_n$ is a K-basis of $\mathrm{cok}(\mathrm{BC}_{K/k}(p))$. Let $d_i \in \mathrm{Dep}(p)$ be the k-linear map $\mathrm{cok}(p) \to k$ that sends c_j to 1 if $i = j$ and to 0 otherwise. The d_i form a k-basis of $\mathrm{Dep}(p)$. Let e_i be the K-linear map $\mathrm{cok}(\mathrm{BC}_{K/k}(p)) \to K$ that sends $1 \otimes c_j$ to 1 if $i = j$ and to 0 otherwise. The e_i form a K-basis of $\mathrm{Dep}(\mathrm{BC}_{K/k}(p))$. It is clear that the map from the lemma sends $1 \otimes d_i$ to e_i for all i, hence it sends a K-basis of the domain to a K-basis of the codomain, so it is an isomorphism. □

The functor $K \otimes_k -$ from k-modules to K-modules is right-exact but not in general left-exact, so the natural map

$$K \otimes_k \ker(p) \to \ker(\mathrm{BC}_{K/k}(p))$$

is not in general an isomorphism. For grid reconstruction systems, we can relate the cases where it is an isomorphism to the cases where the cokernel is free. These are of interest because of corollary (2.1.19), which says that if the cokernel is free, the dependencies are the only obstruction to consistency.

Theorem 2.5.5. *Let A be a finite subset of $\mathbb{Z}^r$ and let D be a sequence of directions in $\mathbb{Z}^r$. For a commutative ring k, denote by p_k the grid reconstruction system over k associated to (A, D). Suppose that $\mathrm{cok}(p_{\mathbb{Z}})$ is finitely generated. Then the following are equivalent*

1. *$\mathrm{cok}(p_{\mathbb{Z}})$ is a free $\mathbb{Z}$-module;*
2. *$\mathrm{cok}(p_k)$ is a free k-module for all k;*
3. *$\ker(p_k) = k \otimes \ker(p_{\mathbb{Z}})$ for all k;*
4. *$\ker(p_k) = k \otimes \ker(p_{\mathbb{Z}})$ for all $k = \mathbb{F}_p$ with p is a prime number.*

Proof. We prove that each point implies the next and that the last implies the first.

$1 \Rightarrow 2$ This follows immediately from lemma (2.5.3).

$2 \Rightarrow 3$ We derive this from the following observation. Let A and B be free $\mathbb{Z}$-modules and $f : A \to B$ be a surjective homomorphism. Then there is a homomorphism $s : B \to A$ such that $f \circ s = \mathrm{id}_B$. Such an s can be found by taking basis for B and for each basis element, picking an element of A that maps to it. Note that there is a well-defined map

$$A \longrightarrow \ker(f) \qquad x \mapsto x - s(f(x))$$

as $f(x - s(f(x))) = f(x) = [f \circ s](f(x)) = 0$ for all $x \in A$. This map is onto, as it restricts to the identity on $\ker(f)$. Its kernel is the image of s. Hence $\ker(f) = \mathrm{cok}(s)$. For any commutative ring k, let f_k and s_k be the base-changes of f and s to k. Then $f_k \circ s_k = \mathrm{id}_{B \otimes_{\mathbb{Z}} k}$ and so $\ker(f_k) = \mathrm{cok}(s_k)$. Hence $\ker(f_k) = \ker(f) \otimes_{\mathbb{Z}} k$.

Apply this observation to the quotient map $\pi_{\mathbb{Z}} : P_{\mathbb{Z}} \to \mathrm{cok}(p_{\mathbb{Z}})$. We conclude that $\ker(\pi_k) = \ker(\pi_{\mathbb{Z}}) \otimes_{\mathbb{Z}} k$. Note that $\ker(\pi_k)$ is the image of p_k. It follows that the surjective map $p_k : T_k \to \mathrm{im}(p_k)$ is the base change to k of the surjective map $p_{\mathbb{Z}} : T_{\mathbb{Z}} \to \mathrm{im}(p_{\mathbb{Z}})$. Since $\mathrm{im}(p_{\mathbb{Z}})$ is a submodule of the free $\mathbb{Z}$-module $P_{\mathbb{Z}}$, it is free. Hence we can apply the observation again to this map and obtain $\ker(p_k) = \ker(p_{\mathbb{Z}}) \otimes_{\mathbb{Z}} k$ as required.

$3 \Rightarrow 4$ Trivial.

$4 \Rightarrow 1$ Since $\mathrm{cok}(p_{\mathbb{Z}})$ is finitely generated, it suffices to show that it is torsion-free (see e.g. [17, Ch. 1, Thm. 8.4]). Let p be a prime number. We will show that $\mathrm{cok}(p_{\mathbb{Z}})$ has no p-torsion. Note that there is a short exact sequence of reconstruction systems

$$0 \to p_{\mathbb{Z}} \to p_{\mathbb{Z}} \to p_{\mathbb{F}_p} \to 0,$$

given by the multiplication-by-p maps $T_{\mathbb{Z}} \to T_{\mathbb{Z}}$ and $P_{\mathbb{Z}} \to P_{\mathbb{Z}}$ and the natural quotient maps $T_{\mathbb{Z}} \to T_{\mathbb{F}_p}$ and $P_{\mathbb{Z}} \to P_{\mathbb{F}_p}$. The given equality $\ker(p_{\mathbb{F}_p}) = \mathbb{F}_p \otimes \ker(p_{\mathbb{Z}})$ is equivalent to saying that the sequence of kernels

$$0 \to \ker(p_{\mathbb{Z}}) \to \ker(p_{\mathbb{Z}}) \to \ker(p_{\mathbb{F}_p}) \to 0$$

is exact. By corollary (2.4.13) this is in turn equivalent to the map $\mathrm{cok}(p_{\mathbb{Z}}) \to \mathrm{cok}(p_{\mathbb{Z}})$ being injective. Since this map is the multiplication

by p on $\mathrm{cok}(p_{\mathbb{Z}})$, it follows that $\mathrm{cok}(p_{\mathbb{Z}})$ has no p-torsion. Since p was arbitrary, we conclude that $\mathrm{cok}(p_{\mathbb{Z}})$ is torsion-free. □

Example 2.5.6. Consider again the configuration of points and lines from example (1.4.2). Let $p_{\mathbb{Z}}$ be the associated grid reconstruction system over $\mathbb{Z}$ and $p_{\mathbb{F}_2}$ the associated grid reconstruction system over $\mathbb{F}_2$.

We have seen before that $\mathrm{cok}(p_{\mathbb{Z}})$ has an element of order 2, so we expect the other criteria from the theorem above to fail as well. It was also established in (1.4.2) that $\ker(p_{\mathbb{Z}})$ is trivial. However, the following table in $\mathbb{F}_2[B]$

$$\begin{array}{ccccc} 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{array}$$

has all line sums 0, so $\ker(p_{\mathbb{F}_2})$ is a non-trivial $\mathbb{F}_2$ vector space. Looking at the generator of the kernel over $\mathbb{Z}$ when we do not exclude the center point, as in example (1.4.1), we see that it contains a value -2 at the central point. When we exclude this central point, the kernel over $\mathbb{Z}$ collapses, but over $\mathbb{F}_2$, the point is not in the support of the kernel generator, so it survives as the table mentioned above.

Galois actions

A morphism of rings $f : k \to K$ also gives us a way to view any reconstruction system over K as one over k. Any K-module M can be viewed as a k-module by putting

$$xm = f(x)m \quad \text{for all } x \in k \text{ and } m \in M.$$

The obvious functoriality of this construction implies it also gives rise to a functor from reconstruction systems over K to those over k. We call this functor, like its counterpart for modules, the *restriction of scalars* from K to k. We write $\mathrm{RES}_{K/k}$ for it.

Note that when the ring homomorphism $f : k \to K$ is injective, the functor $\mathrm{RES}_{K/k}$ expresses something very simple: whenever we have a K-module M, we can 'forget' that we can multiply elements of M with elements

of $K \setminus k$ and be left with a k-module. Very often in the literature such forgetting is done silently and one simply speaks of the k-module M. While this is often convenient, we find the precision that comes with the explicit notation more important than convenience in this case.

We begin this section with a short description of the relation between base change and restriction of scalars, first for modules, then for reconstruction systems. A more precise version of these statements would be that the two functors are *adjoint* to each other. See e.g. [8, App. A5] for more details. Example [A5.2.2 e] in that book treats the case of modules, for which the case of reconstruction systems is derived easily.

Lemma 2.5.7. *Let $f : k \to K$ be a ring homomorphism between commutative rings. Let M be a k-module and N a K-module. Then the map*

$$\begin{array}{ccc} \mathrm{Hom}_k(M, \mathrm{RES}_{K/k}(N)) & \longrightarrow & \mathrm{Hom}_K(\mathrm{BC}_{K/k}(M), N) \\ \phi & \longmapsto & [\lambda \otimes m \mapsto \lambda\phi(m)] \end{array}$$

is an isomorphism of abelian groups.

Proof. The map

$$\begin{array}{ccc} K \times M & \longrightarrow & N \\ (\lambda, m) & \longmapsto & \lambda f(m) \end{array}$$

is clearly k-bilinear, hence it indeed induces a map $\mathrm{BC}_{K/k}(M) \to N$. It is clear that this map is K-linear. This shows the map from the lemma is well-defined. One checks easily it is a group homomorphism. To show it is bijective, one verifies directly that

$$\begin{array}{ccc} \mathrm{Hom}_K(\mathrm{BC}_{K/k}(M), N) & \longrightarrow & \mathrm{Hom}_k(M, \mathrm{RES}_{K/k}(N)) \\ \psi & \longmapsto & [m \mapsto g(1 \otimes m)] \end{array}$$

is a two-sided inverse. □

Remark 2.5.8. Let $f : k \to K$ be a ring homomorphism between commutative rings. Let $p : T \to P$ be a reconstruction system over k and $q : T' \to P'$ a reconstruction system over K. Let $\phi : p \to \mathrm{RES}_{K/k}(q)$ be a morphism of reconstruction systems over k. Then the K-linear maps

$$\begin{array}{rccc} \psi_T : & K \otimes_k T & \longrightarrow & T' \\ & \lambda \otimes t & \longmapsto & \lambda\phi_T(t) \end{array}$$

and

$$\begin{array}{rccc} \psi_P : & K \otimes_k P & \longrightarrow & P' \\ & \lambda \otimes x & \longmapsto & \lambda\phi_P(x) \end{array}$$

give rise to a morphism $\psi : \mathrm{BC}_{K/k}(p) \to q$ of reconstructions systems over K.

Corollary 2.5.9. *Let $f : k \to K$ be a ring homomorphism between commutative rings. Let p be a reconstruction system over k and q a reconstruction system over K. The the map*

$$\mathrm{Hom}_k(p, \mathrm{RES}_{K/k}(q)) \longrightarrow \mathrm{Hom}_K(\mathrm{BC}_{K/k}(p), q)$$

defined in remark (2.5.8) *is an isomorphism of abelian groups.*

For a general ring homomorphism $k \to K$, these results are the extent of what we can say about the relation between a reconstruction system and its base change. However, when $k \to K$ is a finite Galois extension of fields, the restriction of scalars of a base change comes with a rich additional structure, which allows us to give something close to an inverse of the base change construction.

Remark 2.5.10. In any category, the isomorphisms from an object X to itself are called the automorphisms of X. They form a group under composition. We write $\mathrm{Aut}(X)$ for this group. An *action* of a group G on the object X is a group homomorphism $\rho : G \to \mathrm{Aut}(X)$. We usually suppress ρ in the notation and write g for the morphism $\rho(g)$. A morphism $f : X \to Y$ between objects with G-action is called *G-equivariant* or *compatible* with the G-actions, if $\rho_Y(g) \circ f = f \circ \rho_X(g)$ holds for all $g \in G$.

When M is a k-module with an action of the group G, we write M^G for the set of invariant elements, i.e. elements $m \in M$ so that $[\rho(g)](m) = m$ for all $g \in G$.

When $k \to K$ is a Galois extension of fields, the k-module $\mathrm{RES}_{K/k}(K)$ comes with a natural action of $\mathrm{Gal}(K/k)$: any element of the Galois group is a ring homomorphism $K \to K$ which fixes k, so it is a k-linear map $K \to K$. Moreover, Galois theory tells us that $K^{\mathrm{Gal}(K/k)} = k$. This can be generalised to the following result on modules (i.e. vector spaces).

Lemma 2.5.11. *Let $k \to K$ be a Galois extension of fields and let M be a k-module. Let N be the k-module $\mathrm{RES}_{K/k}(K \otimes_k M)$. Then N comes with a natural action of the Galois group $G = \mathrm{Gal}(K/k)$. The set of invariants satisfies*

$$N^G = \{1 \otimes m \mid m \in M\}.$$

Proof. For any $\sigma \in G$, the map

$$\begin{array}{ccc} K \times M & \longrightarrow & K \otimes_k M \\ (\lambda, m) & \longmapsto & \sigma(\lambda) \otimes m \end{array}$$

is well-defined and k-bilinear and thus induces a map $\rho(\sigma)$ from $K \otimes_k M$ to itself. One checks easily that $\rho(\sigma)$ is k-linear and that $\rho(\sigma\tau) = \rho(\sigma) \circ \rho(\tau)$ for all σ and τ in G. It follows that $\rho(\sigma)$ is an automorphism of $\mathrm{Res}_{K/k}(K \otimes_k M)$ and that ρ gives an action of G on this k-module.

Let $(m_i)_{i \in I}$ be a k-basis of M (recall that every k-module, i.e., every k vector space, has a basis). The sequence $(1 \otimes m_i)_{i \in I}$ is a K-basis of $K \otimes_k M$ (see e.g. [17, Ch. 16, Prop. 4.1]), so every element x of the set $K \otimes_k M$ can be represented in a unique way as a sum

$$x = \sum_{i \in I} \lambda_i m_i$$

with the $\lambda_i \in K$ almost all 0. For $\sigma \in G$, we compute

$$[\rho(\sigma)](x) = \sum_{i \in I} \sigma(\lambda_i) m_i.$$

It follows that $\rho(\sigma(x)) = x$ for all $\sigma \in G$ if and only if $\sigma(\lambda_i) = \lambda_i$ for all $\sigma \in G$ and $i \in I$, i.e. if and only if $\lambda_i \in k$ for all $i \in I$. The required result follows. □

For a reconstruction system, one cannot really speak of elements that are fixed by an automorphism. Nevertheless, we can make sense of the reconstruction system fixed by the group action.

Lemma 2.5.12. *Let $p : T \to P$ be a reconstruction system over k and suppose it has an action of a group G. Then G acts on T and P, and one has $p(T^G) \subset P^G$. Thus there is a reconstruction system*

$$p^G : T^G \longrightarrow P^G$$

that comes with a natural injective morphism $\iota : p^G \to p$.

Proof. Let $\sigma \in G$ and let f_σ be the corresponding automorphism of p. Then $f_{\sigma,T}$ is an automorphism of T and we have $f_{\sigma\tau,T} = f_{\sigma,T} \circ f_{\tau,T}$, so there is a well-defined action of G on T by sending σ to $f_{\sigma,T}$. In the same way $\sigma \mapsto f_{\sigma,P}$ defines the action of G on P. Let $t \in T$ and $\sigma \in G$, then we have

$$p(\sigma(t)) = p(f_{\sigma,T}(t)) = f_{\sigma,P}(p(t)) = \sigma(p(t)),$$

so p is compatible with the G-actions on T and P. In particular, if $t \in T$ is fixed by every $\sigma \in G$, then so is $p(t)$.

It follows that the restriction of p to T^G lands inside P^G. We write p^G for this reconstruction system. The inclusion maps $T^G \subset T$ and $P^G \subset G$ give rise to an injective morphism $\iota : p^G \to p$. □

We now come to our central theorem, which tells us how we can recover a reconstruction system over k from its base change to K by using the Galois action. Much stronger versions of this statement are possible, but these require significantly more work and the current statement is sufficient for our needs in chapter 4.

Theorem 2.5.13. *Let $k \to K$ be a Galois extension of fields and let p be a reconstruction system over k. Let $q = \textsc{Res}_{K/k}(\textsc{bc}_{K/k}(p))$, then q has a natural action of $\mathrm{Gal}(K/k)$ and $q^{\mathrm{Gal}(K/k)} = p$.*

Proof. Let $G = \mathrm{Gal}(K/k)$ and write $p : T \to P$. By lemma (2.5.11) there are actions of G on $T' = \textsc{Res}_{K/k}(K \otimes_k T)$ and $P' = \textsc{Res}_{K/k}(K \otimes_k P)$. One checks easily that $q : T' \to P'$ is compatible with these actions. Hence there is an action of G on q.

Moreover, by lemma (2.5.11), there are bijections

$$T \to (T')^G \qquad \text{and} \qquad P \to (P')^G.$$

From the definition of q it is clear that q^G acts the same way on T as p. □

Example 2.5.14. Let $A \subset \mathbb{Z}^r$ be finite and $d_1, \ldots, d_n$ a sequence of directions in $\mathbb{Z}^r$. Let $k \to K$ be a Galois extension of fields. Let p_k and p_K be the grid reconstruction systems associated to (A, D) over k and K respectively.

It was already observed in example (2.5.2) that $p_K = \textsc{bc}_{K/k}(p_k)$. The space of functions $K[A]$ comes with a very obvious action of $\mathrm{Gal}(K/k)$, given $f : A \to k$ and $\sigma \in \mathrm{Gal}(K/k)$, we obtain a function $\sigma(f)$ by putting

$$[\sigma(f)](a) = \sigma(f(a))$$

for all $a \in A$. This action coincides with the action obtained from the isomorphism

$$K[A] = K \otimes_k k[A].$$

Similarly, the action on $K[A_1] \oplus \cdots \oplus K[A_n]$ is also given by simply applying $\sigma \in \mathrm{Gal}(K/k)$ to the function values. It is clear that taking line sums is indeed

compatible with these actions, as elements of $\mathrm{Gal}(K/k)$ are in particular k-linear.

Finally, we look at what the Galois action does on the dependencies. We already saw in lemma (2.5.4) that we can expect some additional complications with the dependencies, but these do not appear if the cokernel has finite dimension.

Lemma 2.5.15. *Let $k \to K$ be a Galois extension of fields and let p be a reconstruction system over k. Suppose that $\mathrm{cok}(p)$ is finite dimensional. Let M be the k-module*

$$M = \textsc{Res}_{K/k}(\mathrm{Dep}(\textsc{bc}_{K/k}(p))).$$

There is a natural inclusion $\mathrm{Dep}(p) \to M$ and a natural action of $\mathrm{Gal}(K/k)$ on M such that $\mathrm{Dep}(p) = M^G$.

Proof. From lemma (2.5.4) we know that there is a natural isomorphism

$$\phi : K \otimes_k \mathrm{Dep}(p) \longrightarrow \mathrm{Dep}(\textsc{bc}_{K/k}(p)).$$

This gives a natural inclusion $\mathrm{Dep}(p) \to M$ sending d to $\phi(1 \otimes d)$. We can apply lemma (2.5.11) to $K \otimes_k \mathrm{Dep}(p)$ and use the isomorphism ϕ to transport the resulting action of $\mathrm{Gal}(K/k)$ to M. From lemma (2.5.11) it is then clear that M^G is $\mathrm{Dep}(p)$. □

Remark 2.5.16. We continue with the notation from the previous lemma, and add that p is a k-linear map $p : T \to P$. The action of $G = \mathrm{Gal}(K/k)$ on M can be defined even when $\mathrm{cok}(p)$ is not finite dimensional. Lemma (2.1.16) identifies $\mathrm{Dep}(\textsc{bc}_{K/k}(p))$ with

$$\{f : K \otimes_k P \to K \mid f \circ \textsc{bc}_{K/k}(p) = 0\}.$$

By lemma (2.5.11), there are G-actions on the sets $K \otimes_k P$ and K. Using these we can define a G-action on the set above as follows. Let $\sigma \in G$ and $f : K \otimes_k P \to K$. Then the map

$$\sigma(f) := \sigma \circ f \circ \sigma^{-1}$$

is again K-linear. Moreover, from the compatibility of the G-action on $K \otimes_k T$ and $K \otimes_k P$ with the map $\textsc{bc}_{K/k}(p)$ it follows that $\sigma(f)$ is again

in $\mathrm{Dep}(\mathrm{BC}_{K/k}(p))$. One verifies easily that $f \mapsto \sigma(f)$ is k-linear and gives rise to a G-action on M. We leave it to the reader to verify that this definition of the action coincides with the one in lemma (2.5.15) in the case that $\mathrm{cok}(p)$ is finite dimensional.

Example 2.5.17. Let $A \subset \mathbb{Z}^r$ and $d_1, \ldots, d_n$ a sequence of directions in $\mathbb{Z}^r$. Let $k \to K$ be a Galois extension of fields. Let p_k and p_K be the grid reconstruction systems associated to (A, D) over k and K respectively.

We saw in lemma (2.1.17) that a dependency d of p_K, corresponds to giving for every line ℓ in

$$\mathcal{L} = A_1 \sqcup \cdots \sqcup A_n$$

a coefficient c_ℓ in K, such that for every $t \in K[A]$, one has

$$\sum_{\ell \in \mathcal{L}} p_\ell(t) c_\ell = 0.$$

Following remark (2.5.16) above, we consider d as K-linear map

$$P_K = K[A_1] \oplus \cdots \oplus K[A_n] \longrightarrow K.$$

The coefficient c_ℓ corresponds to $d(\ell)$ where we view the elements of A_i as generators of $K[A_i]$. In that remark, we saw that $\sigma \in \mathrm{Gal}(K/k)$ acts on d by sending it to $\sigma \circ d \circ \sigma^{-1}$ where the outer σ's act on K and P_K as in example (2.5.14). For $\ell \in \mathcal{L}$, it is clear that $\sigma^{-1}(\ell) = \ell$, so the coefficients corresponding to $\sigma(d)$ are $\sigma(c_\ell)$ for $\ell \in \mathcal{L}$.

In other words, the Galois group acts on the coefficients of a dependency as one would expect it acts on elements of K. From this it is clear that the Galois invariant elements are precisely those for which all the coefficients come from k and that these are precisely the coefficients of dependencies in $\mathrm{Dep}(p_k)$.

---

3 – Finite convex grids

In this chapter, we consider grid reconstruction systems where the grid is either all of $\mathbb{Z}^r$ for some $r \geq 1$ or a finite convex subset. Any grid reconstruction system can be related to a system where the grid A is 'full' (i.e. all of $\mathbb{Z}^r$) using lemma (2.4.4).

The first section of this chapter deals with computing the kernel and cokernel of the full grid reconstruction systems. Section 3.2 then states the results for finite convex grids. In section 3.3 we have collected some basic results about convex sets for the convenience of the reader. The proofs of the results announced in section 3.2 are in section 3.4. The last section focuses on the planar case, i.e. subsets of $\mathbb{Z}^2$.

The material in this chapter is an extension of the paper [2] by the author and Joost Batenburg. In that paper, only the planar case was treated. It also covered some special cases of the general results and machinery set up in the previous chapter.

3.1 Full grid

Throughout this section, let r be a positive integer and $A = \mathbb{Z}^r$. Let d be a lattice direction in $\mathbb{Z}^r$. The set of all lines in direction d is parametrised by $\mathbb{Z}^r/d$, as was remarked before in (2.1.2). Hence both the set of points and the sets of lines we consider have the structure of abelian groups. Moreover, the map that identifies a point with the line passing through it, the natural projection map $\mathbb{Z}^r \to \mathbb{Z}^r/d$, is a group homomorphism.

The upshot of this is that for any ring k the free k-modules $k[\mathbb{Z}^r]$ and $k[\mathbb{Z}^r/d]$ have the additional structure of being *group rings* and that the projection map

$$p_d : k[\mathbb{Z}^r] \to k[\mathbb{Z}^r/d]$$

is a ring homomorphism, even a k-algebra homomorphism. By extension, if D is an n-tuple of directions in $\mathbb{Z}^r$, then the grid reconstruction system associated to $(\mathbb{Z}^r, D)$ is the ring homomorphism

$$p_D : k[\mathbb{Z}^r] \to \prod_{i=1}^{n} k[\mathbb{Z}^r/d_i].$$

Briefly: what is a group ring in this context and how do we represent elements thereof? Let G be an additively written abelian group and let k be

any commutative ring. Let $g \in G$ and define u^g in

$$k[G] = \{f : G \to k \mid f(x) = 0 \text{ for almost all } x \in G\}$$

as the map

$$u^g : h \mapsto \begin{cases} 0 & \text{if } h \neq g \\ 1 & \text{if } h = g \end{cases} .$$

These u^g form a basis for the k-module $k[G]$. We define a k-algebra structure on $k[G]$ by specifying what the product of two of these basis elements is. For g, h in G, we put

$$u^g \cdot u^h = u^{g+h}.$$

One checks that u^0 is the unit element of this ring. We will usually just write 1 for it.

Example 3.1.1. From the examples in section 1.4, example (1.4.3) has the full grid $\mathbb{Z}^2$ as its domain. The group ring $k[\mathbb{Z}^2]$ is fairly easy to understand. There is a ring isomorphism

$$\begin{array}{ccc} k[\mathbb{Z}^2] & \longrightarrow & k[u, u^{-1}, v, v^{-1}] \\ f & \longmapsto & \sum_{(i,j)\in\mathbb{Z}^2} f(i,j) u^i v^j, \end{array}$$

which identifies $k[\mathbb{Z}^2]$ with a Laurent polynomial ring.

For each of the directions $d_1, \ldots, d_4$ from example (1.4.3) the quotient $\mathbb{Z}^2/d_i$ is isomorphic to $\mathbb{Z}$ and so the group rings $k[\mathbb{Z}^2/d_i]$ are all isomorphic to a univariate Laurent polynomial ring $k[w, w^{-1}]$. The table below gives for each direction an explicit isomorphism $\mathbb{Z}^2/d_i \cong \mathbb{Z}$ and the resulting description of the i-th component of the projection map as a ring homomorphism $k[u, u^{-1}, v, v^{-1}] \to k[w, w^{-1}]$.

i	d_i	$\mathbb{Z}^2/d_i \to \mathbb{Z}$	$k[u, u^{-1}, v, v^{-1}] \to k[w, w^{-1}]$
1	$(0,1)$	$(x,y) \mapsto -x$	$u \mapsto w^{-1}, v \mapsto 1$
2	$(1,2)$	$(x,y) \mapsto y - 2x$	$u \mapsto w^{-2}, v \mapsto w$
3	$(2,1)$	$(x,y) \mapsto 2y - x$	$u \mapsto w^{-1}, v \mapsto w^2$
4	$(1,0)$	$(x,y) \mapsto y$	$u \mapsto 1, v \mapsto w$

For rectangular domains $A = \{0, \ldots, n\} \times \{0, \ldots m\} \subset \mathbb{Z}^2$, Hajdu and Tijdeman in [12] associate to a function $f : A \to k$ what they call its generating

function, the bivariate polynomial

$$\sum_{(i,j)\in A} f(i,j)x^i y^j \quad \text{in } k[x,y].$$

This is certainly very similar to the approach we use here. However, they never study the projection map on the full ring as an object in itself. The restriction to non-negative exponents also complicates matters as there is this artificial 'bottom left corner' to deal with.

Using the algebraic structure on these full grid reconstruction systems allows us to describe their kernel and cokernel in a reasonably explicit manner. We state the results first, the final part of the section is devoted to proving them.

Theorem 3.1.2. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let p_D be the grid reconstruction system associated to $(\mathbb{Z}^r, D)$. Then* $\ker(p_D)$ *is the $k[\mathbb{Z}^r]$ ideal generated by*

$$F_D = (u^{d_1} - 1) \cdots (u^{d_n} - 1).$$

This result can also be phrased in a way that is very close to theorem (1.3.2). We include it here as a corollary. Note that multiplying an element $f \in k[\mathbb{Z}^r]$ with u^x for some $x \in \mathbb{Z}^r$ has the effect of translating the corresponding function $f : \mathbb{Z}^r \to k$ by x. We sometimes refer to functions of the form $u^x f$ as *translates* of f.

Corollary 3.1.3. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let p_D be the grid reconstruction system associated to $(\mathbb{Z}^r, D)$. Then a table $f \in k[\mathbb{Z}^r]$ satisfies $p_D(f) = 0$ if and only if it can be written as a k-linear combination of translates of F_D.*

Example 3.1.4. Consider the full grid reconstruction system as described in example (1.4.3) and identify $k[\mathbb{Z}^2]$ with $k[u, u^{-1}, v, v^{-1}]$ as in example (3.1.1). We can now compute F_D as a Laurent polynomial in u and v:

$$\begin{aligned} F_D &= (v-1)(uv^2-1)(u^2v-1)(u-1) \\ &= u^4v^4 - u^4v^3 - u^3v^4 + u^3v^3 - u^3v^2 + u^3v - u^2v^3 \\ &\quad +2u^2v^2 - u^2v + uv^3 - uv^2 + uv - u - v + 1. \end{aligned}$$

Considering this polynomial as a function $\mathbb{Z}^2 \to k$ we obtain, up to translation, the following table, which shows only the non-zero coefficients.

$$\begin{array}{cccccc} & & & -1 & 1 \\ & & 1 & -1 & 1 & -1 \\ & & -1 & 2 & -1 & \\ & -1 & 1 & -1 & 1 & \\ & 1 & -1 & & & \end{array}$$

Thus corollary (3.1.3) gives a description of the kernel of this reconstruction system as was announced in example (1.4.3).

We give two general results about the image of these reconstruction systems. The first one describes a large subspace of the image that plays an important role in the planar case. The second is a general structure result about the cokernel.

Theorem 3.1.5. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let p_D be the grid reconstruction system associated to $(\mathbb{Z}^r, D)$. For $i = 1, \ldots, n$, let $F_i = \prod_{j \neq i}(u^{d_j} - 1) \in k[\mathbb{Z}^r]$. Then the largest $\prod_{i=1}^n k[\mathbb{Z}^r/d_i]$-ideal contained in $\mathrm{im}(p_D)$ is*

$$\bigoplus_{i=1}^{n} p_{d_i}(F_i) \cdot k[\mathbb{Z}^r/d_i].$$

Theorem 3.1.6. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let p_D be the grid reconstruction system corresponding to $(\mathbb{Z}^r, D)$. Then there is an isomorphism of k-modules*

$$\mathrm{cok}(p_D) \cong \bigoplus_{1 \leq i < j \leq n} k[\mathbb{Z}^r]/(u^{d_i} - 1, u^{d_j} - 1).$$

Corollary 3.1.7. *An element*

$$x \in k[\mathbb{Z}^r/d_1] \oplus \cdots \oplus k[\mathbb{Z}^r/d_n]$$

is in the image of p_D if and only if $d(x) = 0$ for all $d \in \mathrm{Dep}(p_D)$.

Remark 3.1.8. A word of warning about the isomorphism from theorem (3.1.6). There is a natural k-linear map

$$\begin{array}{rccc} \phi_{i,j}: & k[\mathbb{Z}^r]/(u^{d_i}-1)\oplus k[\mathbb{Z}^r]/(u^{d_j}-1) & \longrightarrow & k[\mathbb{Z}^r]/(u^{d_i}-1,u^{d_j}-1) \\ & (f,g) & \longmapsto & f-g. \end{array}$$

This map extends to a k-homomorphism

$$\phi_{i,j}: k[\mathbb{Z}^r/d_1]\oplus\cdots\oplus k[\mathbb{Z}^r/d_n]\longrightarrow k[\mathbb{Z}^r]/(u^{d_i}-1,u^{d_j}-1)$$

by taking the zero map on all the extra components, and one easily verifies that it satisfies $\phi_{i,j}\circ p_D=0$ for all i and j. Thus it descends to a k-linear map

$$\operatorname{cok}(p_D)\longrightarrow k[\mathbb{Z}^r]/(u^{d_i}-1,u^{d_j}-1).$$

However, the isomorphism from theorem (3.1.6) is **not** the map sending an x in $\operatorname{cok}(p_D)$ to $(\phi_{i,j}(x))_{i,j}$.

We now turn our attention to proving the theorems stated above. For this, we first recall a general result about group rings. Our standard reference for algebraic results [17] does not say much about group rings beyond their definition, so we refer the reader to [20] for more information. The following lemma is well-known.

Lemma 3.1.9. *Let G be an abelian group and $H\subset G$ a subgroup. Let k be a commutative ring. Then the kernel of the natural ring homomorphism $k[G]\to k[G/H]$ is the $k[G]$-ideal*

$$I_H=(u^h-1\mid h\in H).$$

Moreover, if $h_1,\ldots,h_n$ generate H, then

$$I_H=(u^{h_1}-1,\ldots,u^{h_n}-1).$$

Proof. Write ϕ for the ring homomorphism $k[G]\to k[G/H]$. For $h\in H$ one clearly has $\phi(u^h-1)=0$, so $I_H\subset\ker(\phi)$. It remains to show the other inclusion. Fix a system of representatives $X\subset G$ for the quotient G/H. The quotient map $G\to G/H$ gives rise to a bijection $X\to G/H$ and thus ϕ sends $k[X]\subset k[G]$ bijectively to $k[G/H]$. In particular we have $\ker(\phi)\cap k[X]=\{0\}$.

Let $f\in\ker(\phi)$ and write

$$f=\sum_{i=1}^{m}\lambda_i u^{g_i}.$$

Each $g_i \in G$ can be written as $g_i = h_i + x_i$ with $h_i \in H$ and $x_i \in X$. Put

$$f' = \sum_{i=1}^{m} \lambda_i u^{x_i},$$

then we have

$$f - f' = \sum_{i=1}^{m} \lambda_i (u^{g_i} - u^{x_i})$$

and $u^{g_i} - u^{x_i} = u^{x_i}(u^{h_i} - 1)$ is in I_H. Thus $\phi(f') = \phi(f) = 0$. Moreover, f' is in $k[X]$, so in fact we must have $f' = 0$ and thus $f \in I_H$. This proves the other inclusion.

Finally, note that for any $g \in G$, we have $u^{-g} - 1 = -u^{-g}(u^g - 1)$ and for any $g, g' \in G$ we have

$$u^{gg'} - 1 = u^g(u^{g'} - 1) + (u^g - 1).$$

It follows that if h is in the subgroup generated by $h_1, \ldots, h_n$, then $u^h - 1$ is in the ideal $(u^{h_1} - 1, \ldots, u^{h_n} - 1)$. □

Corollary 3.1.10. *Let d be a direction in $\mathbb{Z}^r$ and let p_d be the projection map*

$$p_d : k[\mathbb{Z}^r] \longrightarrow k[\mathbb{Z}^r/d].$$

Then p_d is onto and its kernel is the ideal $(u^d - 1)$.

Example 3.1.11. Continuing example (3.1.1) above, we obtain from the corollary generators for the ring homomorphisms

$$k[u, u^{-1}, v, v^{-1}] \longrightarrow k[w, w^{-1}]$$

as in the table below. These can easily be verified by elementary means.

i	d_i	$k[u, u^{-1}, v, v^{-1}] \to k[w, w^{-1}]$	kernel
1	$(0, 1)$	$u \mapsto w^{-1}, v \mapsto 1$	$(v - 1)$
2	$(1, 2)$	$u \mapsto w^{-2}, v \mapsto w$	$(uv^2 - 1)$
3	$(2, 1)$	$u \mapsto w^{-1}, v \mapsto w^2$	$(u^2v - 1)$
4	$(1, 0)$	$u \mapsto 1, v \mapsto w$	$(u - 1)$

Lemma 3.1.12. *Let d, e be directions in $\mathbb{Z}^r$ that are independent. Then*

1. *the quotient $k[\mathbb{Z}^r]/(u^d - 1, u^e - 1)$ is a free k-module;*
2. *multiplication by $u^d - 1$ is an injective map on $k[\mathbb{Z}^r]/(u^e - 1)$.*

Proof. For the first part, note that we have

$$k[\mathbb{Z}^r](u^d - 1, u^e - 1) \cong k[\mathbb{Z}^r/\langle d, e\rangle]$$

by lemma (3.1.9) and the latter is a free k-module.

For the second part we use the isomorphism $k[\mathbb{Z}^r](u^e - 1) \cong k[\mathbb{Z}^r/e]$. Note that d has infinite order in $\mathbb{Z}^r/e$, as d and e are independent. Thus we are done if we prove the following. Let G be an abelian group and $g \in G$ of infinite order, then multiplication by $u^g - 1$ is injective on $k[G]$.

Suppose we have

$$(u^g - 1) \sum_{x \in G} \lambda_x u^x = 0,$$

for certain $\lambda_x \in k$ of which only finitely many are non-zero. Expanding the product above, we conclude that $\lambda_x = \lambda_{x+g}$ for all $x \in G$. If follows that if $\lambda_x \neq 0$ for some $x \in G$, then $\lambda_{x+ng} \neq 0$ for all $n \in \mathbb{Z}$. All the elements $x + ng$ are distinct, so we would have infinitely many non-zero λ's, which is impossible. Hence $\lambda_x = 0$ for all $x \in G$ and we conclude that the multplication by $u^g - 1$ is injective as required. □

Using these lemmas, we can now prove the first of the theorems from the start of the section, about the kernel of the full grid reconstruction systems. The second theorem, which describes a large part of the image, follows easily from the first.

Proof of (3.1.2). We proceed by induction on n, the number of directions in D. For $n = 1$, corollary (3.1.10) is the required statement. Suppose now the statement holds for any 2-regular sequence of $n - 1$ directions and we are given a 2-regular sequence $D = (d_1, \ldots, d_n)$. Let $D' = (d_2, \ldots, d_n)$.

Following lemma (2.4.14) we have $p_D(t) = (p_{d_1}(t), p_{D'}(t))$ for all $t \in k[\mathbb{Z}^r]$ and so

$$\ker(p_D) = \ker(p_{d_1}) \cap \ker(p_{D'}) = (u^{d_1} - 1) \cap (F_{D'})$$

by corollary (3.1.10) and the induction hypothesis. For $i = 2, \ldots, n$, the directions d_1 and d_i are independent, so the multiplication by $u^{d_i} - 1$ on $k[\mathbb{Z}^r]/(u^{d_1} - 1)$ is injective by lemma (3.1.12). Hence the multiplication by

$F_{D'}$ is injective on $k[\mathbb{Z}^r]/(u^{d_1}-1)$. So for any $x \in (u^{d_1}-1) \cap (F_{D'})$ we have $x = yF_{D'}$ and

$$yF_{D'} = x = 0 \quad \text{in } k[\mathbb{Z}^r]/(u^{d_1}-1)$$

so y is a multiple of $u^{d_1}-1$. Hence

$$(u^{d_1}-1) \cap (F_{D'}) = ((u^{d_1}-1)F_{D'}) = (F_D)$$

as required. By induction, this completes the proof. □

Proof of (3.1.5). Note that $p_{d_j}(F_i)$ is 0 whenever $j \neq i$. So the image of the ideal $(F_1, \ldots, F_n)$ of $k[\mathbb{Z}^r]$ under p_D is the ideal

$$\prod_{i=1}^{n} p_{d_i}(F_i) k[\mathbb{Z}^r/d_i].$$

Hence this ideal is certainly contained in $\mathrm{im}(p_D)$.

Suppose that I is any $\prod_{i=1}^n k[\mathbb{Z}^r/d_i]$-ideal contained in $\mathrm{im}(p_D)$ and let x be any element of it. Let $e_j \in \prod_{i=1}^n k[\mathbb{Z}^r/d_i]$ be 1 on the j-th coordinate and 0 everywhere else. Put $x_j = xe_j$. Note that $x_j \in I$ and $I \subset \mathrm{im}(p_D)$, so there is a $y_j \in k[\mathbb{Z}^r]$ such that $p_D(y_j) = x_j$.

Note that if we omit the j-th direction, y_j goes to 0, so by theorem (3.1.2), it is a multiple of F_j. Hence the ideal

$$\prod_{i=1}^{n} p_{d_i}(F_i) k[\mathbb{Z}^r/d_i]$$

contains all the x_j, and therefore it also contains $x = x_1 + \cdots + x_n$. □

For the proof of theorem (3.1.6) about the cokernel of full grid reconstruction systems, we require some more general ring-theoretic results. There is a certain similarity to the situation we have here and the formalism of regular sequences and Koszul complexes. For more information on these we refer the reader to [8, Ch. 17]. However, it seems difficult to actually relate the map we want to study to the Koszul complex. This may be related to the observation in remark (3.1.8) that the cokernel cannot described naively as a sum of difference maps.

Lemma 3.1.13. *Let R be a commutative ring and $x, y \in R$ such that multiplication by x is injective on R/y. Then the sequence*

$$\begin{array}{ccccccc} 0 \to & R/xy & \to & R/x \oplus R/y & \to & R/(x,y) & \to 0 \\ & r & \mapsto & (r,r) & & & \\ & & & (r,s) & \mapsto & r-s & \end{array}$$

is exact.

Proof. Let $r \in R$. If xr is in (y) then r is in (y) as multiplication by x is injective on R/y. Hence $(x) \cap (y) = (xy)$, which shows exactness on the left. If for some $r, s \in R$ we have $r - s \in (x, y)$ then there are $a, b \in R$ such that $r - s = ax + by$. This means that $r - ax$ is congruent to r modulo x and to s modulo y, which shows exactness in the middle. The surjectivity of the rightmost arrow is obvious. □

Lemma 3.1.14. *Let R be a commutative ring and let $x, y, z \in R$ be such that the multiplications by x and by y are injective on R/z. Then the sequence*

$$0 \to R/(x, z) \xrightarrow{y\cdot} R/(xy, z) \to R/(y, z) \to 0$$

is exact.

Proof. The diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & R/z & \xrightarrow{x\cdot} & R/z & \longrightarrow & R/(x,z) & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle y\cdot} & & \downarrow{\scriptstyle y\cdot} & & \\ 0 & \longrightarrow & R/z & \xrightarrow[xy\cdot]{} & R/z & \longrightarrow & R/(xy,z) & \longrightarrow & 0 \end{array}$$

is commutative and has exact rows. Hence we can apply the snake lemma to show that the kernel of the rightmost vertical map is 0 and the cokernel is isomorphic to $R/(y, z)$. □

Using these lemmas we will now prove theorem (3.1.6) and the subsequent corollary. We do not give an explicit description of the isomorphism from the theorem in the proof. In theory, such a description could be derived from it. In the planar case, we will derive an explicit description of the dependencies (and through them, the cokernel) in chapter 4.

Proof of (3.1.6). As in the proof of theorem (3.1.2), we proceed by induction on n, the number of directions in D. For $n = 1$, corollary (3.1.10) is the required statement: the map is onto, so the cokernel is trivial, as is the empty direct sum from the statement.

Suppose now the statement holds for any 2-regular sequence of $n - 1$ directions and we are given a 2-regular sequence $D = (d_1, \ldots, d_n)$.

Let $D' = (d_2, \ldots, d_n)$. Following lemma (2.4.14) we have

$$p_D(t) = (p_{d_1}(t), p_{D'}(t)) \qquad \text{for all } t \in k[\mathbb{Z}^r]$$

and so there is a short exact sequence

$$0 \to k[\mathbb{Z}^r]/(\ker(p_{d_1}) + \ker(p_{D'})) \to \operatorname{cok}(p_D) \to \operatorname{cok}(p_{d_1}) \oplus \operatorname{cok}(p_{D'}) \to 0.$$

By theorem (3.1.2) we have

$$k[\mathbb{Z}^r]/(\ker(p_{d_1}) + \ker(p_{D'})) = k[\mathbb{Z}^r]/(u^{d_1} - 1, F_{D'})$$

Note that for $i = 2, \ldots, n$, the directions d_1 and d_i are independent and so multiplication by $u^{d_i} - 1$ is injective on $k[\mathbb{Z}^r]/(u^{d_1} - 1)$. Hence multiplication by

$$D_i = (u^{d_i} - 1) \cdots (u^{d_n} - 1)$$

is injective on $k[\mathbb{Z}^r]/(u^{d_1} - 1)$. Applying lemma (3.1.14) with $x = D_{i+1}$, $y = u^{d_i} - 1$ and $z = u^{d_1} - 1$ we see that for $i = 2, \ldots, n-1$ there is a short exact sequence

$$0 \to k[\mathbb{Z}^r]/(u^{d_1} - 1, D_{i+1}) \to k[\mathbb{Z}^r]/(u^{d_1} - 1, D_i) \to k[\mathbb{Z}^r]/(u^{d_1} - 1, u^{d_i} - 1) \to 0$$

The space on the right is a free k-module by lemma (3.1.12) and so the short exact sequence is split and we have

$$k[\mathbb{Z}^r]/(u^{d_1} - 1, D_i) \cong k[\mathbb{Z}^r]/(u^{d_1} - 1, D_{i+1}) \oplus k[\mathbb{Z}^r]/(u^{d_1} - 1, u^{d_i} - 1).$$

It follows that there is an isomorphism of k-modules

$$k[\mathbb{Z}^r]/(u^{d_1} - 1, F_{D'}) \cong \bigoplus_{j=2,\ldots,n} k[\mathbb{Z}^r]/(u^{d_1} - 1, u^{d_j} - 1).$$

For the cokernels, we observe that by the induction hypothesis and corollary (3.1.10) we have

$$\operatorname{cok}(p_{d_1}) \oplus \operatorname{cok}(p_{D'}) = 0 \oplus \left(\bigoplus_{2 \le i < j \le n} k[\mathbb{Z}^r]/(u^{d_i} - 1, u^{d_j} - 1) \right).$$

All the summands are free k-modules by lemma (3.1.12), so the sum is a free k-module and the exact sequence

$$0 \to k[\mathbb{Z}^r]/(\ker(p_{d_1}) + \ker(p_{D'})) \to \operatorname{cok}(p_D) \to \operatorname{cok}(p_{d_1}) \oplus \operatorname{cok}(p_{D'}) \to 0$$

splits, yielding isomorphisms

$$\begin{aligned}\operatorname{cok}(p_D) &\cong k[\mathbb{Z}^r]/(\ker(p_{d_1}) + \ker(p_{D'})) \oplus \operatorname{cok}(p_{d_1}) \oplus \operatorname{cok}(p_{D'}) \\ &\cong \left(\bigoplus_{j=2,\ldots,n} k[\mathbb{Z}^r]/(u^{d_1} - 1, u^{d_j} - 1) \right) \\ &\quad \oplus \left(\bigoplus_{2 \le i < j \le n} k[\mathbb{Z}^r]/(u^{d_i} - 1, u^{d_j} - 1) \right) \\ &\cong \bigoplus_{1 \le i < j \le n} k[\mathbb{Z}^r]/(u^{d_i} - 1, u^{d_j} - 1)\end{aligned}$$

as required. □

Proof of (3.1.7). By theorem (3.1.6), the cokernel is isomorphic to a direct sum of k-modules that are themselves free by lemma (3.1.12). Hence it is free and we can apply corollary (2.1.19) to obtain the required result. □

3.2 Structure results

The primary way we will study grid reconstruction systems for grids that are not $\mathbb{Z}^r$ is by comparing them to those full grids. The connection between the 'algebraic' description of the kernel of full grids (as in theorem (3.1.2)) and the more combinatorial or geometric nature of other grids is via the *support* of kernel elements.

Definition 3.2.1. Let $f \in k[\mathbb{Z}^r]$ and write $f = \sum_{x \in \mathbb{Z}^r} c_x u^x$. Then the support of f is

$$\operatorname{supp}(f) = \{x \in \mathbb{Z}^r \mid c_x \neq 0\}.$$

Lemma 3.2.2. *Let D be a sequence of directions in $\mathbb{Z}^r$. Let $B \subset A$ be subsets of $\mathbb{Z}^r$. Let k be a commutative ring and let p_A and p_B be the grid reconstruction systems over k corresponding to (A, D) and (B, D) respectively. Then*

$$\ker(p_B) = \{f \in \ker(p_A) \mid \operatorname{supp}(f) \subset B\}.$$

Proof. By lemma (2.4.4) the natural morphism of reconstruction systems $p_B \to p_A$ is injective. Hence

$$\ker(p_B) = \ker(p_A) \cap k[B]$$

as a subset of $k[\mathbb{Z}^r]$. And $k[B]$ is precisely the subset of $f \in k[\mathbb{Z}^r]$ such that $\mathrm{supp}(f) \subset B$. □

Example 3.2.3. If we look at examples (1.4.1) and (1.4.2) we find subsets A and B of $\mathbb{Z}^2$ with $B \subset A$ as in the lemma. It was remarked in example (1.4.1) that the kernel of the reconstruction system p_A is the free abelian group on one generator, which was given in that example. One sees immediately that the support of this generator (and hence any multiple of it) is not contained in B, so the kernel of p_B is trivial.

Comparing examples (1.4.1) and (1.4.3) also yields interesting information. We computed in example (3.1.4) what the kernel of the global map $p_{\mathbb{Z}^r}$ is and we have now proved this result in theorem (3.1.2). By the lemma above $\ker(p_A)$ consists now of those elements in $\ker(p_{\mathbb{Z}^2})$ whose support is contained in A. One would hope that it suffices to look at which translates of the generator of $\ker(p_{\mathbb{Z}^2})$ have support contained in A (of which there is clearly just the one). Theorem (3.2.8) below asserts that this is indeed the case as the rectangle A is a finite, non-empty convex grid set.

There is a problem with just considering the support. As the previous lemma shows, this notion works very well in the combinatorial setting. However, on the algebraic side, the support is a very elusive thing to reason with. For example, the support may depend on the specific ring k used and it is difficult to compute the support of a product of elements of $k[\mathbb{Z}^r]$ given the supports of these elements.

Things on the whole become much nicer if we focus on the *convex hull* of the support, rather than the support itself. The problems mentioned above do not appear for these convex hulls, under some mild conditions that are satisfied by the elements of interest to us.

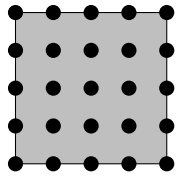
The price paid for looking at these convex hulls is that we cannot use them to describe kernels and cokernels for arbitrary grid sets A, but only for convex grid sets. For the sake of completeness, we first recall the definitions. For more information, including references, about convex sets we refer the reader to the next section.

Definition 3.2.4. A subset X of $\mathbb{R}^r$ is *convex* if the segment between any two points of X is contained in X, i.e. if for all $p, q \in X$ and all $\lambda \in [0,1]$ we have $\lambda p + (1-\lambda)q \in X$.

Definition 3.2.5. The *convex hull*, $H(X)$, of $X \subset \mathbb{R}^r$ is the smallest convex set that contains X.

Definition 3.2.6. A subset A of $\mathbb{Z}^r$ is called a *convex grid set* if it satisfies $H(A) \cap \mathbb{Z}^r = A$.

Example 3.2.7. The set A from example (1.4.1) is a convex grid set in $\mathbb{Z}^2$, as can be seen in the picture below.



It is clear that if we remove $(3,3)$ the resulting set B from example (1.4.2) is no longer a convex grid set.

The first result is about kernels of grid reconstruction systems for convex grid sets. It is a generalisation of theorem (1.3.2) from chapter 1. Note that the set S does not depend on k. From this fact one easily derives the corollary about the cokernel and dependencies. Proofs of these results are in section 3.4.

Theorem 3.2.8. *Let $A \subset \mathbb{Z}^r$ be a finite convex grid set an let D be a 2-regular sequence of directions in $\mathbb{Z}^r$. Then there is a set $S \subset A$ such that the following holds. Let k be a commutative ring that is not the zero ring, let p be the grid reconstruction system corresponding to (A, D) over k and let F_D be as in theorem* (3.1.2)*. Then*

$$S = \{s \in \mathbb{Z}^r \mid \operatorname{supp}(u^s F_D) \subset A\},$$

and the $u^s F_D$ with $s \in S$ form a k-basis of $\ker(p)$.

Example 3.2.9. Using this theorem we can prove that the description of the kernel of the reconstruction system from example (1.4.1) is correct. We computed F_D for these directions in example (3.1.4). It is clear that the set S has only one element, as there is only one translate of F_D that fits inside A. This confirms the description of the kernel given in the original example.

Corollary 3.2.10. *Let $A \subset \mathbb{Z}^r$ be a finite convex grid set, let D be a 2-regular sequence of directions in $\mathbb{Z}^r$, let k be a commutative ring and let $p : T \to P$ be the grid reconstruction system over k associated to (A, D). Then* $\mathrm{cok}(p)$ *and* $\mathrm{Dep}(p)$ *are free k-modules of finite rank. For $x \in P$ we have $x \in \mathrm{im}(p)$ if and only if $d(x) = 0$ for all $d \in \mathrm{Dep}(p)$.*

We can use theorem (3.2.8) to make an efficient reconstruction algorithm for $\mathbb{Z}$-valued functions in the same way as theorem (1.3.3) in chapter 1. This is made explicit in the corollary below.

Corollary 3.2.11. *There is a polynomial time algorithm that on input a finite convex grid set $A \subset \mathbb{Z}^r$, a 2-regular sequence D of directions in $\mathbb{Z}^r$ and $x \in P$, where $p : T \to P$ is the grid reconstruction system over $\mathbb{Z}$ associated to (A, D), outputs $y \in T$ such that $p(y) = x$ or an error if no such y exists.*

The following theorem is a bit imprecise, as some additional theory from the next section is required to formulate it exactly. In section 3.4 we will see the precise version of this in corollary (3.4.14). Intuitively, this theorem splits the consistency problem in two parts: a 'global' part that is related to the full grid and a 'local' part that only involves certain 'corners' of the space. One might hope a similar result holds for the dependencies, but the fact that the space of dependencies for the full grid reconstruction system is in general not of finite rank makes this more complicated.

Theorem 3.2.12. *Let D be a 2-regular sequence of directions in $\mathbb{Z}^r$ and let A be a finite convex grid set which is sufficiently large. Then there is an non-empty $R \subset A$ such that the following holds. Let k be a commutative ring and let p_A, p_R and p_D be the grid reconstruction systems over k corresponding respectively to (A, D), (R, D) and $(\mathbb{Z}^r, D)$. Let $p_{A/R}$ be the reconstruction system defined in lemma* (2.4.4) *for the inclusion $R \subset A$. Then the natural*

map $\mathrm{cok}(p_R) \to \mathrm{cok}(p_D)$ *is injective and there is a* k*-isomorphism*

$$\mathrm{cok}(p_A) \cong \mathrm{cok}(p_R) \oplus \mathrm{cok}(p_{A/R}).$$

3.3 Convex sets and polytopes

In this section, once again let r be a positive integer. We will cover a number of basic results on convex sets, which will be useful in the next section. Convex sets have been studied extensively in the literature, for example in connection to (linear) optimisation problems. For more on this, see e.g. [21, Ch. 2] or [3, Ch. 2]. None of the results in this section are new, though it is often difficult to find the precise statements we need in the literature. We include them here, with proofs, for the convenience of the reader. The last part of this section, starting with definition (3.3.18), describes the polytopes associated to elements of $k[\mathbb{Z}^r]$ and some of the nice properties they have.

Lemma 3.3.1. *Let* $X \subset \mathbb{R}^r$*, then the convex hull of* X *is given by*

$$H(X) = \left\{ \lambda_1 x_1 + \cdots + \lambda_n x_n \;\middle|\; n \geq 1, \begin{array}{c} x_1, \ldots, x_n \in X, \\ \lambda_1, \ldots, \lambda_n \in \mathbb{R}_{\geq 0}, \\ \lambda_1 + \cdots + \lambda_n = 1 \end{array} \right\}.$$

Proof. Let $X \subset \mathbb{R}^r$ be nonempty and put

$$H = \left\{ \lambda_1 x_1 + \cdots + \lambda_n x_n \;\middle|\; n \geq 1, \begin{array}{c} x_1, \ldots, x_n \in X, \\ \lambda_1, \ldots, \lambda_n \in \mathbb{R}_{\geq 0}, \\ \lambda_1 + \cdots + \lambda_n = 1 \end{array} \right\}.$$

It is clear that H contains X. Let $h, h' \in H$ and $\mu \in [0,1]$. Then we can write

$$h = \lambda_1 x_1 + \ldots + \lambda_n x_n$$

and

$$h' = \lambda'_1 x'_1 + \ldots + \lambda'_m x_m$$

by definition of H. Therefore we have

$$\begin{array}{rcl} \mu h + (1-\mu)h' & = & \mu(\lambda_1 x_1 + \ldots + \lambda_n x_n) + (1-\mu)(\lambda'_1 x'_1 + \ldots + \lambda'_m x_m) \\ & = & (\mu\lambda_1)x_1 + \ldots + (\mu\lambda_n)x_n \\ & & +((1-\mu)\lambda'_1)x'_1 + \ldots + ((1-\mu)\lambda'_m)x_m. \end{array}$$

All these coefficients are non-negative, and they sum up to

$$\mu(\lambda_1 + \ldots + \lambda_n) + (1-\mu)(\lambda'_1 + \ldots + \lambda'_m) = \mu + (1-\mu) = 1,$$

so $\mu h + (1-\mu)h'$ is in H. Hence H is convex.

Let $Y \subset \mathbb{R}^r$ be convex and contains X. We want to show that H is contained in Y. We do this by showing any

$$\lambda_1 x_1 + \ldots + \lambda_n x_n$$

with the λ's non-negative and summing to 1 is in Y, by induction on n. For $n = 1$, this is simply saying that $X \subset Y$, which is true by assumption. For $n > 1$, we may assume that $0 < \lambda_1 < 1$. Note that

$$\lambda_1 x_1 + \ldots + \lambda_n x_n = \lambda_1 x_1 + (1-\lambda_1)\Big(\frac{\lambda_2}{1-\lambda_1} x_2 + \ldots + \frac{\lambda_n}{1-\lambda_1} x_n\Big).$$

Note that $\frac{\lambda_i}{1-\lambda_1}$ is non-negative for every i and that

$$\frac{\lambda_2}{1-\lambda_1} + \ldots + \frac{\lambda_n}{1-\lambda_1} = \frac{\lambda_2 + \ldots + \lambda_n}{1-\lambda_1} = 1.$$

Hence by the induction hypothesis,

$$\frac{\lambda_2}{1-\lambda_1} x_2 + \ldots + \frac{\lambda_n}{1-\lambda_1} x_n$$

is in Y, and hence by convexity, so is

$$\lambda_1 x_1 + \ldots + \lambda_n x_n. \qquad \square$$

Definition 3.3.2. Let $X \subset \mathbb{R}^r$ be a convex. A point $x \in X$ is called a *vertex* of X if whenever we write $x = \lambda p + (1-\lambda)q$ with $p, q \in X$ and $\lambda \in [0,1]$, we have $x = p$ or $x = q$.

Definition 3.3.3. A subset X of $\mathbb{R}^r$ is a *polytope* if it is the convex hull of a finite set.

Lemma 3.3.4. *Let $X \subset \mathbb{R}^r$ be a polytope. Then X is closed and bounded.*

Proof. Note that a subset of $\mathbb{R}^r$ is closed and bounded if and only if it is compact. If $X = H(S)$ with $S = \{s_1, \ldots, s_t\}$ finite, then X is the image of the continuous map

$$\{(\lambda_i)_{i=1}^t \in [0,1]^t \mid \lambda_1 + \ldots + \lambda_t = 1\} \longrightarrow \mathbb{R}^r$$

sending $(\lambda_i)_{i=1}^t$ to $\lambda_1 s_1 + \ldots + \lambda_t s_t \in X$. The domain of this map is clearly compact (being closed and bounded in $\mathbb{R}^t$) and therefore its image, X, is compact too. □

Theorem 3.3.5. *Let X be a polytope. Let V be the set of all vertices of X. Then V is the smallest set such that $H(V) = X$.*

Proof. Let $Y \subset X$ with $H(Y) = X$. We want to show that $V \subset Y$. So let $v \in V$. As v is a point of X, we can write

$$v = \lambda_1 y_1 + \cdots + \lambda_n y_n$$

with the $y_i \in Y$ and the λ_i in $\mathbb{R}_{\geq 0}$ with sum 1. We choose such a representation with n minimal.

If $n = 1$, then $v = y_1$ is in Y as required.

If $n > 1$, we rewrite

$$v = \lambda_1 y_1 + (1 - \lambda_1)\Big(\frac{\lambda_2}{1-\lambda_1} y_2 + \cdots + \frac{\lambda_n}{1-\lambda_1} y_n\Big).$$

Note that $\lambda_i/(1-\lambda_1)$ is in $\mathbb{R}_{\geq 0}$ for all i, as we have $\lambda_i \geq 0$ and $\lambda_1 \leq 1$. Furthermore

$$\frac{\lambda_2}{1-\lambda_1} + \cdots + \frac{\lambda_n}{1-\lambda_1} = \frac{\lambda_2 + \cdots + \lambda_n}{1-\lambda_1} = 1,$$

so

$$x = \frac{\lambda_2}{1-\lambda_1} y_2 + \cdots + \frac{\lambda_n}{1-\lambda_1} y_n$$

is in $H(Y) = X$. Now we have $v = \lambda_1 y_1 + (1-\lambda_1)x$ and as v is a vertex, this means either $v = y_1$ or $v = x$, both contradicting the minimality of n.

This shows that any Y with $H(Y) = X$ contains V. It remains now to show that $H(V) = X$. What we will show is the following: if $X = H(Y)$ and $y \in Y$ is not a vertex of X, then $X = H(Y \setminus \{y\})$. Using this, we can remove all $y \in Y \setminus V$ one by one, starting from some finite Y with $H(Y) = X$ to arrive at $H(V) = X$.

So suppose $H(Y) = X$ with Y finite and $y \in Y$ is not a vertex. Then $y = cp + (1-c)q$ for some $c \in [0,1]$ and $p, q \in X$ with $p, q \neq y$. Write

$$p = \sum_{x \in Y} \lambda_x x$$

with the $\lambda_x \in \mathbb{R}_{\geq 0}$ with sum 1. Write

$$q = \sum_{x \in Y} \mu_x x$$

similarly. As neither p nor q is equal to y, we have $\lambda_y < 1$ and $\mu_y < 1$. Hence $\gamma = c\lambda_y + (1-c)\mu_y$ satisfies $\gamma < 1$ and we can write

$$(1-\gamma)y = \sum_{x \in Y \setminus \{y\}} (c\lambda_x + (1-c)\mu_x)x.$$

Rewriting this, we obtain

$$y = \sum_{x \in Y \setminus \{y\}} \frac{c\lambda_x + (1-c)\mu_x}{1-\gamma} x.$$

Note that the coefficients are non-negative and satisfy

$$\begin{aligned} \sum_{x \in Y \setminus \{y\}} \frac{c\lambda_x + (1-c)\mu_x}{1-\gamma} &= \frac{1}{1-\gamma}\left(c \sum_{x \in Y \setminus \{y\}} \lambda_x + (1-c) \sum_{x \in Y \setminus \{y\}} \mu_x \right) \\ &= \frac{c(1-\lambda_y) + (1-c)(1-\mu_y)}{1-\gamma} = \frac{1-\gamma}{1-\gamma} = 1. \end{aligned}$$

Hence $y \in H(Y \setminus \{y\})$ and so $X = H(Y) = H(Y \setminus \{y\})$. □

Lemma 3.3.6. *Let $X \subset \mathbb{R}^r$ be convex, closed and bounded. Let $v \in \mathbb{R}^r$ be a non-zero vector. Put*

$$m = \max_{x \in X} \langle x, v \rangle,$$

where $\langle\, , \rangle$ denotes the standard inner product on $\mathbb{R}^r$. Put

$$F = \{x \in X \mid \langle x, v \rangle = m\}.$$

We call F the face *of X in the direction v and denote it $F_v(X)$. It has the following properties*

1. *F is convex, closed and bounded;*
2. *any vertex of F is a vertex of X;*
3. *if $X = H(Y)$, then $F = H(Y \cap F)$.*

Proof. As X is bounded, any subset of X is also bounded. Let V be the hyperplane

$$V = \{x \in \mathbb{R}^r \mid \langle x, v \rangle = m\}.$$

Note that V is closed and convex. Hence $F = X \cap V$ is closed as the intersection of two closed sets, and is convex as the intersection of two convex sets. This proves (1).

Let x be a vertex of F and suppose we can write $x = \lambda p + (1-\lambda)q$ with $p, q \in X$ and $\lambda \in [0,1]$. If λ is 0 or 1, we have $x = p$ or $x = q$. Otherwise,

$$\langle x, v\rangle = \lambda\langle p, v\rangle + (1-\lambda)\langle q, v\rangle$$

can only be equal to m if both $\langle p, v\rangle = m$ and $\langle q, v\rangle = m$. But then p and q are in F and we must have $x = p$ or $x = q$ as x is a vertex of F. Hence x is a vertex of X. This proves (2).

Suppose that $X = H(Y)$ and let $x \in X$. Then we can write

$$x = \lambda_1 y_1 + \cdots + \lambda_n y_n$$

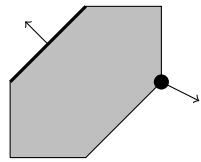
with the y_i in Y and the $\lambda_i \in \mathbb{R}_{>0}$ summing to 1. It follows that

$$\langle x, v\rangle \leq \max\{\langle y_1, v\rangle, \ldots, \langle y_n, v\rangle\}.$$

Hence $m = \max_{y\in Y}\langle y, v\rangle$ and we have $\langle x, v\rangle = m$ if and only if $\langle y_i, v\rangle = m$ for all i. This proves (3). □

Corollary 3.3.7. *Let $X \subset \mathbb{R}^r$ be a polytope and $v \in \mathbb{R}^r$ a non-zero vector. Then $F_v(X)$ is a polytope.*

Example 3.3.8. In the picture below is a polytope in $\mathbb{R}^2$ that has some of its faces highlighted.



We see that the face in direction $(-1, 1)$ is a line segment and the face in direction $(2, -1)$ is a point.

Recall that a (closed) *halfspace* in $\mathbb{R}^r$ is a set of the form

$$\{x \in \mathbb{R}^r \mid \langle x, v\rangle \geq c\}$$

for some non-zero $v \in \mathbb{R}^r$ and $c \in \mathbb{R}$.

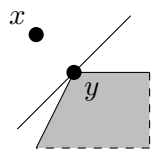
Lemma 3.3.9. *Let $X \subset \mathbb{R}^r$ be closed, bounded and convex. Then X is equal to the intersection of all halfspaces containing X.*

Proof. It is clear that X is contained in the intersection of all halfspaces containing X. Suppose that $x \in \mathbb{R}^r \setminus X$. We will show that there is a halfspace H containing X such that $x \notin H$.

As X is a compact subset of $\mathbb{R}^r$ and $x \notin X$ there is a $y \in X$ such that

$$d(x, y) = \inf \{d(x, y') \mid y' \in Y\}$$

and this distance is non-zero. Let H be the halfspace not containing x whose boundary hyperplane is the hyperplane through y perpendicular to $y - x$.



If there is some $z \in \mathbb{R}^r \setminus H$ such that $z \in X$, then the line segment between y and z contains points in X that are closer to x than y is. This contradicts the minimality of $d(x, y)$, hence X is contained in H. □

Theorem 3.3.10. *Let $X \subset \mathbb{R}^r$ be closed, bounded and convex and let ∂X be the topological boundary of X. Then ∂X is the union of all faces of X.*

Proof. Let $x \in X$. Denote by S^{r-1} the unit sphere around the origin inside $\mathbb{R}^r$. Let δ be the function

$$\begin{array}{rcl} S^{r-1} & \longrightarrow & \mathbb{R}_{\geq 0} \\ d & \mapsto & \max \{\langle v - x, d \rangle \mid v \in X\}. \end{array}$$

From the definition of the face of X in direction d, we have $x \in F_d(X)$ if and only if $\delta(d) = 0$. On the other hand, if there is an $\epsilon > 0$ such that $\delta(d) > \epsilon$ for all $d \in S^{n-1}$, then $B_\epsilon(x)$, the open ball of radius ϵ around x, is contained in all halfspaces containing X, and hence, by the previous lemma, is contained in X, so x is in the interior of X.

To prove the theorem, we therefore have to show that if $\delta(d) > 0$ for all $d \in S^{r-1}$, then there is an $\epsilon > 0$ such that $\delta(d) > \epsilon$ for all $d \in S^{r-1}$. Let

$$U_\epsilon = \{d \in S^{r-1} \mid \delta(d) > \epsilon\}$$

and suppose that $d \in U_\epsilon$. As X is closed and bounded, there is a $v \in X$ such that $\delta(d) = \langle v - x, d\rangle$. Put

$$\eta = \frac{\delta(d) - \epsilon}{||v - x||}$$

and note that for any $d' \in S^{r-1}$ with $||d - d'|| < \eta$, we have

$$|\langle v - x, d - d'\rangle| \leq ||v - x|| \cdot ||d - d'|| < ||v - x||\eta = \delta(d) - \epsilon,$$

so $\langle v - x, d'\rangle > \epsilon$. Hence $\delta(d') > \epsilon$. We conclude that the ball with radius η around d is contained in U_ϵ. Since $d \in U_\epsilon$ was arbitrary, we conclude U_ϵ is an open subset of S^{r-1}.

If $\delta(d) > 0$ for all $d \in S^{r-1}$, the U_ϵ with $\epsilon > 0$ then form an open cover of S^{r-1}. As S^{n-1} is compact, this means there must be a finite subcover. Moreover, since $U_\epsilon \subset U_{\epsilon'}$ whenever $\epsilon > \epsilon'$, there must in fact be some $\epsilon > 0$ such that $U_\epsilon = S^{r-1}$, i.e. $\delta(d) > \epsilon$ for all $d \in S^{r-1}$. □

Corollary 3.3.11. *Let $X \subset \mathbb{R}^r$ be a polytope. Let ∂X be the topological boundary of X. Then there is a finite set of directions $\{v_1, \ldots, v_n\}$ such that*

$$\partial X = \bigcup_{i=1}^{n} F_{v_i}(X).$$

Proof. After the previous theorem, what remains to be shown is that a polytope has only finitely many distinct faces. Note that by lemma (3.3.6), a face of X is uniquely identified by the set of vertices of X that lie in the face. Since a polytope only has finitely many vertices, it can only have finitely many distinct faces. □

Lemma 3.3.12. *Let $X \subset \mathbb{R}^r$ be a polytope and let $x \in X$. Then x is a vertex of X if and only if there is a non-zero $v \in \mathbb{R}^r$ such that $F_v(X) = \{x\}$.*

Proof. The 'if' part is straightforward. Suppose we have $v \in \mathbb{R}^r$ such that $F_v(X) = \{x\}$ and suppose that we also have $p, q \in X$ and $\lambda \in [0, 1]$ such that $x = \lambda p + (1 - \lambda)q$. If $p \neq x$ and $q \neq x$, we have $\langle p, v\rangle < \langle x, v\rangle$ and $\langle q, v\rangle < \langle x, v\rangle$, and so

$$\langle x, v\rangle = \lambda\langle p, v\rangle + (1 - \lambda)\langle q, v\rangle < \lambda\langle x, v\rangle + (1 - \lambda)\langle x, v\rangle < \langle x, v\rangle,$$

which is clearly a contradiction.

For the 'only if' part, we proceed by induction on r.

If $r = 1$, X is a closed and bounded interval in $\mathbb{R}$. The vertices of such an interval are the points $\max X$ and $\min X$, so either $\langle x, 1\rangle$ or $\langle x, -1\rangle$ is maximal.

For $r > 1$, note that any interior point of X is not a vertex, hence the vertex x is in ∂X. By theorem (3.3.10) there is a direction v_0 such that x is in $F_{v_0}(X)$. Let H the hyperplane perpendicular to v_0 that contains $F_{v_0}(X)$. By lemma (3.3.6) x is a vertex of $F_{v_0}(X) \subset H$ too. By the induction hypothesis there is a direction v_1 parallel to H such that x is the unique maximal point in direction v_1 of $F_{v_0}(X)$.

Let Y be the set of vertices of X that do not lie in H. Let $m = \langle x, v_0\rangle$. For all $y \in Y$, we have $\langle y, v_0\rangle < m$. Since there are only finitely many of these y, there is a $\delta > 0$ such that $\langle y, v_0\rangle < m - \delta$ for all $y \in Y$. Let $\epsilon > 0$ be such that $\epsilon|\langle p, v_1\rangle| < \frac{1}{2}\delta$ holds for every vertex p of X. Again this is possible since X has only finitely many vertices.

Now put $v = v_0 + \epsilon v_1$. Let $y \in Y$, then we have

$$\langle y, v\rangle = \langle y, v_0\rangle + \epsilon\langle y, v_1\rangle < (m - \delta) + \frac{1}{2}\delta = m - \frac{1}{2}\delta.$$

Similarly, if p is a vertex of $F_{v_0}(X)$, we have

$$\langle p, v\rangle = \langle p, v_0\rangle + \epsilon\langle p, v_1\rangle \geq m - \frac{1}{2}\delta.$$

It follows that $F_v(X)$ is a subset of $F_{v_0}(X)$. Moreover, if $p \in F_{v_0}(X)$ is not x, then

$$\langle p, v\rangle = m + \epsilon\langle p, v_1\rangle < m + \epsilon\langle x, v_1\rangle = \langle x, v\rangle,$$

so $F_v(X) = \{x\}$ as required. □

Lemma 3.3.13. *Let $X, Y \subset \mathbb{R}^r$ be convex sets. Then*

$$X + Y = \{x + y \mid x \in X, y \in Y\}$$

is also convex. Any vertex of $X + Y$ is can be written in a unique way as $x + y$ with $x \in X$ and $y \in Y$. Moreover, these x and y are vertices of X and Y respectively.

Proof. Let p be in $H(X + Y)$. Then we can write

$$p = \lambda_1(x_1 + y_1) + \cdots + \lambda_n(x_n + y_n)$$

with the x_i in X, the y_i in Y and the λ_i in $\mathbb{R}_{\geq 0}$ with sum 1. Note that

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n$$

is in $H(X) = X$, and likewise

$$y = \lambda_1 y_1 + \cdots + \lambda_n y_n$$

is in Y. It follows that $p = x + y$ is in $X + Y$. Hence $X + Y$ is convex.

Let $v \in X + Y$ be a vertex. Suppose we can write $v = x_1 + y_1$ and $v = x_2 + y_2$ with $x_{1,2} \in X$ and $y_{1,2} \in Y$. Put $p = x_1 + y_2$ and $q = x_2 + y_1$. Then we have

$$\frac{1}{2}p + \frac{1}{2}q = \frac{1}{2}(x_1 + y_1 + x_2 + y_2) = v,$$

so we must have $v = p$ or $v = q$, as v is a vertex of $X + Y$. Either one implies $x_1 = x_2$ and $y_1 = y_2$. Hence v can be represented as $x + y$ in a unique way.

Suppose x is not a vertex of X. Then x can be written as

$$x = \lambda x_1 + (1 - \lambda)x_2$$

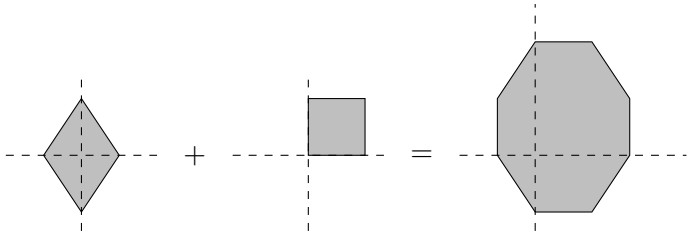
with $\lambda \in [0, 1]$ and $x_{1,2}$ not equal to x. But then

$$v = \lambda(x_1 + y) + (1 - \lambda)(x_2 + y)$$

and neither $x_1 + y$ nor $x_2 + y$ is equal to $x + y = v$. This contradicts the fact that v is a vertex. So x is a vertex of X and by a similar argument, y is a vertex of Y. □

Corollary 3.3.14. *Let $X, Y \subset \mathbb{R}^r$ be polytopes. Then $X + Y$ is a polytope.*

Example 3.3.15. Below are pictures of two polytopes in $\mathbb{R}^2$ and their sum.



Note that every vertex of the sum is indeed the sum of a vertex from each summand.

Definition 3.3.16. Let X and Y be convex, closed and bounded subsets of $\mathbb{R}^r$. Then X is Y-rounded if there is a convex, closed and bounded $Z \subset \mathbb{R}^r$ such that $X = Y + Z$.

Lemma 3.3.17. *Let $X, Y \subset \mathbb{R}^r$ be convex, closed and bounded. Then $X+Y$ is closed and bounded and for $v \in \mathbb{R}^r$ non-zero we have*

$$F_v(X + Y) = F_v(X) + F_v(Y).$$

Proof. Note that there is a surjective continuous map $X \times Y \to X + Y$ sending (x, y) to $x+y$. Since X and Y are compact (i.e. closed and bounded), so is $X \times Y$ and therefore, so is its image, $X + Y$.

Let $m_X = \max_{x \in X} \langle x, v \rangle$ and $m_Y = \max_{y \in Y} \langle y, v \rangle$. Then for $p \in X + Y$ we can write $p = x + y$ with $x \in X$ and $y \in Y$ and so we have

$$\langle p, v \rangle = \langle x, v \rangle + \langle y, v \rangle \leq m_X + m_Y$$

with equality if and only if $x \in F_v(X)$ and $y \in F_v(Y)$. Hence we have $F_v(X + Y) = F_v(X) + F_v(Y)$ as required. □

Definition 3.3.18. Let k be a commutative ring that is not the zero ring and let $f \in k[\mathbb{Z}^r]$. We call

$$P(f) := H(\mathrm{supp}(f))$$

the polytope of f.

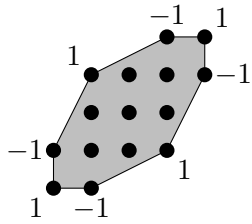
Remark 3.3.19. As the support of $f \in k[\mathbb{Z}^r]$ is always finite, $P(f)$ is indeed a polytope. Moreover, all vertices of this polytope are points in $\mathbb{Z}^r$. Thus it makes sense to talk about the coefficient of f at a vertex v of $P(f)$: it is the coefficient f_v when we write $f = \sum_{x \in \mathbb{Z}^r} f_x u^x$.

Definition 3.3.20. We say $f \in k[\mathbb{Z}^r]$ has *strong vertices* if for every vertex v of $P(f)$, the coefficient f_v is not a zero divisor in k.

Example 3.3.21. In example (3.1.4) we considered the Laurent polynomial F_D in $k[u, u^{-1}, v, v^{-1}] = k[\mathbb{Z}^2]$ given by

$$\begin{aligned} F_D &= (v-1)(uv^2-1)(u^2v-1)(u-1) \\ &= u^4v^4 - u^4v^3 - u^3v^4 + u^3v^3 - u^3v^2 + u^3v - u^2v^3 \\ &\quad +2u^2v^2 - u^2v + uv^3 - uv^2 + uv - u - v + 1. \end{aligned}$$

Suppose that k is not the zero ring. Then we see in the picture below, the polytope $P(F_D)$. Each vertex is labeled with the coefficient of the corresponding monomial.



We see that F_D has strong vertices, as 1 and -1 are not zero divisors in any non-zero commutative ring k.

Lemma 3.3.22. *Let $f, g \in k[\mathbb{Z}^r]$ and suppose f has strong vertices. Then $P(fg) = P(f) + P(g)$. Moreover, if g also has strong vertices, fg has strong vertices.*

Proof. Note that for every $d \in \mathbb{Z}^r$ the coefficient of fg at u^d is

$$\sum_{e \in \mathbb{Z}^r} f_e g_{d-e}$$

which can only be non-zero if there are $e, e' \in \mathbb{Z}^r$ with $e + e' = d$ such that $f_e \neq 0$ and $g_{e'} \neq 0$. Hence $\operatorname{supp}(fg) \subset \operatorname{supp}(f) + \operatorname{supp}(g)$, and therefore

$$P(fg) \subset P(f) + P(g).$$

For the other inclusion, let d be a vertex of $P(f) + P(g)$. Then, by lemma (3.3.13), there is only one way to write $d = e + e'$ with $e \in P(f)$ and $e' \in P(g)$. Moreover e and e' are vertices of these respective sets. Hence the coefficient of fg at u^d is $f_e g_{e'}$. As f_e is not a zero-divisor, this coefficient is nonzero, hence $d \in P(fg)$. Thus all vertices of $P(f) + P(g)$ are contained in $P(fg)$ and hence, by theorem (3.3.5), $P(f) + P(g) \subset P(fg)$. If in addition to f_e, $g_{e'}$ is also not a zero-divisor, then $f_e g_{e'}$ is not a zero-divisor. Hence if f and g both have strong vertices, then fg also has strong vertices. □

3.4 Proving the structure results

In this section we prove the structure results from section 3.2. The basic results on the kernel and cokernel follow easily by looking at the polytopes of kernel elements. The decomposition of the module of dependencies is much more involved and requires more results on convex sets that we discussed in the previous section.

Lemma 3.4.1. *Let D be a 2-regular sequence of directions in $\mathbb{Z}^r$, let k be a commutative ring that is not the zero ring and let p_D be the grid reconstruction system corresponding to $(\mathbb{Z}^r, D)$. Let F_D be the generator of* $\ker(p_D)$ *as in* (3.1.2)*. Then F_D has strong vertices and $\Delta_D = P(F_D)$ is independent of k.*

Proof. Note that F_D is the image of

$$F = (u^{d_1} - 1) \cdots (u^{d_n} - 1) \in \mathbb{Z}[\mathbb{Z}^r]$$

under the natural map $\mathbb{Z}[\mathbb{Z}^r] \to k[\mathbb{Z}^r]$. This implies $P(F_D)$ is a subset of $P(F)$. To show the reverse inclusion, we will prove that for every vertex v of $P(F)$, the coefficients F_v is ± 1. Note that this property holds for each $u^{d_i} - 1$ and that a product of two polynomials with this property again has this property, as a the coefficient at each vertex is the product of the coefficients of a vertex from each of the factors by lemma (3.3.13). Hence for every vertex v of $P(F)$, the coefficient F_v is ± 1 and so the coefficient $(F_D)_v$ is the image of ± 1 in k, which means this coefficients of F_D is non-zero, so $P(F) \subset P(F_D)$. Hence $P(F) = P(F_D)$ and the coefficients at every vertex of F_D is not a zero divisor, i.e. F_D has strong vertices. □

Proof of (3.2.8). As A is a convex grid set, the support of an element $f \in k[\mathbb{Z}^r]$ is contained in A if and only if $P(f)$ is a subset of $H(A)$. Let F_D be as in the theorem. Applying lemma (3.2.2) to the inclusion $A \subset \mathbb{Z}^r$, we conclude that $\ker(p)$ consists of those $f \in k[\mathbb{Z}^r]$ satisfying $f = F_D g$ for some $g \in k[\mathbb{Z}^r]$ and $\operatorname{supp}(f) \subset A$. As we have just observed, this will happen if and only if $P(f) \subset H(A)$.

By lemma (3.4.1), F_D has strong vertices and so, by lemma (3.3.22) we have $P(f) = P(F_D) + P(g)$. Lemma (3.4.1) also showed that $\Delta_D = P(F_D)$ does not depend on the choice of k. It follows that $P(f) \subset H(A)$ if and only if $\operatorname{supp}(g) \subset S$, where S is the set

$$S = \{s \in \mathbb{Z}^r \mid s + \Delta_D \subset H(A)\}.$$

The monomials u^s with $s \in S$ form a basis for the k-module of polynomials g with $\mathrm{supp}(g) \subset S$ and thus the functions $u^s F_D$ with $s \in S$ form a basis of $\ker(p)$. □

Proof of (3.2.10). From the description of the kernel in theorem (3.2.8) it follows that these kernels satisfy point 3 of theorem (2.5.5). Hence $\mathrm{cok}(p)$ is a free k-module of finite rank and we can apply corollary (2.1.19) to obtain the rest of the result. □

Proof of (3.2.11). We assume that A is given by enumerating its points. Since we give the output as a function $A \to \mathbb{Z}$, it is unreasonable to expect a runtime that does not involve $\#A$.

The algorithm works analogously to that from theorem (1.3.3). First, compute a solution to the reconstruction problem over $\mathbb{Q}$ using linear algebra. Since all the spaces involved have size polynomial in $\#A \cdot \#D$, this can be done in the stated runtime. If no solution over $\mathbb{Q}$ is found, output an error (any solution over $\mathbb{Z}$ would also be a solution over $\mathbb{Q}$).

Let p_D be the grid reconstruction system over $\mathbb{Q}$ associated to $(\mathbb{Z}^r, D)$. Let f_D be the generator of $\ker(p_D)$ described in theorem (3.1.2). We have to be a little careful here, as it may not be possible to write down p_D in acceptable time, unless A is large enough.

Let $f_1 = u^{d_1} - 1$ in $\mathbb{Q}[\mathbb{Z}^r]$ determine $A_1 \subset A$ such that $\mathrm{supp}(u^x f_1) \subset A$ if and only if $x \in A_1$. Note that $0 \in P(f_1)$ and f_1 divides F_D, so any $x \in A$ such that $\mathrm{supp}(u^x F_D) \subset A$, is in A_1.

For $i \in \{2, \ldots, n\}$, let $f_i = f_{i-1} \cdot (u^{d_i} - 1)$ and determine $A_i \subset A_{i+1}$ such that $\mathrm{supp}(u^x f_i) \subset A$ if and only if $x \in A_i$. As with f_1, we have $0 \in P(f_i)$ and $f_i | F_D$ so that every $x \in A$ such that $\mathrm{supp}(u^x f_D) \subset A$ is in A_i.

Suppose the computation got interrupted as A_i is empty. Then there is no $x \in \mathbb{Z}^r$ such that $x + P(f_i) \subset H(A)$. Since f_D is a multiple of f_i, it follows that there is no $x \in \mathbb{Z}^r$ such that $x + P(f_D) \subset H(A)$, and so, by theorem (3.2.8) we have $\ker(p) = 0$. It follows that a solution to the reconstruction problem over $\mathbb{Q}$, and over $\mathbb{Z}$, is unique. Moreover, since $\mathrm{cok}(p)$ is a free $\mathbb{Z}$-module, there is a solution over $\mathbb{Z}$ if and only if there is a solution over $\mathbb{Q}$. Hence the unique solution over $\mathbb{Q}$, which we have already determined, is also the unique solution over $\mathbb{Z}$. Output it and stop the algorithm.

Suppose the computation did not get interrupted. Since all the A_i are non-empty subsets of A, all the f_i have a support of which at least one translate is contained in A, hence they have no more than $\#A$ non-zero coefficients. It follows that we can compute the f_i and A_i in polynomial time

in the length of the original input. Note that $f_n = f_D$ and A_n is the set S from theorem (3.2.8).

The rest of the algorithm is identical to that in the proof of theorem (1.3.3) and we do not reproduce it here. □

We now turn our attention to the decomposition result for the cokernel as described in theorem (3.2.12). We derive a precise version of this result in corollary (3.4.14). But first, we state the technical result that this corollary will follow from, and spend some time proving it.

Definition 3.4.2. Let D be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let Δ_D be as in (3.4.1). Then a finite convex grid set $A \subset \mathbb{Z}^r$ is called *D-rounded* if $H(A)$ is Δ_D-rounded in the sense of definition (3.3.16). The *D-rounded part* of a convex grid set B is the largest D-rounded subset of B.

Theorem 3.4.3. *Let D be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let A be a finite convex grid set in $\mathbb{Z}^r$ that is D-rounded and suppose that $H(A)$ has non-empty interior. Let k be a commutative ring and let p_A and p_D be the grid reconstruction systems over k corresponding to (A, D) and $(\mathbb{Z}^r, D)$ respectively. Then the natural map $\operatorname{cok}(p_A) \to \operatorname{cok}(p_D)$ is injective.*

The idea of the proof is as follows. We construct finite convex subsets A_i of $\mathbb{Z}^r$ such that

$$A = A_0 \subset A_1 \subset A_2 \subset \dots$$

and

$$\bigcup_{i=0}^{\infty} A_i = \mathbb{Z}^r.$$

We then show that $\operatorname{cok}(p_{A_{i-1}}) \to \operatorname{cok}(p_{A_i})$ is injective for all i and conclude from this that $\operatorname{cok}(p_A) \to \operatorname{cok}(p_D)$ is injective.

Definition 3.4.4. Let $x \in \mathbb{R}^r$ be a point and $\lambda \in \mathbb{R}$ a non-zero scalar. Then the *point scaling map* with center x and scaling factor λ is the map

$$\begin{array}{rccc} s_{x,\lambda} : \mathbb{R}^r & \longrightarrow & \mathbb{R}^r \\ v & \longmapsto & x + \lambda(v - x). \end{array}$$

Lemma 3.4.5. *Let $X \subset \mathbb{R}^r$ be convex and have a non-empty interior. Let x*

be in X°, the interior of X. For $\lambda \in \mathbb{R}_{>1}$, put $X_\lambda = s_{x,\lambda}(X)$. Then one has

$$\bigcup_{\lambda>1} X_\lambda = \mathbb{R}^r$$

and for every $\lambda > 1$ the interior of X_λ is given by

$$X_\lambda^\circ = \bigcup_{1<\mu<\lambda} X_\mu.$$

Proof. As x is in the interior of X, there is some positive ϵ such that $B_\epsilon(x)$, the open ball of radius ϵ around x, is contained in X. It follows that $B_{\lambda\epsilon}(x)$ is contained in X_λ for all $\lambda \in \mathbb{R}_{>1}$. We conclude that

$$\bigcup_{\lambda>1} X_\lambda = \mathbb{R}^r$$

as any point $y \in \mathbb{R}^r$ is in $B_{\lambda\epsilon}(x)$ for some sufficiently large λ.

Suppose that $y \in X_\lambda^\circ$. Then for some $\delta > 0$, $B_\delta(y)$ is contained in X_λ. Let $y_1 \in X$ be the point such that $y = s_{x,\lambda}(y_1)$. Pick $\mu > 1$ such that $\mu < \lambda$ and $\mu > \lambda\frac{||x-y||}{||x-y||+\delta}$. This is always possible as $\lambda > 1$ and the fraction involving $||x-y||$ is always between 0 and 1. We will show that $y \in X_\mu$.

Let $y' = s_{x,\mu}(y_1)$ and $\delta' = \frac{\mu}{\lambda}\delta$. Note that $B_{\delta'}(y')$ is contained in X_μ, as $B_\delta(y)$ is contained in X_λ. Now we compute

$$||y-y'|| = ||x-y|| - ||x-y'|| = (1-\frac{\mu}{\lambda})||x-y||.$$

Rewriting $\mu > \lambda\frac{||x-y||}{||x-y||+\delta}$ one obtains $(\lambda-\mu)||x-y|| < \mu\delta$ and so

$$(1-\frac{\mu}{\lambda})||x-y|| < \frac{\mu}{\lambda}\delta = \delta'.$$

It follows that $||y-y'|| < \delta'$, i.e. $y \in B_{\delta'}(y')$. Hence y is in X_μ as required.

Conversely, suppose that $y \in X_\mu$ for some $\mu < \lambda$. We want to show that y is in the interior of X_λ. Let $y_1 \in X$ be the point such that $y = s_{x,\mu}(y_1)$ and let $y' = s_{x,\lambda}(y_1)$. Let $\delta = \lambda^{-1}(\lambda-\mu)$. Note that $0 < \delta < 1$. It follows that $s_{y',\delta}$ moves a point $z \in B_\epsilon(x)$ to a point on the line segment between z and y'. As X_λ is convex, this point is in X_λ. Note that

$$s_{y',\delta}(B_\epsilon(x)) = B_{\delta\epsilon}(s_{y',\delta}(x)) = B_{\delta\epsilon}(y)$$

and there is an open ball around y contained in X_λ, i.e, y is in the interior of X_λ. □

Lemma 3.4.6. *Let $X \subset \mathbb{R}^r$ be convex, closed and bounded and suppose it has a non-empty interior. Let $x \in X^\circ$ and for $\lambda \in \mathbb{R}_{>1}$ let $X_\lambda = s_{x,\lambda}(X)$ as in lemma* (3.4.5)*. Then the set*

$$\Lambda = \{\lambda \in \mathbb{R}_{>1} \mid \forall \mu < \lambda : X_\mu \cap \mathbb{Z}^r \neq X_\lambda \cap \mathbb{Z}^r\}$$

can be enumerated in order, i.e.

$$\Lambda = \{\lambda_1, \lambda_2, \lambda_3, \ldots\}$$

with $\lambda_i < \lambda_j$ whenever $i < j$.

Proof. It follows from lemma (3.4.5) that the boundaries ∂X_λ with $\lambda > 1$ are all disjoint. The set Λ from the theorem can also be described as

$$\Lambda = \{\lambda \in \mathbb{R}_{>1} \mid \mathbb{Z}^r \cap \partial X_\lambda \text{ is non-empty}\}.$$

It follows that this set is countable. Moreover, for every $m > 1$, the sets X_λ with $1 < \lambda < m$ are all contained in some big closed box B around the origin, as X is bounded. Hence, for $1 < \lambda < m$ we have

$$\mathbb{Z}^r \cap \partial X_\lambda = (B \cap \mathbb{Z}^r) \cap \partial X_\lambda$$

and $(B \cap \mathbb{Z}^r)$ is finite, so there are only finitely many $\lambda \in \Lambda$ such that $\lambda < m$. It follows that Λ can be enumerated in order. □

The idea is now to take $X = H(A)$, fix some $x \in X^\circ$ and put $A_i = X_{\lambda_i} \cap \mathbb{Z}^r$. But first we will show how to use the freedom we have in choosing $x \in X^\circ$ to have a lot of control over $A_i \setminus A_{i-1}$.

Lemma 3.4.7. *Let V and V' be parallel but distinct hyperplanes in $\mathbb{R}^r$ and let W, W' be another pair of parallel, distinct hyperplanes in $\mathbb{R}^r$, such that $(V, V') \neq (W, W')$. Then the set*

$$\{p \in \mathbb{R}^r \mid \exists \lambda \in \mathbb{R} : s_{p,\lambda} V = V' \text{ and } s_{p,\lambda} W = W'\}$$

is contained in a hyperplane in $\mathbb{R}^r$.

Proof. There are non-zero vectors $v, w \in \mathbb{R}^r$ and scalars α, α', β and β' such that

$$\begin{aligned} V &= \{x \mid \langle x, v\rangle = \alpha\}, \\ V' &= \{x \mid \langle x, v\rangle = \alpha'\}, \\ W &= \{x \mid \langle x, w\rangle = \beta\}, \\ W' &= \{x \mid \langle x, w\rangle = \beta'\}. \end{aligned}$$

Hence $s_{p,\lambda}V = V'$ if and only if

$$\langle (p + \lambda(x - p)), v\rangle = \alpha'$$

whenever $\langle x, v\rangle = \alpha$. This is equivalent to

$$\alpha' = \lambda\alpha + (1 - \lambda)\langle p, v\rangle$$

and to

$$(\alpha' - \langle p, v\rangle) = \lambda(\alpha - \langle p, v\rangle).$$

Similarly, $s_{p,\lambda}W = W'$ if and only if

$$(\beta' - \langle p, w\rangle) = \lambda(\beta - \langle p, w\rangle).$$

Note that $\alpha - \langle p, v\rangle = 0$ if and only if $p \in V$. Since V and V' are distinct, for such p there never is a λ such that $s_{p,\lambda}V = V'$. Likewise, if one of the other factors above is 0, p is contained in one of the other hyperplanes V', W or W', and for such p there never is a λ. From now on we restrict to p outside these hyperplanes and so all these factors are non-zero. In that case, there is a λ satisfying these equations if and only if

$$(\alpha' - \langle p, v\rangle)(\beta - \langle p, w\rangle) = (\alpha - \langle p, v\rangle)(\beta' - \langle p, w\rangle).$$

After expanding and simplifying this equation is equivalent to

$$\alpha'\beta - \beta\langle p, v\rangle - \alpha'\langle p, w\rangle = \alpha\beta' - \beta'\langle p, v\rangle - \alpha\langle p, w\rangle.$$

This is a linear equation in the coordinates of p. The solutions to such an equation are contained in a hyperplane as required, unless all the terms in the equation cancel. We will show that this cannot happen under the conditions from the lemma. Let $i = 1, \ldots, r$, then the coefficient of p_i, the i-th coordinate of p in the equation above is zero if and only if

$$-\beta v_i - \alpha' w_i = -\beta' v_i - \alpha w_i,$$

which can be rewritten as $(\beta' - \beta)v_i = (\alpha' - \alpha)w_i$. It follows that p vanishes from the equation altogether if and only if

$$(\beta' - \beta)v = (\alpha' - \alpha)w.$$

If this is the case, we will either have no solutions, when $\alpha'\beta \neq \alpha\beta'$, or every p is a solution, when $\alpha'\beta = \alpha\beta'$. However, in the latter case we have for $x \in \mathbb{R}^r$ that

$$\langle x, v\rangle = \frac{\beta' - \beta}{\alpha' - \alpha}\langle x, w\rangle$$

and as

$$\alpha\frac{\beta'-\beta}{\alpha'-\alpha}=\frac{\alpha\beta'-\alpha\beta}{\alpha'-\alpha}=\frac{\alpha'\beta-\alpha\beta}{\alpha'-\alpha}=\beta$$

we have $V=W$ and by a similar argument, $V'=W'$. Since, by assumption $(V,V')\neq(W,W')$, this cannot happen. □

Lemma 3.4.8. *Let B be a non-empty open ball in $\mathbb{R}^r$. Then B is not contained in a countable union of hyperplanes in $\mathbb{R}^r$.*

Proof. We prove this by induction on r. For $r=1$, B is an open interval in $\mathbb{R}$ and a hyperplane is a single point in $\mathbb{R}$. Hence the claim is that an interval is not countable, which is indeed true.

Let $r>1$. For any point in B, uncountably many hyperplanes in $\mathbb{R}^r$ pass though that point. Hence there is a hyperplane V that passes through a point in B and that is distinct from the countably many given hyperplanes. If we intersect B and the given hyperplanes with V, we obtain a similar situation in $\mathbb{R}^{r-1}$, so by the induction hypothesis, there is a point in V that lies inside $B\cap V$ but doesn't lie in the intersection of any of the given hyperplanes with V. Hence this point is in B and not in any of the given hyperplanes. □

Theorem 3.4.9. *Let $X\subset\mathbb{Z}^r$ be a polytope and suppose that X has a non-empty interior. Then there is an $x\in X^\circ$ so that if one defines $X_\lambda=s_{x,\lambda}(X)$ as in lemma* (3.4.5) *and let $\lambda_1,\lambda_2,\dots$ be as in lemma* (3.4.6)*, there is for every $i\geq 1$ a $v_i\in\mathbb{R}^r\setminus\{0\}$ such that $(\partial X_{\lambda_i})\cap\mathbb{Z}^r$ is $F_{v_i}(X_{\lambda_i})\cap\mathbb{Z}^r$.*

Proof. By corollary (3.3.11) we can write

$$\partial X=F_{w_1}(X)\cup\cdots\cup F_{w_n}(X).$$

Without loss of generality, the w_i can be chosen such that the faces $F_{w_i}(X)$ are all distinct. Let H_i be the hyperplane perpendicular to w_i so that $H_i\cap X=F_{w_i}(X)$. It follows that these hyperplanes are all distinct. For every $x\in X^\circ$ and every $\lambda>1$ we have

$$\begin{aligned}\partial(s_{x,\lambda}(X)) &= s_{x,\lambda}(\partial X)=s_{x,\lambda}(F_{w_1}(X)\cup\cdots\cup F_{w_n}(X))\\ &= F_{w_1}(s_{x,\lambda}(X))\cup\cdots\cup F_{w_n}(s_{x,\lambda}(X))\\ &= (s_{x,\lambda}(X)\cap s_{x,\lambda}(H_1))\cup\cdots\cup(s_{x,\lambda}(X)\cap s_{x,\lambda}(H_n)).\end{aligned}$$

For $i=1,\dots,n$, the set

$$\left\{H\;\middle|\;\begin{array}{l}H\text{ hyperplane parallel to }H_i\\ H\cap\mathbb{Z}^r\text{ is non-empty}\\ H\neq H_i\end{array}\right\}$$

is countable, so we can label the elements as $H_{i,j}$ with $j \in \mathbb{Z}_{\geq 0}$. Hence the set of pairs (H, H') where $H = H_{i,j}$ and $H' = H_{i',j'}$ for some i, j, i' and j' satisfying $(i,j) \neq (i',j')$ is also countable. For each such a pair, we have $H \neq H_i$ and $H' \neq H_{i'}$ and $(H_i, H) \neq (H_{i'}, H')$ as $(i,j) \neq (i',j')$. By lemma (3.4.7) it follows there is a hyperplane $B_{H,H'}$ containing the set

$$\{x \mid \exists \lambda : s_{x,\lambda}(H_i) = H \text{ and } s_{x,\lambda}(H_{i'}) = H'\}.$$

As X° is non-empty, it contains some open ball. By lemma (3.4.8) there is a point x in this ball such that x lies outside all the hyperplanes $B_{H,H'}$ defined above. We will show this point x satisfies the requirements from the theorem.

For $\lambda > 1$, let $X_\lambda = s_{x,\lambda}(X)$ as in lemma (3.4.5). Let $\lambda_1, \lambda_2, \ldots$ be as in lemma (3.4.6). For every $i > 0$ the set $\mathbb{Z}^r \cap \partial X_{\lambda_i}$ is non-empty and so at least one of the sets $\mathbb{Z}^r \cap (X_{\lambda_i} \cap s_{x,\lambda_i}(H_j))$ with $j = 1, \ldots, n$ is non-empty. If more than one of them is non-empty, we have hyperplanes $H = s_{x,\lambda_i}(H_j)$ and $H' = s_{x,\lambda_i}(H_{j'})$, which are perpendicular to H_j and $H_{j'}$ respectively and both of which contain points in $\mathbb{Z}^r$. But we chose x outside the hyperplane $B_{H,H'}$, so this cannot happen. It follows that there is a j such that

$$\mathbb{Z}^r \cap \partial X_{\lambda_i} = \mathbb{Z}^r \cap (X_{\lambda_i} \cap s_{x,\lambda_i}(H_j)) = F_{w_j}(X_{\lambda_i}).$$

We put $v_i = w_j$. By lemma (3.4.5), the interior of X_{λ_i} is the union of the X_μ with $1 < \mu < \lambda_i$, so $A_i \setminus A_{i-1}$ consists precisely of the integral points on the boundary of X_{λ_i}. Hence we have $A_i \setminus A_{i-1} = F_{v_i}(X_{\lambda_i})$ as required. □

Let A and D be as in theorem (3.4.3) and put $X = H(A)$. Fix $x \in X^\circ$ as in theorem (3.4.9) above and let $A_0 = A$ and $A_i = X_{\lambda_i} \cap \mathbb{Z}^r$ for $i \geq 1$. Let $v_i \in \mathbb{Z}^r$ be as in the theorem the direction such that $A_i \setminus A_{i-1} = F_{v_i}(X_{\lambda_i}) \cap \mathbb{Z}^r$. Let p_{A_i} be the grid reconstruction system corresponding to (A_i, D) and let $p_{A_i/A_{i-1}}$ be the reconstruction system defined in lemma (2.4.4) for $A_{i-1} \subset A_i$.

Lemma 3.4.10. *Let $i \geq 1$. Then $p_{A_i/A_{i-1}}$ is equal to the grid reconstruction system corresponding to $(A_i \setminus A_{i-1}, D')$, where D' is the sub-sequence of D consisting of those d_j perpendicular to v_i.*

Proof. What we need to show is that the lines in directions from D that go through a point in $A_i \setminus A_{i-1}$ and don't go through any point in A_{i-1} are precisely those lines in directions from D' that go through a point in $A_i \setminus A_{i-1}$.

We know that $A_i \setminus A_{i-1}$ is the set of integral points in $F_{v_i}(X_{\lambda_i})$. Let H be the hyperplane perpendicular to v_i containing this face. By definition

$F_{v_i}(X_{\lambda_i})$ is the intersection of X_{λ_i} with H. Note that

$$A_{i-1} \cap H \subset A_i \cap H \subset X_{\lambda_i} \cap H = F_{v_i}(X_{\lambda_i}) = A_i \setminus A_{i-1}$$

so it must be empty. Any line in a direction from D' going through a point in $A_i \setminus A_{i-1}$ is contained in H and therefore by the above it does not go through any point in A_{i-1}. This proves one inclusion.

Conversely, suppose we have a point p in $A_i \setminus A_{i-1}$ and a line ℓ in direction d_j that is not perpendicular to v_i. What we will show is that ℓ contains another point of A_i. This point cannot be in $H \cap A_i = A_i \setminus A_{i-1}$, so it must be in A_{i-1}, proving the other inclusion.

Let $L = P(u^{d_j} - 1)$. Note L is a line segment parallel to ℓ with endpoints a and b satisfying $\{a\} = F_{v_i}(L)$ and $\{b\} = F_{-v_i}(L)$. Note that Δ_D is L-rounded and X_{λ_i} is Δ_D-rounded, so X_{λ_i} is L-rounded and we can write $X_{\lambda_i} = L + Y$ for some convex set Y. Hence, by lemma (3.3.17) we have

$$F_{v_i}(X_{\lambda_i}) = F_{v_i}(Y) + F_{v_i}(L) = F_{v_i}(Y) + a.$$

In particular, p can be written as $q+a$ for some $q \in Y$. Note that $L+q \subset X_{\lambda_i}$, so the point $q + b$ is in X_{λ_i}. Moreover, it is integral, so $q + b$ is in A_i and it also lies on ℓ, i.e. it is a second point in $A_i \cap \ell$. As noted before, this completes the proof. □

Lemma 3.4.11. *Let D be a 2-regular sequence of directions in $\mathbb{Z}^r$ and let $v \in \mathbb{R}^r$ be a non-zero direction. Let D' be the sub-sequence of D consisting of those d_j perpendicular to v. Let Δ_D and $\Delta_{D'}$ be as in lemma* (3.4.1)*. Let $B = F_v(\Delta_D) \cap \mathbb{Z}^r$ and let G be the restriction of F_D to B. Then there is a $t \in \mathbb{Z}^r$ such that $G = \pm u^t F_{D'}$. In particular, $F_v(\Delta_D) = t + \Delta_{D'}$.*

Proof. Let D'' be the complement of D' in D. Let F_D, $F_{D'}$ and $F_{D''}$ be as in theorem (3.1.2). Let Δ_D, $\Delta_{D'}$ and $\Delta_{D''}$ be their polytopes as in lemma (3.4.1). Note that $F_D = F_{D'}F_{D''}$ and these elements have strong corners, so

$$\Delta_D = \Delta_{D'} + \Delta_{D''}.$$

We will prove below that $F_v(\Delta_{D'}) = \Delta_{D'}$ and that $F_v(\Delta_{D''})$ is a single point, but we first show how to complete the proof from this observation. By lemma (3.3.17) this implies that $F_v(\Delta_D) = t + \Delta_{D'}$, where $t \in \mathbb{Z}^r$ is the unique point in $F_v(\Delta_{D''})$. In lemma (3.4.1) we showed that the coefficient of $F_{D''}$ at t is ± 1. Every contribution to coefficients of F_D in $F_v(\Delta_D)$ is a product of this coefficient of $F_{D''}$ at t times a coefficient from $F_{D'}$. It follows that $G = \pm u^t F_{D'}$ as required.

Let

$$H_v = \{x \in \mathbb{R}^r \mid \langle x, v\rangle = 0\}$$

be the hyperspace orthogonal to v.

Let $d \in D$. If $\langle d, v\rangle = 0$, then $H(u^d - 1)$ is contained in H_v. It follows that $\Delta_{D'}$ is contained in this hyperspace, and so $F_v(\Delta_{D'}) = \Delta_{D'}$ as all points in it have the same inner product with v.

On the other hand if $d \in D$ has $\langle d, v\rangle \neq 0$, then $H(u^d - 1)$ has a unique vertex with maximal inner product with v. Any sum of such polytopes will again have a unique vertex with maximal inner product with v. Thus $F_v(\Delta_{D''})$ consists of a single point. □

Theorem 3.4.12. *The natural map* $\ker(p_{A_i}) \to \ker(p_{A_i/A_{i-1}})$ *is surjective.*

Proof. By definition of $p_{A_i/A_{i-1}}$ there is a short exact sequence

$$0 \to p_{A_{i-1}} \to p_{A_i} \to p_{A_i/A_{i-1}} \to 0.$$

By lemma (3.4.10) the $p_{A_i/A_{i-1}}$ is equal to the grid reconstruction system associated to $(A_i \setminus A_{i-1}, D')$, where D' is the sub-sequence of D consisting of those d_j perpendicular to v_i. Hence, by theorem (3.2.8) there is a finite set S such that $\ker(p_{A_i/A_{i-1}})$ is generated by elements of the form $u^s F_{D'}$ with $s \in S$.

To prove the theorem it therefore suffices to show that each of these $u^s F_{D'}$ arise as the restriction of an element of $\ker(p_{A_i})$ to $A_i \setminus A_{i-1}$. By lemma (3.4.11) there is a $t \in \mathbb{Z}^r$ such that $u^t F_{D'} = G$, where G is the restriction of F_D to $F_{v_i}(\Delta_D) \cap \mathbb{Z}^r$. It follows that $u^s F_{D'} = u^{s-t} G$. In particular, we have

$$(s - t) + F_{v_i}(\Delta_D) = P(u^s \Delta_{D'}) \subset H(A_i \setminus A_{i-1}) \subset F_{v_i}(X_{\lambda_i}).$$

Recall that A is D-rounded and so there is some convex set Y such that $X_{\lambda_i} = Y + \Delta_D$. By lemma (3.3.17) we then have

$$F_{v_i}(X_{\lambda_i}) = F_{v_i}(Y) + F_{v_i}(\Delta_D).$$

It follows that $(s-t) \in F_{v_i}(Y)$ and in particular $(s-t) \in Y$. Hence $(s-t)+\Delta_D$ is contained in X_{λ_i}. It follows that the support of $u^{s-t} F_D$ is contained in A_i, hence $u^{s-t} F_D$ is in $\ker(p_{A_i})$ by theorem (3.2.8). By construction, the restriction of $u^{s-t} F_D$ to $A_i \setminus A_{i-1}$ is equal to $u^s F_{D'}$. □

Corollary 3.4.13. *The natural map* $\operatorname{cok}(p_{A_{i-1}}) \to \operatorname{cok}(p_{A_i})$ *is injective.*

Proof. Apply corollary (2.4.13) to the short exact sequence

$$0 \to p_{A_{i-1}} \to p_{A_i} \to p_{A_i/A_{i-1}} \to 0$$

to derive this from the theorem above. □

Proof of (3.4.3). Fixing notation, write $p_{A_i} : T_{A_i} \to P_{A_i}$. Similarly, write $p_A : T_A \to P_A$ and $p_D : T_D \to P_D$. We view all the T_{A_i} as subspaces of $T_D = k[\mathbb{Z}^r]$ and all the P_{A_i} as subspaces of $P_D = k[\mathbb{Z}^r/d_1] \oplus \cdots \oplus k[\mathbb{Z}^r/d_n]$.

Let $x \in P_A$. We want to prove that if $x \in \mathrm{im}(p_D)$ (i.e. x maps to 0 in $\mathrm{cok}(p_D)$), then $x \in \mathrm{im}(P_A)$ (i.e. x represents the 0 element in $\mathrm{cok}(p_A)$). Note that x is in p_{A_i} for all i, so it gives rise to an element of each $\mathrm{cok}(p_{A_i})$. By corollary (3.4.13), if x maps to 0 in some $\mathrm{cok}(p_{A_i})$, it must map to 0 also in all $\mathrm{cok}(p_{A_j})$ with $j < i$, in particular it must map to 0 in $\mathrm{cok}(p_{A_0}) = \mathrm{cok}(p_A)$.

Now suppose there is a $y \in T_D$ such that $p_D(y) = x$. Then $\mathrm{supp}(y)$ is contained in A_i for some sufficiently large i. Hence $x \in \mathrm{im}(p_{A_i})$, that is, x maps to 0 in $\mathrm{cok}(p_{A_i})$. As we just observed, this completes the proof. □

Corollary 3.4.14. *Let D be a 2-regular sequence of directions in $\mathbb{Z}^r$ and let A be a convex grid set in $\mathbb{Z}^r$ and R its D-rounded part. Suppose R has non-empty interior. Let k be a commutative ring. Let p_A, p_R and p_D be the grid reconstruction systems over k corresponding respectively to (A, D), (R, D) and $(\mathbb{Z}^r, D)$. Let $p_{A/R}$ be the reconstruction system defined in* (2.4.4) *for the extension $R \subset A$. Then* $\mathrm{cok}(p_R) \to \mathrm{cok}(p_D)$ *is injective,* $\mathrm{cok}(p_{A/R})$ *is a free k-module and there is an isomorphism* $\mathrm{cok}(p_A) \cong \mathrm{cok}(p_R) \oplus \mathrm{cok}(p_{A/R})$.

Proof. By assumption, we can apply theorem (3.4.3) to the set R and obtain that the map $\mathrm{cok}(p_R) \to \mathrm{cok}(p_D)$ is injective. We can write this map as a composition

$$\mathrm{cok}(p_R) \to \mathrm{cok}(p_A) \to \mathrm{cok}(p_D)$$

and so the map $\mathrm{cok}(p_R) \to \mathrm{cok}(p_A)$ must be injective. The short exact sequence of reconstruction systems

$$0 \to p_R \to p_A \to p_{A/R} \to 0$$

gives rise to an exact sequence of cokernels

$$\mathrm{cok}(p_R) \to \mathrm{cok}(p_A) \to \mathrm{cok}(p_{A/R}) \to 0$$

by lemma (2.4.7). By the observation above, the first map is injective and so this sequence is again short exact.

It now suffices to show that $\operatorname{cok}(p_{A/R})$ is a free k-module, for a short exact sequence that ends in a free module is split and hence we obtain

$$\operatorname{cok}(p_A) \cong \operatorname{cok}(p_R) \oplus \operatorname{cok}(p_{A/R}).$$

By lemma (2.5.3) it suffices to show this in the case that $k = \mathbb{Z}$. As $\operatorname{cok}(p_{A/R,\mathbb{Z}})$ is a finitely generated $\mathbb{Z}$-module, it suffices to show it is torsion-free (see e.g. [17, Ch. 1, Thm. 8.4]). Let k be $\mathbb{Z}$ or $\mathbb{F}_p$ for some prime p. Then the rank of $\operatorname{cok}(p_{A/R,k})$ as a k-module is equal to

$$\operatorname{rk}(\operatorname{cok}(p_{A,k})) - \operatorname{rk}(\operatorname{cok}(p_{R,k}))$$

by the short exact sequence we derived before. From corollary (3.2.10) we know that the ranks of $\operatorname{cok}(p_{A,k})$ and $\operatorname{cok}(p_{R,k})$ do not depend on k. Hence the rank of $\operatorname{cok}(p_{A/R,k})$ does not depend on k. This implies that $\operatorname{cok}(p_{A/R,\mathbb{Z}})$ cannot have p-torsion for any prime p and hence it is torsion-free as required. □

3.5 The planar case

In this section, we focus on the case $r = 2$. Many of the important results from earlier in the chapter are somewhat nicer in this case. The key observation is the following.

Theorem 3.5.1. *Let $D = ((a_1, b_1), \ldots, (a_n, b_n))$ be a 2-regular sequence of directions in $\mathbb{Z}^2$. Let k be a commutative ring and let $p : T \to P$ be the grid reconstruction system associated to $(\mathbb{Z}^2, D)$. Then $\operatorname{cok}(p)$ is a free k-module of rank*

$$\sum_{1 \le i < j \le n} |a_i b_j - a_j b_i|.$$

Hence $\operatorname{Dep}(p)$ is free of the same rank and for $x \in P$ we have $x \in \operatorname{im}(p)$ if and only if $d(x) = 0$ for all d in a basis of $\operatorname{Dep}(p)$.

Proof. In theorem (3.1.6) we saw that

$$\operatorname{cok}(p) \cong \bigoplus_{1 \le i < j \le n} k[\mathbb{Z}^r]/(u^{d_i} - 1, u^{d_j} - 1),$$

where $d_i = (a_i, b_i)$ for all $i = 1, \ldots, n$.

Let $1 \le i < j \le n$. By lemma (3.1.9) there is an isomorphism

$$k[\mathbb{Z}^r]/(u^{d_i} - 1, u^{d_j} - 1) \cong k[\mathbb{Z}^r/\langle d_i, d_j \rangle].$$

The quotient group $\mathbb{Z}^r/\langle d_i, d_j\rangle$ is finite and consists of $|a_ib_j - a_jb_i|$ elements. The result follows. □

Example 3.5.2. Consider the full grid reconstruction system defined in example (1.4.3). Recall that the projection directions in this example are $(1,0)$, $(2,1)$, $(1,2)$ and $(0,1)$. The table below collects the quotient groups $Z^2/\langle d_i, d_j\rangle$ for all pairs of directions d_i, d_j.

d_i, d_j	$\mathbb{Z}^2/\langle d_i, d_j\rangle \cong$
$(1,0),(2,1)$	$\{0\}$
$(1,0),(1,2)$	$\mathbb{Z}/2\mathbb{Z}$
$(1,0),(0,1)$	$\{0\}$
$(2,1),(1,2)$	$\mathbb{Z}/3\mathbb{Z}$
$(2,1),(0,1)$	$\mathbb{Z}/2\mathbb{Z}$
$(1,2),(0,1)$	$\{0\}$

Adding the orders of these groups together we conclude that the cokernel (and hence the space of dependencies) has rank 10.

Theorem (3.5.1) shows that the module of dependencies for the full $\mathbb{Z}^2$ grid is always of finite rank. Using this, one would expect there is an efficient algorithm to decide the consistency problem. But there is a small complication that still needs to be overcome, namely, it is not at all clear how to make an algorithm to compute the value of a dependency on a vector of line sums, as there are infinitely many lines and they can each have a different non-zero weight. Fortunately, we can use theorem (3.1.5) to show these weights have additional structure, which allows us to compute those weights we need efficiently.

Remark 3.5.3. A primitive direction $d = (a, b) \in \mathbb{Z}^r$ gives rise to a surjective group homomorphism

$$\phi_d : \mathbb{Z}^2 \longrightarrow \mathbb{Z} \qquad (x, y) \mapsto ay - bx$$

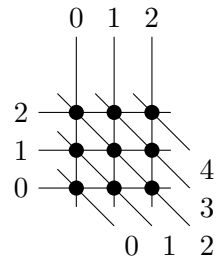
whose kernel is precisely the subgroup generated by d. Hence ϕ_d induces an isomorphism $\mathbb{Z}^2/d \cong \mathbb{Z}$. This gives an identification of the set of lines in the direction d with $\mathbb{Z}$.

Let D be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$, k a commutative ring and p the grid reconstruction system over k corresponding to $(\mathbb{Z}^2, D)$. We saw in lemma (2.1.17) that we can associate to any dependency $d \in \mathrm{Dep}(p)$ a sequence of coefficients $c_\ell \in k$ for $\ell \in \mathcal{L}$, where $\mathcal{L}$ is the set of all lines in the directions from D. As we have just seen, $\mathcal{L}$ is the disjoint union of n copies of $\mathbb{Z}$. Hence this sequence $(c_\ell)_{\ell\in\mathcal{L}}$ can be represented as n two-sided infinite sequences $c_1, \ldots, c_n : \mathbb{Z} \to k$.

Example 3.5.4. Consider the directions $(0,-1)$, $(1,0)$ and $(1,-1)$ in $\mathbb{Z}^2$. The maps ϕ_{d_i} are given in the following table.

i	d_i	ϕ_{d_i}
1	$(0,-1)$	$(x,y) \mapsto x$
2	$(1,0)$	$(x,y) \mapsto y$
3	$(1,-1)$	$(x,y) \mapsto x+y$

This corresponds to a numbering of the lines as in the following picture



Let k be a commutative ring and let p be the grid reconstruction system corresponding to $(\mathbb{Z}^2, (d_1, d_2, d_3))$ over k.

Consider the following sequences:

$$\begin{aligned} c_1 &: \mathbb{Z} \to k \qquad n \mapsto \quad n; \\ c_2 &: \mathbb{Z} \to k \qquad n \mapsto \quad n; \\ c_3 &: \mathbb{Z} \to k \qquad n \mapsto -n. \end{aligned}$$

They are the coefficient sequences corresponding to the k-linear map

$$d : k[\mathbb{Z}^2/d_1] \oplus k[\mathbb{Z}^2/d_2] \oplus k[\mathbb{Z}^2/d_3] \to k$$

that counts the j-th line in direction d_i with weight $c_i(j)$.

One sees easily that d satisfies $d \circ p = 0$. Suppose we have a function $f : \mathbb{Z}^2 \to k$ that is 0 almost everywhere. Let $(x, y) \in \mathbb{Z}^2$. The coefficient of f at (x, y) contributes to the x-th line in direction d_1, the y-th line in direction d_2 and the $(x + y)$-th line in direction d_3. Hence the total contribution of $f(x, y)$ to $d(p(f))$ is

$$(c_1(x) + c_2(y) + c_3(x + y))f(x, y) = (x + y - (x + y))f(x, y) = 0.$$

As f and (x, y) were arbitrary, we conclude that $d \circ p = 0$. Hence d descends to a dependency $\mathrm{cok}(p) \to k$.

Theorem 3.5.5. *Let D be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. Let $n = \#D$ and assume $n \geq 1$. Then there are positive integers $r_1, \ldots, r_n$ and*

$$a_{1,1}, \ldots, a_{1,r_1}, a_{2,1}, \ldots, a_{n,r_n} \in \mathbb{Z}$$

with $a_{i,1} = \pm 1$ and $a_{i,r_i} = \pm 1$ for $i = 1, \ldots, n$ such that the following holds. Let k be a commutative ring and let p be the grid reconstruction system over k corresponding to (A, D). Let $d \in \mathrm{Dep}(p)$ and let $c_1, \ldots, c_n$ be the corresponding coefficient sequences as in remark (3.5.3)*. Then we have*

$$\forall m \in \mathbb{Z} : \sum_{j=1}^{r_i} a_{i,j} c_i(m + j) = 0$$

for $i = 1, \ldots, n$.

Proof. If k is the zero ring, any linear combination of ring elements is 0 and so any linear recurrence relation will hold over this ring. From now on we restrict ourselves to the case that k is not the zero ring.

Let $F_1, \ldots, F_n$ be as in theorem (3.1.5). Note that the image of F_i in $k[\mathbb{Z}^2/d_j]$ is 0 whenever $j \neq i$. The map ϕ_{d_i} from remark (3.5.3) gives rise to an isomorphism of rings $k[\mathbb{Z}^2/d_i] \cong k[\mathbb{Z}] = k[w, w^{-1}]$. The image of $p_{d_i}(F_i)$ under this isomorphism is a Laurent polynomial

$$a_{i,1} w^{s+1} + \cdots + a_{i,r_i} w^{s+r_i}$$

where r_i and s are such that $a_{i,1}$ and a_{i,r_i} are non-zero.

Let $i \in \{1, \ldots, n\}$ and let $m \in \mathbb{Z}$. As $\phi_{d_i} : \mathbb{Z}^2 \to \mathbb{Z}$ is onto, there is a $t \in \mathbb{Z}^2$ such that $\phi_{d_i}(t) = m - s$. Now consider the image of $u^t F_i \in k[\mathbb{Z}^2]$ under the

line sum map p. In $k[\mathbb{Z}^2/d_j]$ with $j \neq 0$ it maps 0 and in $k[\mathbb{Z}^2/d_i] \cong k[w, w^{-1}]$ it maps to

$$w^{m-s} \sum_{j=1}^{r_i} a_{i,j} w^{s+j} = \sum_{j=1}^{r_i} a_{i,j} w^{m+j}.$$

The dependency d must send $p(u^t F_i)$ to 0. The line corresponding to the monomial w^j above has weight $c_i(j)$ by definition of the coefficient sequences. Hence for all $i \in \{1, \ldots, n\}$ and all $m \in \mathbb{Z}$ we have

$$\sum_{j=1}^{r_i} a_{i,j} c_i(m+j) = 0.$$

It remains to be shown that the coefficients $a_{i,j}$ are all integers and that the leading and trailing coefficient of each $p_{d_i}(F_i)$ are ± 1. Note that F_i is a product of factors $u^{d_j} - 1$. One computes that such a factor maps to $w^m - 1$ where $m = \det(d_i, d_j)$ is non-zero. So each of these factors has integer coefficients and the leading and trailing coefficients are ± 1. Moreover, one easily sees that when two polynomials have this property, so does their product. □

Example 3.5.6. We continue with the setup from example (3.5.4) above. Using the isomorphism $k[\mathbb{Z}^2] \cong k[u, u^{-1}, v, v^{-1}]$ the F_i are given as Laurent polynomials in the following table.

i	F_i
1	$(u-1)(uv^{-1}-1)$
2	$(v^{-1}-1)(uv^{-1}-1)$
3	$(v^{-1}-1)(u-1)$

Using the isomorphisms $k[\mathbb{Z}^2/d_i] \cong k[w, w^{-1}]$ the line sum maps p_{d_i} are given by the following ring homomorphisms.

i	$k[u, u^{-1}, v, v^{-1}] \to k[w, w^{-1}]$
1	$u \mapsto w, v \mapsto 1$
2	$u \mapsto 1, v \mapsto w$
3	$u \mapsto w, v \mapsto w$

It follows that the elements $p_{d_i}(F_i)$ are as in the table below.

i	$p_{d_i}(F_i)$
1	$(w-1)(w-1) = w^2 - 2w + 1$
2	$(w^{-1}-1)(w^{-1}-1) = 1 - 2w^{-1} + w^{-2}$
3	$(w^{-1}-1)(w-1) = w + 2 + w^{-1}$

Hence, according to theorem (3.5.5), the sequences c_1, c_2 and c_3 corresponding to a dependency, such as for example those given in example (3.5.4) should each satisfy the linear recurrence

$$c_i(m) - 2c_i(m-1) + c_i(m-2) = 0$$

for all $m \in \mathbb{Z}$. Rewriting this equation as

$$c_i(m) - c_i(m-1) = c_i(m-1) - c_i(m-2)$$

we conclude that each c_i must be given by a linear function $c_i(n) = an + b$ for some $a, b \in k$.

Corollary 3.5.7. *Let D be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. Let $p : T \to P$ be the grid reconstruction system over $\mathbb{Z}$ associated to $(\mathbb{Z}^2, D)$. Then, for every fixed D, there is an algorithm that given an element $x \in P$ decides if $x \in \mathrm{im}(p)$ whose runtime is polynomial in the length of the input.*

Proof. We assume the element $x \in P$ is given to us as a finite set of lines $\mathcal{L}$ and a coefficient x_ℓ for every line $\ell \in L$ and that for all $\ell \notin \mathcal{L}$ we have $x_\ell = 0$.

Since our algorithm may depend on D, we assume we know the following:

- the numbers $r_1, \ldots, r_n$ and $a_{i,j}$ associated to D in theorem (3.5.5);
- a basis for $\mathrm{Dep}(p)$ (we will compute such bases in chapter 4);
- for every d in this basis, the coefficients $c_i(j)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, r_j - 1$.

What we need to show is, given the data we've just described, we can compute $d(x)$ for every d in the basis in time polynomial in the input length.

Using the linear recurrence relation for each c_i, we can compute $c_i(n)$ for every $n \in \mathbb{Z}$ in time polynomial in $\log(n)$. This can for example be done

using a closed formula for the solution, or using a matrix representation of the recurrence. In general, one expects the size of these numbers to grow as λ^n, where λ is the largest absolute value of a complex root of the characteristic polynomial of the recurrence, and so the size of the output would grow exponentially in $\log(n)$.

However, we computed in the proof of theorem (3.5.5) that these polynomials are products of polynomials of the form $w^m - 1$, hence all their roots have absolute value 1 and so the size of the n-th coefficient is at most a polynomial in $\log(n)$.

Hence we can find for each line the weight of that line in time polynomial in the input length. What remains is to multiply these weights by the line sums and add them all together, which can also clearly be done in polynomial time. □

In order to make this result computationally useful, we need a way to compute a basis (or generating set) for the dependencies. This is what we will do in chapter 4. In fact, we will describe the dependencies there in such a way we won't even need the linear recurrences to compute coefficients.

The finiteness of the dependencies also makes it plausible that for a 'sufficiently large' $A \subset \mathbb{Z}^2$ we should be able to distinguish between all the 'global' dependencies by only their restriction to A. The following theorem makes this exact.

Theorem 3.5.8. *Let D be a 2-regular sequence of directions in $\mathbb{Z}^2$ and suppose $A \subset \mathbb{Z}^2$ is such that $H(A)$ contains $t + \Delta_D$ for some $t \in \mathbb{Z}^2$. Let p_D and p_A be the grid reconstruction systems associated to $(\mathbb{Z}^2, D)$ and (A, D) respectively. Then the natural map $\mathrm{cok}(p_A) \to \mathrm{cok}(p_D)$ is surjective and so the map $\mathrm{Dep}(p_D) \to \mathrm{Dep}(p_A)$ is injective.*

Proof. Fixing notation, write $p_D : T_D \to P_D$ and $p_A : T_A \to P_A$. We want to show that for every $x \in P_D$ there is a $y \in T_D$ such that $x + p_D(y)$ is in P_A (viewed as a subset of P_D). So, let $x \in P_D$ and write $x = (x_1, \ldots, x_n)$ with $x_i \in k[\mathbb{Z}^2/d_i]$. Let $F_1, \ldots, F_n$ be as in theorem (3.1.5). Then F_i maps to 0 in $k[\mathbb{Z}^2/d_j]$ whenever $j \neq i$, so it suffices to show that for every i, there is a $y_i \in k[\mathbb{Z}^2]$ such that $x_i + p_{d_i}(y_i F_i)$ is contained in $k[A_i] \subset k[\mathbb{Z}^2/d_i]$.

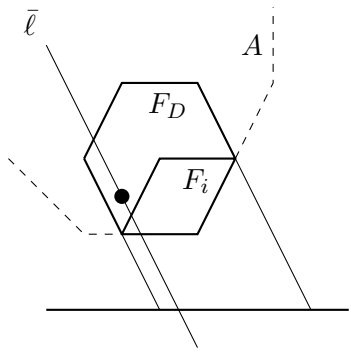
As in theorem (3.5.5), there is an isomorphism $\phi_{d_i} : k[\mathbb{Z}^2/d_i] \to k[w, w^{-1}]$. Following the proof of theorem (3.5.5), the image of $p_{d_i}(F_i)$ in $k[w, w^{-1}]$ is

the polynomial

$$f_i = \sum_{j=1}^{r_i} a_{i,j} w^{s+j}$$

and moreover, $a_{i,1}$ and a_{i,r_i} are both ± 1. It is clear that if we have any r_i consecutive powers of w, we can add to x_i suitable multiples of translates of f_i to obtain a polynomial whose support is contained in these consecutive powers of w. So what we need to show is that the image of $k[A_i]$ in $k[w, w^{-1}]$ contains such a set of r_i consecutive powers of w.

Our assumption on A is that $t + P(F_D)$ is contained in $H(A)$. Note that $P(F_i) \subset P(F_D)$, so $u^t F_i$ has support in A and $p_{d_i}(u^t F_i)$ is $w^t f_i$, which hits lines $s+t+1$ and $s+t+r_i$. What remains to be shown is that all the (lattice) lines between these two are also in A_i.



Let $\ell = \{p + \lambda d_i \mid \lambda \in \mathbb{Z}\}$ be such a line. By the convexity of $P(F_D)$, the geometric line $\bar{\ell} = \{p + \lambda d_i \mid \lambda \in \mathbb{R}\}$ intersects $P(F_D)$.

Moreover, since F_D is a multiple of $u^{d_i} - 1$, $P(F_D)$ is L-rounded, where L is the segment $P(u^{d_i} - 1)$. It follows that the intersection of $\bar{\ell}$ and $P(F_D)$ is at least as wide as the line segment L. Hence at least one integral point lies in this segment, that is, the intersection of ℓ and $P(F_D)$ is non-empty. As A is a convex grid set, this point must also be in A, hence ℓ is in A_i. □

Corollary 3.5.9. *Let D be a 2-regular sequence of at least 2 directions in $\mathbb{Z}^2$ and let A be a non-empty D-rounded convex grid set in $\mathbb{Z}^2$. Let k be a commutative ring and let p_A and p_D be the grid reconstruction systems over k corresponding to (A, D) and $(\mathbb{Z}^r, D)$ respectively. Then the natural map $\mathrm{cok}(p_A) \to \mathrm{cok}(p_D)$ is an isomorphism.*

Proof. By theorem (3.5.8) the map is surjective, and by theorem (3.4.3) it is injective. □

Corollary 3.5.10. *Let D be a 2-regular sequence of at least 2 directions in $\mathbb{Z}^2$ and let A be a convex grid set in $\mathbb{Z}^2$ such that the D-rounded part R of A is non-empty. Denote by p_A and p_D respectively the grid reconstruction systems corresponding to (A, D) and $(\mathbb{Z}^r, D)$. Let $p_{A/R}$ be the reconstruction system defined in* (2.4.4) *for the extension $R \subset A$. Then* $\mathrm{Dep}(p_A)$ *decomposes as a direct sum*

$$\mathrm{Dep}(p_A) \cong \mathrm{Dep}(p_D) \oplus \mathrm{Dep}(p_{A/R}).$$

Proof. As D contains at least 2 independent directions, Δ_D has non-empty interior, and so R has as well. Hence corollary (3.4.14) applies and there is a direct sum decomposition

$$\mathrm{cok}(p_A) \cong \mathrm{cok}(p_R) \oplus \mathrm{cok}(p_{A/R}),$$

From the previous corollary we know that $\mathrm{cok}(p_R) \cong \mathrm{cok}(p_D)$. Taking homomorphisms to k we obtain the required decomposition of the dependencies. □

4 – Computing dependencies

4.1 Introduction

The aim of this chapter is to give an explicit way to calculate the dependencies for the grid reconstrution system associated to $(\mathbb{Z}^2, D)$, where D is a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. We know from theorem (3.5.1) that the cokernel and the space of dependencies are free of finite rank over k in this case, for every commutative ring k. Hence, by lemma (2.5.4), the space of dependencies over k is just the base change of the space of dependencies over $\mathbb{Z}$, so it suffices to give a basis over $\mathbb{Z}$.

As it turns out, the best way to do this is to first consider the reconstruction system over $\overline{\mathbb{Q}}$, where techniques from commutative algebra and algebraic geometry can be used to give an explicit description of the dependencies. This will be done in section 4.2. Using Galois theory we can then find the dependencies over $\mathbb{Q}$ and within this space we can recognise the dependencies over $\mathbb{Z}$ in a straightforward way. This will be the subject of section 4.3.

Remark 4.1.1. Let k be a commutative ring and $D = (d_1, \ldots, d_n)$ a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. Let $p : T \to P$ be the grid reconstruction system over k associated to $(\mathbb{Z}^2, D)$. Following example (3.1.1) we can identify $k[\mathbb{Z}^2]$ with the Laurent polynomial ring $k[u, u^{-1}, v, v^{-1}]$. Similarly, for each $d_i = (a_i, b_i)$ in D there is a group isomorphism

$$\mathbb{Z}^2/d_i \longrightarrow \mathbb{Z} \quad \overline{(x,y)} \mapsto a_i y - b_i x.$$

This gives rise to an identification of $k[\mathbb{Z}^2/d_i]$ with $k[w, w^{-1}]$. Under these identifications, the projection map p_{d_i} becomes the ring homomorphism

$$\begin{array}{ccc} k[u, u^{-1}, v, v^{-1}] & \longrightarrow & k[w, w^{-1}] \\ u & \mapsto & w^{-b_i} \\ v & \mapsto & w^{a_i}. \end{array}$$

The entire reconstruction system p can in this way be viewed as a map

$$k[u, u^{-1}, v, v^{-1}] \longrightarrow k[w, w^{-1}]^n.$$

In remark (3.5.3) we described dependencies $d \in \mathrm{Dep}(p)$ via sequences $c_i : \mathbb{Z} \to k$ for $i = 1, \ldots, n$. Note that if we consider the dependency d as a

k-linear map $k[w, w^{-1}]^n \to k$ following the identifications above, then $c_i(j)$ is the image under d of w^j from the i-th copy of $k[w, w^{-1}]$.

It is convenient to give names to some (Laurent) polynomials that come up frequently in the study of these reconstruction systems. For $i = 1, \ldots, n$, let f_i be $u^{a_i} v^{b_i} - 1$. Note that this is the element corresponding to $u^{d_i} - 1$ under the identifications above. Write

$$F = f_1 \cdots f_n$$

and for $i = 1, \ldots, n$,

$$F_i = \prod_{j \neq i} f_j.$$

Note that F corresponds to the generator of $\ker(p)$ described in theorem (3.1.2) and the F_i correspond to their namesakes from theorem (3.1.5).

Before we begin our study of the dependencies of these reconstruction systems, we consider the situation briefly from the point of view of algebraic geometry. The reader who is not familiar with this subject should skip ahead to the start of section 4.2. No details or rigorous statements are provided for this discussion of the geometric situation; it serves as a 'big picture' look on the technical material in section 4.2.

A geometric perspective

Let k be an algebraically closed field of characteristic 0 and let

$$D = (d_1, \ldots, d_n)$$

be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. Let p_D be the grid reconstruction system over k associated to $(\mathbb{Z}^2, D)$. Following remark (4.1.1), we can view it as a map between Laurent polynomial rings. Its kernel is the principal ideal generated by F as in the remark. Dividing out this kernel we obtain a ring homomorphism

$$p : k[u, u^{-1}, v, v^{-1}]/F \longrightarrow \prod_{i=1}^{n} k[w, w^{-1}]$$

which is injective and whose cokernel is the same as that of p_D.

Let $T = \mathrm{Spec}(k[u, u^{-1}, v, v^{-1}])$. This is the torus $\mathbb{G}_{\mathrm{m}}^2$ over $\mathrm{Spec}(k)$. Let $X_i = \mathrm{Spec}(k[u, u^{-1}, v, v^{-1}]/f_i)$. Then $X_i \subset T$ is a closed subscheme. As

d_i is primitive, f_i is an irreducible element of $k[u, u^{-1}, v, v^{-1}]$ and so X_i is reduced. The map induced by p_{d_i} is an isomorphism between X_i and $\mathbb{G}_{\mathrm{m}} = \mathrm{Spec}(k[w, w^{-1}])$. Let

$$X = \bigcup_{i=1}^{n} X_i = \mathrm{Spec}(k[u, u^{-1}, v, v^{-1}]/F)$$

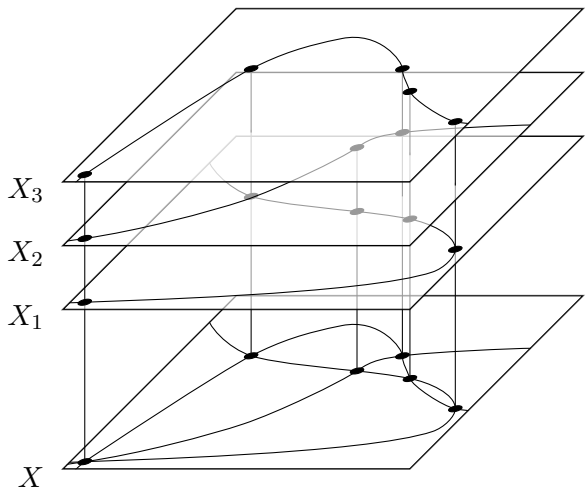
and note that

$$\mathrm{Spec}(\prod_{i=1}^{n} k[w, w^{-1}]) = \coprod_{i=1}^{n} \mathbb{G}_{\mathrm{m}}.$$

The map p corresponds to the projection map

$$\coprod_{i=1}^{n} \mathbb{G}_{\mathrm{m}} \longrightarrow \bigcup_{i=1}^{n} X_i.$$

In a picture, we have



From this it is clear that outside the intersection points of the X_i the map p is an isomorphism. Hence the cokernel as a sheaf is supported only above these intersection points and can be understood by computing the stalk over each of them.

Computing these stalks is a local affair. As the directions d_i are pairwise independent, the intersections of the X_i are transversal. It is therefore not surprising that we can study the local behaviour at the intersection point by just looking at the tangent lines of the X_i at this point. Formally, this can be

achieved by going to the complete local ring at the intersection point, where there is a natural logarithm map that identifies a formal neighbourhood of the point in T with a formal neighbourhood of the affine plane. We lose no information by going to the completion as the stalks we are trying to compute have finite dimension over k.

Lines in the plane have the nice property that they intersect in no more that one point. Hence looking locally at an intersection of lines is the same as looking globally at the same intersection. At this point, the module we are trying to compute has such an easy explicit description that one can just write down a basis for it. Tracing back through all the isomorphisms this allows us to give explicit dependencies for p.

4.2 Dependencies over $\overline{\mathbb{Q}}$

In this section we make a detailed study of grid reconstruction systems over $\overline{\mathbb{Q}}$ associated to the full grid $\mathbb{Z}^2$. While the central result, theorem (4.2.2) is perfectly elementary, the proof is firmly rooted in commutative algebra. Effort has been taken to keep all the steps as simple as possible and references will be given for general results that are needed.

Remark 4.2.1. Roots of unity come up naturally in the computations that we are going to do. It is important that we fix a 'compatible choice' of all roots of unity, which we do in the following way. Fix an embedding $\overline{\mathbb{Q}} \subset \mathbb{C}$ and let ζ be the group homomorphism

$$\begin{array}{rccc} \zeta : \mathbb{Q}/\mathbb{Z} & \longrightarrow & \overline{\mathbb{Q}}^{\times} \\ x & \longmapsto & e^{2\pi \mathrm{i} x}. \end{array}$$

The aim of this section is to prove the following theorem.

Theorem 4.2.2. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$ and p_D the grid reconstruction system associated to $(\mathbb{Z}^2, D)$ over $k = \overline{\mathbb{Q}}$. Write $d_i = (a_i, b_i)$ for all $i \in \{1, \ldots, n\}$. Let $(m_x, m_y) \in (\mathbb{Q}/\mathbb{Z})^2$ and $s, t \in \mathbb{Z}_{\geq 0}$. Put*

$$S = \{i \mid a_i m_x + b_i m_y \in \mathbb{Z}\},$$

put $d = \#S - 2 - s - t$ and suppose that $d \geq 0$. As in remark (4.1.1), view p_D as a map

$$k[u, u^{-1}, v, v^{-1}] \longrightarrow k[w, w^{-1}]^n.$$

Let

$$\delta : k[w, w^{-1}]^n \longrightarrow k$$

be the k-linear map given by the following formulas. On the i-th component, put

$$\begin{aligned} w^m &\longmapsto 0 && \text{if } i \notin S \\ &\longmapsto \frac{r_i(-b_i)^s a_i^t}{d!}\zeta_i^m m^d && \text{if } i \in S, \end{aligned}$$

where

$$\zeta_i = \zeta(a_i' m_x + b_i' m_y)$$

with $a_i', b_i' \in \mathbb{Z}$ such that $a_i\, b_i' - b_i\, a_i' = 1$ and

$$r_i = (-1)^{\#\{j \in S \;\mid\; j \leq i\}} \prod_{j_1, j_2} (a_{j_1} b_{j_2} - a_{j_2} b_{j_1}),$$

with the product running over $j_1, j_2 \in S \setminus \{i\}$ such that $j_1 < j_2$. Then $\delta \circ p_D = 0$, so it corresponds to an element of $\mathrm{Dep}(p_D)$. Moreover, running over all (m_x, m_y), s and t satisfying $d \geq 0$ produces a basis of $\mathrm{Dep}(p_D)$.

We will see what this theorem says about example (1.4.3) later in this section. This example will be built up along with the various intermediate steps of the proof.

The theorem gives a very explicit description of a basis for the space of dependencies. It is not surprising that this explicit nature translates into an algorithmic statement. One has to be a little careful with this, as the output a priori is infinite. The following corollary gives an algorithmic result that is useful for a consistency algorithm similar to the one discussed in corollary (3.5.7).

Corollary 4.2.3. *There is an algorithm that given a 2-regular sequence of primitive directions in D produces for the grid reconstruction system p_D over $\overline{\mathbb{Q}}$ associated to $(\mathbb{Z}^2, D)$ in time polynomial in $\dim \mathrm{Dep}(p_D)$ an explicit description of a basis for $\mathrm{Dep}(p_D)$ such that for every dependency in the basis and every line ℓ in a direction $d_i \in D$ the weight of that line in that dependency can be computed in polynomial time in $\log(|m|)$ where m is the image of ℓ in $\mathbb{Z}^2/d_i \xrightarrow{\sim} \mathbb{Z}$.*

The rest of this section is devoted to proving theorem (4.2.2). We use methods from commutative algebra to do this. We refer the reader to [8] and [1] for more background on commutative algebra. The strategy of the proof is as outlined in the geometric discussion from the previous section.

Local decomposition of the cokernel

Lemma 4.2.4. *Let k be a commutative ring, let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$ and let p_D be the grid reconstruction system over k corresponding to $(\mathbb{Z}^2, D)$. Let $F_1, \ldots, F_n$ be as in remark* (4.1.1) *and put*

$$A = k[u, u^{-1}, v, v^{-1}]/(F_1, \ldots, F_n)$$

and

$$B = \bigoplus_{i=1}^{n} k[w, w^{-1}]/p_{d_i}(F_i).$$

Then p_D induces a ring homomorphism p on the quotients

$$p : A \longrightarrow B.$$

The reconstruction system p is injective and has cokernel $\operatorname{cok}(p) = \operatorname{cok}(p_D)$.

Proof. Note that as in theorem (3.1.5) the image of the ideal $(F_1, \ldots, F_n)$ under p_D is the ideal

$$I = \bigoplus_{i=1}^{n} p_{d_i}(F_i)k[w, w^{-1}].$$

It follows that there is indeed a well-defined map p induced by p_D between the quotients A and B.

Let $x \in A$ and suppose that $p(x) = 0$. Let $\hat{x}$ be any lift of x to $k[u, u^{-1}, v, v^{-1}]$. From the definition of p, we know that $p_D(\hat{x})$ is in I, hence there is some $y \in (F_1, \ldots, F_n)$ such that $p_D(\hat{x}) = p_D(y)$. Hence $\hat{x} - y$ is in $\ker(p_D)$, which is generated by F by theorem (3.1.2). Note that $F \in (F_1, \ldots, F_n)$, so in fact $\hat{x} - y$ is in $(F_1, \ldots, F_n)$ and therefore $\hat{x}$ is. This means that $x = 0$ in A. It follows that the map p is injective.

For the equality of the cokernels, note that dividing out the entire image of p_D at once or first dividing out the image of the ideal $(F_1, \ldots, F_n)$ and then the rest of the image doesn't make any difference. □

Example 4.2.5. As was mentioned at the start of the section, we consider the situation from example (1.4.3) as a running example in this section. The description of the reconstruction system p_D as a map

$$k[u, u^{-1}, v, v^{-1}] \longrightarrow k[w, w^{-1}]^4$$

was given in example (3.1.1).

The generators for the ideal defining A are

$$\begin{aligned} F_1 = &\ u^4v^3 - u^3v^3 - u^3v - u^2v^2 + u^2v + uv^2 + u - 1, \\ F_2 = &\ u^3v^2 - u^3v - u^2v^2 + u^2v - uv + u + v - 1, \\ F_3 = &\ u^2v^3 - u^2v^2 - uv^3 + uv^2 - uv + u + v - 1, \\ F_4 = &\ u^3v^4 - u^3v^3 - u^2v^2 + u^2v - uv^3 + uv^2 + v - 1. \end{aligned}$$

Another basis for this ideal (a Gröbner basis, in fact) is given by

$$\begin{aligned} &u^4 - 2u^3 + 2u - v^4 + 2v^3 - 2v, \\ &u^2v - u^2 - 2uv + 2u + v^5 - 2v^4 + 2v^2 - 1, \\ &uv^2 - 2uv + u + v^5 - 2v^4 + v^2 + v - 1, \\ &v^6 - v^5 - v^4 + v^2 + v - 1. \end{aligned}$$

From this basis one easily sees that A is of dimension 10 as a vector space over k and the following monomials form a basis of A:

$$\{1, v, v^2, v^3, v^4, v^5, u, uv, u^2, u^3\}.$$

The images of the $p_{d_i}(F_i)$ inside $k[w, w^{-1}]$ are given by

$$\begin{aligned} p_{d_1}(F_1) = &\ w^{-4} - 2w^{-3} + 2w^{-1} - 1, \\ p_{d_2}(F_2) = &\ -w^{-5} + w^{-4} + w^{-3} - w^{-1} - 1 + w, \\ p_{d_3}(F_3) = &\ -w^5 + w^4 + w^3 - w - 1 + w^{-1}, \\ p_{d_4}(F_4) = &\ w^4 - 2w^3 + 2w - 1. \end{aligned}$$

It follows that the vector spaces $k[w, w^{-1}]/p_{d_i}(F_i)$ have dimensions 4, 6, 6 and 4 respectively for $i = 1, 2, 3, 4$. Hence B has dimension 20. It follows that $\mathrm{cok}(p)$ has dimension $20 - 10 = 10$, which is consistent with what we showed in example (3.5.2).

For the rest of this section, fix a 2-regular sequence $D = (d_1, \ldots, d_n)$ of primitive directions in $\mathbb{Z}^2$. Let p_D be the grid reconstruction system over $k = \overline{\mathbb{Q}}$ associated to $(\mathbb{Z}^2, D)$. Let $p : A \to B$ be the reconstruction system between quotient rings as described in lemma (4.2.4) above.

Theorem 4.2.6. *Write $d_i = (a_i, b_i)$ for $i = 1, \ldots, n$. Then there is a bijection between*

$$M := \left\{ (x, y) \in (\mathbb{Q}/\mathbb{Z})^2 \;\middle|\; \exists\, i, j : \begin{array}{c} a_i x + b_i y \in \mathbb{Z} \\ a_j x + b_j y \in \mathbb{Z} \\ i \neq j \end{array} \right\}$$

and

$$\{\mathfrak{m} \subset A \mid \mathfrak{m} \textit{ maximal ideal}\}$$

sending (x, y) *to the ideal* $(u - \zeta(x), v - \zeta(y))$.

Proof. The Nullstellensatz (see, e.g. [8, Sect. 1.6, Cor. 1.9]) implies that there is a bijective map

$$\left\{(\alpha,\beta) \in k^2 \;\middle|\; \begin{array}{c} F_i(\alpha,\beta) = 0 \\ \text{for all } i \end{array}\right\} \longrightarrow \{\mathfrak{m} \subset A \mid \mathfrak{m} \text{ maximal ideal }\}$$
$$(\alpha,\beta) \longmapsto (u-\alpha, v-\beta).$$

What remains to be shown is that the pairs (α, β) that come up are precisely $(\zeta(x), \zeta(y))$ for $(x, y) \in M$.

Let $(\alpha, \beta) \in k^2$. Note that $F_i(\alpha, \beta) = 0$ if and only if there is a $j \neq i$ such that $f_j(\alpha, \beta) = 0$. It follows that (α, β) is a common zero of $(F_1, \ldots, F_n)$ if and only if

$$\exists\, i, j : i \neq j \text{ and } f_i(\alpha,\beta) = f_j(\alpha,\beta) = 0.$$

Suppose that (α, β) is common zero of f_i and f_j. Then we have

$$\alpha^{a_i}\beta^{b_i} = 1 \quad \text{and} \quad \alpha^{a_j}\beta^{b_j} = 1$$

so

$$\alpha^{a_ib_j - a_jb_i} = \alpha^{a_ib_j}(\alpha^{-a_j})^{b_i} = \alpha^{a_ib_j}(\beta^{b_j})^{b_i} = (\alpha^{a_i}\beta^{b_i})^{b_j} = 1$$

and similarly $\beta^{a_ib_j - a_jb_i} = 1$. It follows that α and β are roots of unity. Hence there are x and y in $\mathbb{Q}/\mathbb{Z}$ such that $\alpha = \zeta(x)$ and $\beta = \zeta(y)$. Moreover, these satisfy

$$1 = \zeta(x)^{a_i}\zeta(y)^{b_i} = \zeta(a_ix + b_iy)$$

and

$$1 = \zeta(x)^{a_j}\zeta(y)^{b_j} = \zeta(a_jx + b_jy),$$

that is, $a_ix + b_iy \in \mathbb{Z}$ and $a_jx + b_jy \in \mathbb{Z}$. So (x, y) is in M.

Conversely, if we have any $(x, y) \in M$ with $a_ix+b_iy \in \mathbb{Z}$ and $a_jx+b_jy \in \mathbb{Z}$ then it is clear that $(\zeta(x), \zeta(y))$ is a common zero of f_i and f_j. This completes the proof. □

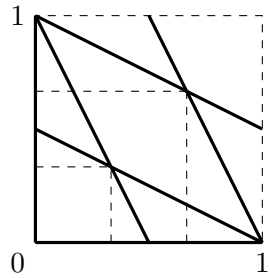
Example 4.2.7. In this example we will compute the set M for the sequence D from our running example, (1.4.3). A fundamental domain for the action of $\mathbb{Z}^2$ on $\mathbb{Q}^2$ is given by

$$\{(x, y) \in \mathbb{Q}^2 \mid 0 \leq x < 1 \text{ and } 0 \leq y < 1\}.$$

In the picture below are the images of the solution sets

$$\{(x, y) \mid a_i x + b_i y \in \mathbb{Z}\}$$

inside this fundamental domain.



We see that the set M is given by

$$M = \left\{(0,0), \left(0, \frac{1}{2}\right), \left(\frac{1}{2}, 0\right), \left(\frac{1}{3}, \frac{1}{3}\right), \left(\frac{2}{3}, \frac{2}{3}\right)\right\}.$$

One verifies easily that the polynomials $F_1, \ldots, F_4$ as given in example (4.2.5) indeed all have roots in the corresponding points in k^2.

One of the consequences of the previous theorem is that the ring A has only finitely many maximal ideals. The idea is now to study $p : A \to B$ 'one maximal ideal at a time'. This is formalised by the concept of *localisation*. For general results on localisations, we refer the reader to [8, Ch. 2]. A brief introduction is also given in [17, Sect. 2.4]. We follow the standard notation in the literature and write $R_{\mathfrak{m}}$ for the localisation of a commutative ring R at the maximal ideal $\mathfrak{m}$.

Lemma 4.2.8. *For $(x, y) \in M$, let $\mathfrak{m}_{(x,y)}$ be the corresponding maximal ideal of A, let $A_{(x,y)} = A_{\mathfrak{m}_{(x,y)}}$ be the localisation of A at $\mathfrak{m}_{(x,y)}$, let $B_{(x,y)}$ be $A_{\mathfrak{m}_{(x,y)}} \otimes_A B$ and let $p_{(x,y)}$ be the reconstruction system*

$$\begin{array}{rccc} p_{(x,y)} : A_{(x,y)} & \longrightarrow & B_{(x,y)} \\ a & \mapsto & a \otimes 1. \end{array}$$

Then the maps

$$\begin{array}{ccc} A & \longrightarrow & \bigoplus_{(x,y)\in M} A_{(x,y)} \\ a & \mapsto & (a, \ldots, a) \end{array}$$

and

$$\begin{array}{ccc} B & \longrightarrow & \bigoplus_{(x,y)\in M} B_{(x,y)} \\ b & \mapsto & (1\otimes b, \ldots, 1\otimes b) \end{array}$$

are isomorphisms. Together they form an isomorphism

$$p \cong \bigoplus_{(x,y)\in M} p_{(x,y)}.$$

Proof. Note that for each i, $k[w,w^{-1}]/p_{d_i}(F_i)$ is a finite dimensional k-vector space. It follows that B, and hence by the injectivity of p also A, is a finite dimensional k-vector space. Hence A is an Artinian ring. So the natural map

$$A \longrightarrow \bigoplus_{\mathfrak{m}} A_{\mathfrak{m}}$$

where the sum runs over the maximal ideals $\mathfrak{m}$ of A, is an isomorphism (see e.g. [8, Sect. 2.4, Cor. 2.16]). Recall from theorem (4.2.6) that the maximal ideals of A are in bijection with the elements of M.

As p is a ring homomorphism, the tensor product $A \otimes_A B$ via p is well-defined and isomorphic to B. As $A = \bigoplus_{(x,y)\in M} A_{(x,y)}$ and tensor products commute with direct sums (see e.g. [17, Ch. 16, Prop. 2.1]), this is in turn isomorphic to

$$\bigoplus_{(x,y)\in M} A_{(x,y)} \otimes B = \bigoplus_{(x,y)\in M} B_{(x,y)}.$$

Moreover, $p : A \to B$ is isomorphic as a reconstruction system to

$$\begin{array}{ccc} A & \longrightarrow & A \otimes_A B \\ a & \longmapsto & a \otimes 1, \end{array}$$

and this is in turn isomorphic to

$$\begin{array}{ccc} \bigoplus_{(x,y)\in M} A_{(x,y)} & \longrightarrow & \bigoplus_{(x,y)\in M} (A_{(x,y)} \otimes_A B) \\ (a_{(x,y)})_{(x,y)} & \longmapsto & (a_{(x,y)} \otimes 1)_{(x,y)}, \end{array}$$

yielding the isomorphism of reconstruction systems as claimed. □

Example 4.2.9. We computed in example (4.2.7) that the maximal ideals of A in our running example are

$$\begin{array}{lll} \mathfrak{m}_1 = & (u-1, v-1) & \text{corresponding to the point } (0,0), \\ \mathfrak{m}_2 = & (u-1, v+1) & \text{corresponding to the point } (0,1/2), \\ \mathfrak{m}_3 = & (u+1, v-1) & \text{corresponding to the point } (1/2,0), \\ \mathfrak{m}_4 = & (u-\mu, v-\mu) & \text{corresponding to the point } (1/3,1/3), \\ \mathfrak{m}_5 = & (u-\mu^2, v-\mu^2) & \text{corresponding to the point } (2/3,2/3), \end{array}$$

where $\mu = \zeta(1/3)$.

Since all four f_i have a root in $(1,1)$, every F_i is in the ideal $(u-1, v-1)^3$, and so the ideal $(F_1, \ldots, F_4)$ is contained in $(u-1, v-1)^3$ and there is a natural map

$$A \longrightarrow k[u, u^{-1}, v, v^{-1}]/(u-1, v-1)^3.$$

The other elements of M correspond to points in k^2 that are a root of precisely two of the f_i. Therefore each F_i is contained in the corresponding maximal ideal. Altogether this means there are natural maps $A \longrightarrow A_i$ for $i = 1, \ldots, 5$, with the A_i given by

$$\begin{array}{ll} A_1 = & k[u, u^{-1}, v, v^{-1}]/(u-1, v-1)^3, \\ A_2 = & k[u, u^{-1}, v, v^{-1}]/(u-1, v+1), \\ A_3 = & k[u, u^{-1}, v, v^{-1}]/(u+1, v-1), \\ A_4 = & k[u, u^{-1}, v, v^{-1}]/(u-\mu, v-\mu), \\ A_5 = & k[u, u^{-1}, v, v^{-1}]/(u-\mu^2, v-\mu^2). \end{array}$$

Using some additional theory (the fact that the f_i intersect transversally) one can show that for every i, A_i is the localisation of A at $\mathfrak{m}_i$.

Note that the dimension of A_1 as a k-vector space is 6: a basis is given by the monomials $1, u, v, u^2, uv, v^2$. The other A_i each have dimension 1 as k-vector spaces. Put together, these give a map

$$A \longrightarrow A_1 \oplus \cdots \oplus A_5$$

between two vector spaces of dimension 10. A computation shows that his map is indeed an isomorphism.

The following two lemmas will show that if we have a way to compute the localisation at $(u-1, v-1)$, we can use that to compute the localisations at other maximal ideals of A as well.

Lemma 4.2.10. *Let $(x, y) \in M$ and Let D' be the sub-sequence of D consisting of those d_i such that $a_i x + b_i y \in \mathbb{Z}$. Let p'_D be the grid reconstruction system over k corresponding to $(\mathbb{Z}^2, D')$ and $p' : A' \to B'$ the corresponding map on the quotients as defined in lemma (4.2.4). Then there is a natural map $p \to p'$ and it induces an isomorphism of reconstruction systems $p_{(x,y)} \cong p'_{(x,y)}$.*

Proof. The set $I = \{i \mid a_i x + b_i y \in \mathbb{Z}\}$ is the natural index set for elements of D'. Using this numbering, for $i \in I$ let F'_i and B'_i be the analogues of F_i and B_i, i.e.

$$F'_i = \prod_{j \in I \setminus \{i\}} f_j$$

and

$$B'_i = k[w, w^{-1}]/p_{d_i}(F'_i).$$

We extend these definitions for $i \notin I$ by putting $F'_i = \prod_{j \in I} f_j$ and $B'_i = 0$.

Using these notations, we have

$$A' = k[u, u^{-1}, v, v^{-1}]/(F'_1, \ldots, F'_n)$$

and

$$B' = \prod_{i=1}^{n} B'_i.$$

Moreover, as F'_i is a divisor of F_i for all i, there are natural surjective maps $A \to A'$ and $B_i \to B'_i$ for all i. It is clear that the diagram

$$\begin{array}{ccc} A & \xrightarrow{p} & B \\ \downarrow & & \downarrow \\ A' & \xrightarrow[p']{} & B' \end{array}$$

commutes, i.e., there is a map of reconstructions systems $p \to p'$. We now need to show that after localisation, the maps become isomorphisms.

Let $\mathfrak{m}$ be the maximal ideal of $k[u, u^{-1}, v, v^{-1}]$ generated by $u - \zeta(x)$ and $v - \zeta(y)$. Let R be the localisation of $k[u, u^{-1}, v, v^{-1}]$ at $\mathfrak{m}$. For $i \notin I$, the polynomial f_i doesn't have a root at $(\zeta(x), \zeta(y))$, so f_i is not in $\mathfrak{m}$. Hence f_i is a unit in R for all $i \notin I$.

It follows that for all $i = 1, \ldots, n$, the quotient F_i/F'_i is a unit in R. Therefore,

$$(F_1, \ldots, F_n) = (F'_1, \ldots, F'_n)$$

holds in R. Hence

$$A_{(x,y)} = R/(F_1, \ldots, F_n) = R/(F_1', \ldots, F_n') = A'_{(x,y)}.$$

For $i = 1, \ldots, n$, put

$$S_i = R \otimes_{k[u,u^{-1},v,v^{-1}]} k[w, w^{-1}],$$

where the tensor product is taken relative to p_{d_i}. Then

$$(B_i)_{(x,y)} = A_{(x,y)} \otimes_A B_i$$

can be seen as the quotient of S_i first by $(F_1 \otimes 1, \ldots, F_n \otimes 1)$ and then by $(1 \otimes p_{d_i}(F_i))$. But in S_i we have $(1 \otimes p_{d_i}(F_i)) = F_i \otimes 1$, so in fact

$$(B_i)_{(x,y)} = S_i/(F_1 \otimes 1, \ldots, F_n \otimes 1).$$

For $i \in I$, we have by a similar argument

$$(B_i')_{(x,y)} = S_i/(F_1' \otimes 1, \ldots, F_n' \otimes 1).$$

As remarked before F_j/F_j' is a unit in R for all j. It follows at once that for $i \in I$, we have

$$(B_i)_{(x,y)} = S_i/(F_1 \otimes 1, \ldots, F_n \otimes 1) = S_i/(F_1' \otimes 1, \ldots, F_n' \otimes 1) = (B_i')_{(x,y)}.$$

If $i \notin I$, then f_i is a unit in R, so $f_i \otimes 1$ is a unit in S_i. But

$$f_i \otimes 1 = 1 \otimes p_{d_i}(f_i) = 1 \otimes 0 = 0,$$

so S_i is the zero ring. In particular $(B_i)_{(x,y)} = 0 = (B_i')_{(x,y)}$ in this case. □

Lemma 4.2.11. *Let $(x, y) \in M$ and suppose that $a_i x + b_i y \in \mathbb{Z}$ for all i. Let $\mathfrak{m}$ be the maximal ideal of A corresponding to (x, y) Then there is an isomorphism $s : p \to p$ such that $s_A(\mathfrak{m}) = (u - 1, v - 1)$.*

Proof. Let S be the ring automorphism

$$\begin{array}{rcl} k[u, u^{-1}, v, v^{-1}] & \longrightarrow & k[u, u^{-1}, v, v^{-1}] \\ u & \longmapsto & \zeta(x)u \\ v & \longmapsto & \zeta(y)v. \end{array}$$

Note that $S(f_i) = f_i$ for all $i = 1, \ldots, n$, as

$$\begin{array}{ll} (\zeta(x)u)^{a_i}(\zeta(y)v)^{b_i} - 1 & = \zeta(x)^{a_i}\zeta(y)^{b_i}u^{a_i}v^{b_i} - 1 \\ & = \zeta(a_i x + b_i y)u^{a_i}v^{b_i} - 1 = u^{a_i}v^{b_i} - 1. \end{array}$$

It follows that there is a well-defined ring automorphism $s_A : A \to A$ induced by S. Moreover, it is clear that $s_A(\mathfrak{m}) = (u-1, v-1)$. It remains to be shown that there is a ring automorphism $s_B : B \to B$ such that the diagram

$$\begin{array}{ccc} A & \xrightarrow{p} & B \\ {\scriptstyle s_A}\downarrow & & \downarrow{\scriptstyle s_B} \\ A & \xrightarrow[p]{} & B \end{array}$$

commutes.

As a_i and b_i are coprime, there are a_i' and b_i' such that $a_i b_i' - b_i a_i' = 1$. Let T_i be the ring automorphism

$$\begin{array}{rcl} k[w, w^{-1}] & \longrightarrow & k[w, w^{-1}] \\ w & \longmapsto & \zeta(a_i' x + b_i' y) w. \end{array}$$

Then we have $T_i(p_{d_i}(u)) = p_{d_i}(S(u))$ as

$$\begin{array}{ll} \zeta(a_i' x + b_i' y)^{-b_i} & = \zeta(-b_i a_i' x - b_i b_i' y) = \zeta((1 - a_i b_i')x - b_i b_i' y) \\ & = \zeta(x - b_i'(a_i x + b_i y)) = \zeta(x), \end{array}$$

for $b_i'(a_i x + b_i y)$ is in $\mathbb{Z}$. Likewise, we have $T_i(p_{d_i}(v)) = p_{d_i}(S(v))$. It follows that

$$T_i(p_{d_i}(F_i)) = p_{d_i}(S(F_i)) = p_{d_i}(F_i).$$

Hence T_i induces a ring automorphism $s_{B,i} : B_i \to B_i$. Putting these together yields a ring automorphism $s_B : B \to B$. The required relations with s_A come from the computation we did for T_i and S. □

The upshot of these two lemmas is that to think about all $p_{(x,y)}$, it is enough to just think about $p_{(0,0)}$. Hence the logical next step is to determine $\mathrm{Dep}(p_{(0,0)})$. The strategy for this is to 'take logarithms', which turns the problem into a linear one, and solve this linear problem explicitly.

Linearising the problem

Consider the following 'linear analogue' of the reconstruction system p. Let

$$\hat{A} = k[x, y]/(x, y)^{n-1} \quad \text{and} \quad \hat{B} = \left(k[z]/z^{n-1}\right)^n.$$

Let $\hat{p}$ be the reconstruction system given by the ring homomorphism

$$\begin{array}{rccc} \hat{p} : \hat{A} & \longrightarrow & \hat{B} \\ x & \longmapsto & (-b_i z)_i \\ y & \longmapsto & (a_i z)_i \end{array}$$

We will now construct a morphism of reconstruction systems

$$\log : p \longrightarrow \hat{p}$$

and show it induces an isomorphism $p_{(0,0)} \cong \hat{p}$.

Lemma 4.2.12. *Let R be a commutative $\mathbb{Q}$-algebra. For $x \in R$ nilpotent, define*

$$e^x = \sum_{n \geq 0} \frac{1}{n!} x^n.$$

If $x, y \in R$ are nilpotent, then $x + y$ is nilpotent and

$$e^x e^y = e^{x+y}.$$

Proof. First note that e^x is indeed well-defined as the sum is always finite. Now suppose $x, y \in R$ are nilpotent, and that $x^n = y^m = 0$. Then

$$(x+y)^{n+m} = \sum_{i+j=n+m} \binom{n+m}{i} x^i y^j$$

is 0 as in each term either $i \geq n$ or $j \geq m$. Hence $x + y$ is nilpotent. We prove $e^x e^y = e^{x+y}$ by expanding both sides as polynomials in x and y and comparing coefficients. On the lefthand side the coefficient of $x^a y^b$ is $\frac{1}{a!}\frac{1}{b!}$. On the righthand side, it is $\binom{a+b}{a}\frac{1}{(a+b)!}$. These are equal by the well-known formula for binomial coefficients. □

Lemma 4.2.13. *There are ring homomorphisms given by*

$$\begin{array}{rcl} A \longrightarrow \hat{A} & & u \mapsto e^x \\ & & v \mapsto e^y \end{array}$$

and

$$B \longrightarrow \hat{B} \qquad w \mapsto e^z \text{ on every component.}$$

These give rise to a morphism of reconstruction systems

$$\log : p \longrightarrow \hat{p}.$$

Proof. These formulas clearly define ring homomorphisms

$$L_A : k[u, u^{-1}, v, v^{-1}] \longrightarrow \hat{A}$$

and

$$L_B : k[w, w^{-1}]^n \longrightarrow \hat{B}.$$

Note that for $i = 1, \ldots, n$, we have

$$\begin{aligned} L_A(f_i) &= (e^x)^{a_i}(e^y)^{b_i} - 1 = e^{a_i x + b_i y} - 1 \\ &= \textstyle\sum_{n \geq 1} \frac{1}{n!}(a_i x + b_i y)^n, \end{aligned}$$

which is in the ideal (x, y). Hence $L_A(F_i)$ is in $(x, y)^{n-1}$, that is, $L_A(F_i) = 0$. It follows that L_A induces a ring homomorphism $\log_A : A \to \hat{A}$.

In the same way,

$$L_{B,i}(p_{d_i}(f_j)) = e^{(a_i b_j - a_j b_i)z} - 1$$

is in the ideal (z), hence $L_{B,i}(p_{d_i}(F_i))$ is in $(z)^{n-1}$, so it is 0. It follows that L_B induces a ring homomorphism $\log_B : B \to \hat{B}$.

It remains to be shown that $\log_B \circ p = \hat{p} \circ \log_A$. As all the maps involved are ring homomorphisms, it suffices to check this for u and v. We compute

$$\log_B(p(u)) = (e^{-b_i z})_i = \hat{p}(e^x)$$

and

$$\log_B(p(v)) = (e^{a_i z})_i = \hat{p}(e^y).$$

This completes the proof. □

Example 4.2.14. Continuing our running example, we follow some of the computations surrounding the reconstruction system $\hat{p} : \hat{A} \to \hat{B}$ for the directions from example (1.4.3).

The space $\hat{A} = k[x, y]/(x, y)^3$ is a 6-dimensional k-vector space. A basis for this space is given by the monomials $1, x, y, x^2, xy, y^2$. The map L_A from the proof of lemma (4.2.13) is given by

$$\begin{array}{ccc} k[u, u^{-1}, v, v^{-1}] & \longrightarrow & k[x, y]/(x, y)^3 \\ u & \mapsto & 1 + x + \frac{1}{2}x^2 \\ v & \mapsto & 1 + y + \frac{1}{2}y^2. \end{array}$$

Applying this to $F_1 = (uv^2-1)(u^2v-1)(u-1)$ from example (4.2.5) we obtain

$$\begin{aligned} L_A(F_1) &= L_A(uv^2-1)L_A(u^2v-1)L_A(u-1) \\ &= (x+2y+\tfrac{1}{2}x^2+2xy+2y^2) \\ &\quad \cdot(2x+y+2x^2+2xy+\tfrac{1}{2}y^2)(x+\tfrac{1}{2}x^2) \\ &= 0, \end{aligned}$$

and similarly for the other F_i. Hence L_A indeed induces a ring homomorphism $\log_A : A \to \hat{A}$.

The space $\hat{B} = (k[z]/z^3)^4$ is a 12-dimensional k-vector space. Focusing on the fourth coordinate, corresponding to the direction $(1,0)$, we see that the map L_B is given by

$$\begin{array}{ccc} k[w,w^{-1}] & \longrightarrow & k[z]/z^3 \\ w & \mapsto & 1+z+\frac{1}{2}z^2. \end{array}$$

The image of $p_{d_4}(F_4) = w^4 - 2w^3 + 2w - 1$ under this map is

$$\begin{aligned} L_{B,4}(p_{d_4}(F_4)) &= (1+z+\tfrac{1}{2}z^2)^4 - 2(1+z+\tfrac{1}{2}z^2)^3 + 2(1+z+\tfrac{1}{2}z^2) - 1 \\ &= (1+4z+8z^2) - 2(1+3z+\tfrac{9}{2}z^2) + 2(1+z+\tfrac{1}{2}z^2) - 1 \\ &= (1-2+2-1) + (4-6+2)z + (8-9+1)z^2 \\ &= 0, \end{aligned}$$

so it induces a homomorphism

$$\log_{B,4} : k[w,w^{-1}]/p_{d_4}(F_4) \longrightarrow k[z]/z^3.$$

A verification of the compatibility of log with p and $\hat{p}$ in this case is left to the reader. Note that it suffices to verify that u and v in A map to the same element of $\hat{B}$ under $\log_B \circ p$ and $\hat{p} \circ \log_A$.

Theorem 4.2.15. *The morphism* log *induces an isomorphism* $p_{(0,0)} \cong \hat{p}$.

Proof. Note that the natural maps $A \to A_{(0,0)}$ and $B \to B_{(0,0)}$ are surjective, so what we have to show is that $\log_A$ induces an isomorphism $A_{(0,0)} \to \hat{A}$ and that $\log_B$ induces an isomorphism $B_{(0,0)} \to \hat{B}$. Compatibility with $p_{(0,0)}$ and $\hat{p}$ then follows from the fact that the original maps are compatible with p and $\hat{p}$.

First note that $\log_A(u-1)$ is a multiple of x and $\log_A(v-1)$ is a multiple of y, so the ideal $(u-1, v-1)$ of A is mapped into the ideal (x,y) of $\hat{A}$. Moreover,

any $f \in A$ can be written as $f = f(1,1) + g$ with $g \in (u-1, v-1)$. It follows that $\log_A(f) = f(1,1) + \log_A(g)$. Hence any $f \in A$ that is not in $(u-1, v-1)$ has $\log_A(f) \notin (x, y)$. As $\hat{A}$ is a local ring, this implies $\log_A(f)$ is a unit. The universal property of localisations implies that $\log_A$ factors uniquely as the natural map $A \to A_{(0,0)}$ followed by a map $\log_{A_{(0,0)}} : A_{(0,0)} \to \hat{A}$.

The map $\log_{A_{(0,0)}}$ is clearly surjective, as $\log_A$ is. To show that it is also injective, we examine $A_{(0,0)}$ a more closely. Let R be the localisation of $k[u, u^{-1}, v, v^{-1}]$ at $(u-1, v-1)$. Recall that $A_{(0,0)} = R/(F_1, \ldots, F_n)$. We will now show that

$$(F_1, \ldots, F_n) = (u-1, v-1)^{n-1}$$

as ideals of R. Note that

$$u^{a_i} v^{b_i} - 1 = v^{b_i}(u^{a_i} - 1) + (v^{b_i} - 1)$$

is in $(u-1, v-1)$, so F_i is in $(u-1, v-1)^{n-1}$ for all i. This proves one inclusion. By Nakayama's lemma (see e.g. [8, Cor. 4.8]) it now suffices to show that the images of $F_1, \ldots, F_n$ generate $(u-1, v-1)^{n-1}/(u-1, v-1)^n$ as an $R/(u-1, v-1)$-module. Note that

$$\begin{aligned} u^{a_i} v^{b_i} - 1 &= ((u-1)+1)^{a_i}((v-1)+1)^{b_i} - 1 \\ &\equiv a_i(u-1) + b_i(v-1) \qquad \mod (u-1, v-1)^2. \end{aligned}$$

As the (a_i, b_i) are pairwise independent, it follows that f_1 and f_2 generate $(u-1, v-1)/(u-1, v-1)^2$.

Now we proceed by induction on n. The remark we just made shows that for $n = 2$, F_1 and F_2 indeed generate $(u-1, v-1)/(u-1, v-1)^2$. For $n > 2$, put $F_{1,i} = F_i/f_1$ for $i \neq 1$ and $F_{2,i} = F_i/f_2$ for $i \neq 2$. By the induction hypothesis the $F_{1,i}$ generate $(u-1, v-1)^{n-2}$, and so do the $F_{2,i}$. Note that

$$\begin{aligned} (F_1, \ldots, F_n) &= (f_1 F_{1,2}, \ldots, f_1 F_{1,n}) + (f_2 F_{2,1}, \ldots, f_2 F_{2,n}) \\ &= (f_1)(u-1, v-1)^{n-2} + (f_2)(u-1, v-1)^{n-2} \\ &= (f_1, f_2)(u-1, v-1)^{n-2} \end{aligned}$$

and hence $(F_1, \ldots, F_n)/(u-1, v-1)^n = (u-1, v-1)^{n-1}/(u-1, v-1)^n$.

It follows that

$$A_{(0,0)} = k[u, u^{-1}, v, v^{-1}]/(u-1, v-1)^{n-1}$$

as this ring is already local. Moreover, we have

$$u^{-1} = \frac{1}{1-(1-u)} = \sum_{i=0}^{n-2} (1-u)^i$$

and similarly for v^{-1}, so actually $A_{(0,0)} = k[u,v]/(u-1,v-1)^{n-1}$.

Note that for any $i,j \geq 0$,

$$\log_{A_{(0,0)}}((u-1)^i(v-1)^j)$$

is of the form

$$x^i y^j + \text{lexicographically higher terms.}$$

Now let $f \in A_{(0,0)}$ non-zero. Write f as a polynomial in $u-1$ and $v-1$. Suppose $c_{i,j}(u-1)^i(v-1)^j$ is the lexicographically smallest non-vanishing term of f. Then $\log_{A_{(0,0)}}(f)$ has a term $c_{i,j}x^iy^j$ and is therefore non-zero. Hence $\log_{A_{(0,0)}}$ is injective.

We have now shown that $\log_A$ induces an isomorphism

$$\log_{A_{(0,0)}} : A_{(0,0)} \to \hat{A}.$$

What remains is to show the analogous statement for $\log_B$. Specifically, we want to show that for $i = 1, \ldots, n$,

$$\log_{B_i} : B_i \longrightarrow \hat{B}_i$$

induces an isomorphism

$$\log_{(B_i)_{(0,0)}} : (B_i)_{(0,0)} \longrightarrow \hat{B}_i.$$

By definition $(B_i)_{(0,0)} = B_i \otimes_A A_{(0,0)}$ where the tensor product is taken relative to the map p_i. Let

$$\begin{array}{rccl} \lambda : B_i \times A_{(0,0)} & \longrightarrow & \hat{B}_i \\ (b,a) & \longmapsto & \log_{B_i}(b) \cdot \hat{p}_i(\log_{A_{(0,0)}}(a)). \end{array}$$

One checks that λ is bilinear, in fact A-bilinear, as $\log_A$ and $\log_B$ are compatible with p and $\hat{p}$. It follows that λ induces a map on the tensor products

$$\log_{(B_i)_{(0,0)}} : B_i \otimes_A A_{(0,0)} \longrightarrow \hat{B}_i.$$

It is clear that the composition of the natural map $B_i \to B_i \otimes_A A_{(0,0)}$ with this map is indeed equal to $\log_{B_i}$.

Note that $\log_{(B_i)_{(0,0)}}$ is surjective, as $\log_{B_i}$ is, so everything in $\hat{B}$ is the image of an element of the form $b \otimes 1$. As $A \to B_i$ is onto, any tensor, and hence any element of $B_i \otimes_A A_{(0,0)}$ can be written in the form $1 \otimes a$. To show that $\log_{(B_i)_{(0,0)}}$ is injective, we must show that for every $a \in A_{(0,0)}$ such that $\hat{p}(\log_{A_{(0,0)}}(a)) = 0$, we have $1 \otimes a = 0$ in $B_i \otimes_A A_{(0,0)}$.

To this end, we compute the kernel of $\hat{p}_i$. Clearly

$$\hat{p}_i(a_i x + b_i y) = a_i(-b_i z) + b_i(a_i z) = 0,$$

so $(a_i x + b_i y) \subset \ker(\hat{p}_i)$. Suppose that a_i is non-zero (the case $b_i \neq 0$ is analogous) and let V be the subspace of $\hat{A}$ generated by the powers of y. It is clear that $\hat{p}_i$ is injective on V. Note that for every positive integer m,

$$(a_i)^{-m}(a_i x + b_i y)^m = x^m + \text{terms of lower degree in } x,$$

so any $f \in \hat{A}$ can be written as $f = f_0 + f_y$, where f_0 is in $(a_i x + b_i y)$ and f_y is a polynomial in y, i.e., $f_y \in V$. As

$$\hat{p}_i(f) = \hat{p}_i(f_0) + \hat{p}_i(f_y) = 0 + \hat{p}_i(f_y),$$

we see that $f \in \ker(\hat{p})$ if and only if $f_y = 0$, i.e., if and only if $f \in (a_i x + b_i y)$.

Let g be the unique inverse image of $a_i x + b_i y$ under $\log_{A_{(0,0)}}$. Then an $a \in A_{(0,0)}$ satisfies $\hat{p}(\log_{A_{(0,0)}}(a)) = 0$ if and only if $a \in (g)$. Note that

$$\log_{A_{(0,0)}}(f_i) = (e^x)^{a_i})(e^y)^{b_i} - 1 = e^{a_i x + b_i y} - 1 = \sum_{j \geq 1} \frac{1}{j!}(a_i x + b_i y)^j.$$

This is a multiple of $a_i x + b_i y$ and the quotient is 1 modulo (x, y), so is a unit in $\hat{A}$. Hence $\log_{A_{(0,0)}}(f_i)$ and $\log_{A_{(0,0)}}(g)$ differ by a unit, and therefore $(f_i) = (g)$ in $A_{(0,0)}$.

Now it is clear that if $a \in A_{(0,0)}$ is such that $\hat{p}(\log_{A_{(0,0)}}(a)) = 0$, then $a \in (g) = (f_i)$ hence

$$1 \otimes a = 1 \otimes f_i a' = p_{d_i}(f_i) \otimes a' = 0 \otimes a' = 0$$

in $B_i \otimes_A A_{(0,0)}$. Therefore $\log_{(B_i)_{(0,0)}}$ is injective, hence bijective, which completes the proof. □

Example 4.2.16. We continue our running example by verifying that $\log_A$ indeed induces an isomorphism between the local ring A_1 from example (4.2.9) and the ring $\hat{A}$ from example (4.2.14).

Consider the ring homomorphism

$$\begin{array}{ccc} k[u, u^{-1}, v, v^{-1}] & \longrightarrow & k[x, y]/(x, y)^3 \\ u & \mapsto & 1 + x + \frac{1}{2}x^2 \\ v & \mapsto & 1 + y + \frac{1}{2}y^2. \end{array}$$

It is clear that $(u-1, v-1)^3$ is in the kernel of this map, and so it induces a ring homomorphism $\phi : A_1 \to \hat{A}$. Moreover $\log_A$ is the composition of the natural map $A \to A_1$ and ϕ, hence ϕ is the localisation of the map $\log_A$ at $(u-1, v-1)$. We want to show that ϕ is an isomorphism. The table below represents ϕ as a matrix relative to the obvious monomial bases of A_1 and $\hat{A}$.

	1	x	y	x^2	xy	y^2
1	1	0	0	0	0	0
u	1	1	0	$\frac{1}{2}$	0	0
v	1	0	1	0	0	$\frac{1}{2}$
u^2	1	2	0	2	0	0
uv	1	1	1	$\frac{1}{2}$	1	$\frac{1}{2}$
v^2	1	0	2	0	0	2

The determinant of this matrix is 1, thus it is invertible and ϕ is an isomorphism.

Computing dependencies for the linear problem

Having established the relation between $p_{(0,0)}$ and $\hat{p}$, we now turn our attention to computing the dependencies of $\hat{p}$. The explicit, linear nature of this map makes them relatively easy to compute. The approach we take is via a pairing in $\hat{B}$, which identifies it with its dual $\mathrm{Hom}_k(\hat{B}, k)$.

Lemma 4.2.17. *The k-linear map*

$$\begin{array}{rccc} \langle\,,\rangle : & \hat{B} \times \hat{B} & \longrightarrow & k \\ & ((f_i)_i, (g_i)_i) & \longmapsto & \sum_{i=1}^{n} \mathrm{coef}_{z^{n-2}}(f_i g_i) \end{array}$$

is a perfect pairing.

Proof. It suffices to show that

$$\begin{array}{rccc} \langle\,,\rangle_i : & k[z]/z^{n-1} \times k[z]/z^{n-1} & \longrightarrow & k \\ & f, g & \longmapsto & \mathrm{coef}_{z^{n-2}}(fg) \end{array}$$

is a perfect pairing for all i. Note that the maps

$$\mathrm{coef}_{z^j} : k[z]/z^{n-1} \longrightarrow k$$

for $j = 0, \ldots, n-2$ form a basis of $\mathrm{Hom}(k[z]/z^{n-1}, k)$. Furthermore,

$$\mathrm{coef}_{z^j}(f) = \mathrm{coef}_{z^{n-2}}(z^{n-j-2} f) = \langle z^{n-j-2}, f \rangle_i,$$

so $\langle , \rangle_i$ identifies that basis $z^0, \ldots, z^{n-2}$ of $k[z]/z^{n-1}$ with a basis of its dual. Hence the pairing is perfect. □

Lemma 4.2.18. *Let $\hat{A}^\perp$ be given by*

$$\hat{A}^\perp = \Big\{ b \in \hat{B} \;\Big|\; \forall a \in \hat{A} : \langle \hat{p}(a), b \rangle = 0 \Big\}.$$

Then the pairing gives rise to a bijection

$$\begin{array}{rcl} \hat{A}^\perp & \longrightarrow & \mathrm{Dep}(\hat{p}) \\ b & \longmapsto & \langle -, b \rangle. \end{array}$$

Proof. The pairing gives rise to a bijection

$$\begin{array}{rcl} \hat{B} & \longrightarrow & \mathrm{Hom}(\hat{B}, k) \\ b & \longmapsto & \langle -, b \rangle. \end{array}$$

The dependencies are precisely those $\phi : \hat{B} \to k$ such that $\mathrm{im}(\hat{p}) \subset \ker(\phi)$, i.e. such that $\phi(\hat{p}(a)) = 0$ for all $a \in \hat{A}$. □

Example 4.2.19. We continue our running example by computing $\mathrm{im}(\hat{p})$ and $\hat{A}^\perp$ inside $\hat{B}$ for the directions from example (1.4.3). We begin by giving the image under $\hat{p}$ of the monomials that form a basis of $\hat{A}$.

$\hat{A}$	$\hat{B}_1$			$\hat{B}_2$			$\hat{B}_3$			$\hat{B}_4$		
	1	z	z^2	1	z	z^2	1	z	z^2	1	z	z^2
1	1	0	0	1	0	0	1	0	0	1	0	0
x	0	−1	0	0	−2	0	0	−1	0	0	0	0
y	0	0	0	0	1	0	0	2	0	0	1	0
x^2	0	0	1	0	0	4	0	0	1	0	0	0
xy	0	0	0	0	0	−2	0	0	−2	0	0	0
y^2	0	0	0	0	0	1	0	0	4	0	0	1

The pairing on each $\hat{B}_i$ is given by

$$\langle a_0 + a_1 z + a_2 z^2, b_0 + b_1 z + b_2 z^2 \rangle = a_0 b_2 + a_1 b_1 + a_2 b_0.$$

The space $\hat{A}^\perp$ as defined in lemma (4.2.18) above has dimension $12 - 6 = 6$. One checks easily that the following 6 elements of $\hat{B}$ are in $\hat{A}^\perp$ and are

independent, thus they must form a basis for this space.

$\hat{B}_1$			$\hat{B}_2$			$\hat{B}_3$			$\hat{B}_4$		
1	z	z^2	1	z	z^2	1	z	z^2	1	z	z^2
3	0	0	−1	0	0	1	0	0	−3	0	0
0	−3	0	0	2	0	0	−1	0	0	0	0
0	0	0	0	−1	0	0	2	0	0	−3	0
0	0	3	0	0	−4	0	0	1	0	0	0
0	0	0	0	0	2	0	0	−2	0	0	0
0	0	0	0	0	−1	0	0	4	0	0	−3

The astute reader may notice a strong similarity between this table and the previous one. This similarity is confirmed in theorem (4.2.21) below.

The following lemma is a technical result that is needed in the proof of the subsequent theorem. It is a homogeneous analogue of Vandermonde determinants.

Lemma 4.2.20. *Let d be a positive integer. Then, over the commutative ring $R = \mathbb{Z}[\alpha_1, \ldots, \alpha_d, \beta_1, \ldots, \beta_d]$ we have*

$$\det \begin{pmatrix} \alpha_1^{d-1} & \cdots & \alpha_d^{d-1} \\ \alpha_1^{d-2}\beta_1 & & \alpha_d^{d-2}\beta_d \\ \vdots & & \vdots \\ \beta_1^{d-1} & \cdots & \beta_d^{d-1} \end{pmatrix} = \prod_{1 \le i < j \le d} (\alpha_i\beta_j - \alpha_j\beta_i).$$

Proof. The righthand side of the equation is a homogeneous polynomial of degree $d(d-1)$. Note that each of the factors $\alpha_i\beta_j - \alpha_j\beta_i$ is irreducible in R and that no two of them differ by ± 1.

Suppose $x = (a_1, \ldots, a_d, b_1, \ldots, b_d) \in \mathbb{C}^{2d}$ is a zero of the righthand side, then there are i, j such that $a_i b_j = a_j b_i$. If $a_i = b_i = 0$, then the matrix on the lefthand side has a row of zeroes when evaluated at x, so clearly x is a zero of the determinant. If $a_i \neq 0$, put $\lambda = a_j/a_i$, otherwise, if $b_i \neq 0$, put $\lambda = b_j/b_i$. Since $a_i b_j = a_j b_i$, it follows we have $a_j = \lambda a_i$ and $b_j = \lambda b_i$. Hence, for every $e = 0, \ldots, d-1$ we have

$$a_j^e b_j^{d-1-e} = \lambda^{d-1} a_i^e b_i^{d-1-e},$$

so the j-th column of the matrix on the lefthand side is λ^{d-1} times the i-th column, and therefore x is again a zero of the determinant.

Hence every zero over $\mathbb{C}$ of the righthand side is also a zero of the lefthand side, and so, by the Nullstellensatz (see e.g. [8, Thm. 1.6]), some power of the lefthand side is a multiple of the righthand side. But the righthand side is a product of distinct irreducible elements of the unique factorisation domain R, so if it divides some power of the lefthand side, it in fact divides the lefthand side itself. Moreover, the lefthand side is also homogeneous of degree $d(d-1)$, so the quotient is a constant in $\mathbb{Z}$. Considering one monomial or evaluating at one non-zero point shows that they are in fact equal. □

Theorem 4.2.21. *There is an $r \in \hat{B}$ such that there is a well-defined isomorphism*

$$\begin{array}{ccc} \hat{A} & \longrightarrow & \mathrm{Dep}(\hat{p}) \\ a & \longmapsto & \langle -, r\hat{p}(a)\rangle. \end{array}$$

Proof. By lemma (4.2.18), what we need to show is the existence of an r such that $\hat{A}^{\perp} = r \cdot \mathrm{im}(\hat{p})$.

Consider the matrix

$$M = \begin{pmatrix} (-b_1)^{n-2} & \cdots & (-b_n)^{n-2} \\ (-b_1)^{n-3}a_1 & & (-b_n)^{n-3}a_n \\ \vdots & & \vdots \\ a_1^{n-2} & \cdots & a_n^{n-2} \end{pmatrix}$$

and let M_i be the square matrix obtained from M by removing the i-th column. Put $r_i = (-1)^{i+1} \det M_i$ and let $r = (r_1, \ldots, r_n) \in \hat{B}$.

For $i, j \geq 0$ with $i + j < n - 2$ we have

$$\langle r, \hat{p}(x^i y^j)\rangle = 0$$

as all of the polynomials that appear have degree less than $n - 2$. For the other cases, we have

$$\langle r, \hat{p}(x^{n-2-i}y^i)\rangle = \sum_{j=0}^{n} (-1)^{j+1} \Big((-b_j)^{n-2-i} a_j^i \Big) \det M_i,$$

where $i \in \{0, \ldots, n-2\}$. This is the first row expansion of the determinant

of the square matrix

$$\begin{pmatrix} (-b_1)^{n-2-i}a_1^i & \cdots & (-b_n)^{n-2-i}a_n^i \\ (-b_1)^{n-2} & \cdots & (-b_n)^{n-2} \\ (-b_1)^{n-3}a_1 & & (-b_n)^{n-3}a_n \\ \vdots & & \vdots \\ a_1^{n-2} & \cdots & a_n^{n-2} \end{pmatrix}$$

obtained from M by repeating the i-th row at the top of the matrix. This determinant is 0, as the matrix has two identical rows.

It follows that $\langle r, \hat{p}(a)\rangle = 0$ for all $a \in \hat{A}$, i.e. that $r \in \hat{A}^\perp$. For $a, a' \in \hat{A}$ we have

$$\langle r\hat{p}(a), \hat{p}(a')\rangle = \sum_{i=0}^{n} \operatorname{coef}_{z^{n-2}}(r_i \hat{p}_i(a)\hat{p}_i(a')) = \langle r, \hat{p}(aa')\rangle = 0,$$

so $r \cdot \operatorname{im}(\hat{p}) \subset \hat{A}^\perp$.

Note that by lemma (4.2.20),

$$r_i = (-1)^{i+1} \det M_i = (-1)^{i+1} \prod (a_s b_t - a_t b_s)$$

where the product runs over all $1 \leq s < t \leq n$ such that $s, t \neq i$. Hence r_i is non-zero, as the directions are pairwise independent. It follows that multiplication by r is a bijection $\hat{B} \to \hat{B}$. Finally, we compare dimensions

$$\begin{aligned} \dim \hat{A}^\perp &= \dim \hat{B} - \dim(\operatorname{im}(\hat{p})) \\ &= n(n-1) - \tfrac{1}{2}n(n-1) = \tfrac{1}{2}n(n-1) \\ &= \dim(r \cdot \operatorname{im}(\hat{p})) \end{aligned}$$

and conclude that $r \cdot \operatorname{im}(\hat{p}) = \hat{A}^\perp$. □

Example 4.2.22. Continuing our running example, we'll compute the vector r for the four directions from example (1.4.3). Recall that these directions are $d_1 = (0, 1)$, $d_2 = (1, 2)$, $d_3 = (2, 1)$ and $d_4 = (1, 0)$. Hence the coefficients

r_i are given by

$$\begin{aligned}
r_1 &= (-1)^2(1\cdot 1 - 2\cdot 2)(1\cdot 0 - 2\cdot 1)(2\cdot 0 - 1\cdot 1)\\
&= 1\cdot -3\cdot -2\cdot -1\\
&= -6,\\
r_2 &= (-1)^3(0\cdot 1 - 1\cdot 2)(0\cdot 0 - 1\cdot 1)(2\cdot 0 - 1\cdot 1)\\
&= -1\cdot -2\cdot -1\cdot -1\\
&= 2,\\
r_3 &= (-1)^4(0\cdot 2 - 1\cdot 1)(0\cdot 0 - 1\cdot 1)(1\cdot 0 - 2\cdot 1)\\
&= 1\cdot -1\cdot -1\cdot -2\\
&= -2,\\
r_4 &= (-1)^5(0\cdot 2 - 1\cdot 1)(0\cdot 1 - 1\cdot 2)(1\cdot 1 - 2\cdot 2)\\
&= -1\cdot -1\cdot -2\cdot -3\\
&= 6.
\end{aligned}$$

Hence we have $r = (-6, 2, -2, 6)$. Looking back at the tables in example (4.2.19), one sees that the rows of the second table (for $\hat{A}^{\perp}$) are those of the first table (for $\mathrm{im}(\hat{p})$) multiplied by $-\frac{1}{2}r$.

Putting it all together

We now have all the results at our disposal to prove the theorem that was announced at the start of the section.

Proof of (4.2.2). Let (m_x, m_y), s, t and S as in the statement of the theorem. Let $D_S = (d_i | i \in S)$ and let p_S be the reconstruction system corresponding to D_S defined in lemma (4.2.4). Then we have

$$p_{(m_x,m_y)} \cong p_{S,(m_x,m_y)}$$

by lemma (4.2.10) and

$$p_{S,(m_x,m_y)} \cong p_{S,(0,0)}$$

by lemma (4.2.11) and

$$p_{S,(0,0)} \cong \hat{p}_S$$

by theorem (4.2.15).

Let $\delta_S : \hat{B}_S \to k$ be the dependency given by

$$f \mapsto \langle b, r\hat{p}(x^s y^t)\rangle,$$

with r as in theorem (4.2.21). From the proof of that theorem we know that $r = (r_i)_{i \in S}$ with the r_i as in the statement of this theorem. Via the isomorphisms mentioned above, δ_S induces a dependency $\delta : B_{(m_x,m_y)} \to k$.

If $i \notin S$, then w^m in the i-th component of $B_{(m_x,m_y)}$ maps to 0 in $B_{S,(m_x,m_y)}$, so the i-th component of δ is 0. If $i \in S$, then w^m in the i-th component of $B_{(m_x,m_y)}$ maps to w^m in $B_{S,(m_x,m_y)}$. This maps to $(\zeta_i w)^m$ in $B_{S,(0,0)}$, with ζ_i as in the theorem. Applying log, this maps to $\zeta_i^m e^{mz}$. Finally, applying δ_S, we get

$$\langle \zeta_i^m e^{mz}, r_i \hat{p}_i(x^s y^t) \rangle_i$$

as the righthand side is a monomial of degree $s + t$, this evaluates to

$$\frac{\zeta_i^m m^d}{d!} r_i (-b_i)^s a_i^t.$$

Running over all s and t, $x^s y^t$ runs over a basis for $\hat{A}_S$ and therefore the δ_S's run over a basis of $\mathrm{Dep}(\hat{p}_S)$. Hence the δ's run over a basis of $\mathrm{Dep}(p_{(m_x,m_y)})$. Running over all (m_x, m_y) runs over all the local factors and so produces a basis for $\mathrm{Dep}(p_D) = \mathrm{Dep}(p)$ as claimed. □

Example 4.2.23. To conclude this section, we give the basis for the dependencies over $\overline{\mathbb{Q}}$ of our running example, the full grid reconstruction system from example (1.4.3), viewed as a map

$$p_D : k[u, u^{-1}, v, v^{-1}] \longrightarrow k[w, w^{-1}]^4$$

as described in remark (4.1.1). We have computed M in example (4.2.7). The points (x, y) that are admissible in theorem (4.2.2) are all from M, as we must have $\#S \geq 2$ for these points. The table below shows for every point in M the corresponding set S and the corresponding (r_i) and (ζ_i). For convenience of notation, let $\mu = \zeta(1/3)$.

M	S	(r_i)	(ζ_i)
$(0,0)$	$\{1,2,3,4\}$	$(-6,\ 2,-2,\ 6)$	$(\ 1,\ 1,\ 1,\ 1)$
$(0,1/2)$	$\{2,4\}$	$(\ 1,\ -1)$	$(\ -1,\ -1)$
$(1/2,0)$	$\{1,3\}$	$(\ 1,\ -1\ )$	$(-1,\ -1\ )$
$(1/3,1/3)$	$\{2,3\}$	$(\ 1,-1\ )$	$(\ \mu\ ,\ \mu^2\ )$
$(2/3,2/3)$	$\{2,3\}$	$(\ 1,-1\ )$	$(\ \mu^2,\ \mu\ )$

The table below lists the admissible (x, y), s and t from theorem (4.2.2) and the coefficient vectors of the corresponding dependency. What is shown in the table is the weight of the line corresponding to w^m in each of the four directions.

M	s	t	$(0,1)$	$(1,2)$	$(2,1)$	$(1,0)$
$(0,0)$	0	0	$-3m^2$	m^2	$-m^2$	$3m^2$
	1	0	$6m$	$-4m$	$2m$	0
	0	1	0	$2m$	$-4m$	$6m$
	2	0	-6	8	-2	0
	1	1	0	-4	4	0
	0	2	0	2	-8	6
$(0,1/2)$	0	0	0	$(-1)^m$	0	$(-1)^{m+1}$
$(1/2,0)$	0	0	$(-1)^m$	0	$(-1)^{m+1}$	0
$(1/3,1/3)$	0	0	0	μ^m	$-\mu^{2m}$	0
$(2/3,2/3)$	0	0	0	μ^{2m}	$-\mu^m$	0

We can verify that these are indeed dependencies by computing for each one the image under the dependency of

$$p_D(u^a v^b) = (w^{-a}, w^{b-2a}, w^{2b-a}, w^b)$$

for all $(a, b) \in \mathbb{Z}^2$. For example, the value for the first dependency will be

$$-3(-a)^2 + (b-2a)^2 - (2b-a)^2 + 3(b)^2,$$

which expands to 0, and similarly for the other dependencies.

4.3 Rational and integral dependencies

We continue our study of the dependencies for grid reconstruction systems associated to $(\mathbb{Z}^2, D)$, where D is a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. In this section we use our understanding of these dependencies over $\overline{\mathbb{Q}}$ to give a description of the dependencies over $\mathbb{Q}$ and $\mathbb{Z}$.

For the rational dependencies, we will prove a result similar to theorem (4.2.2). We do not state it here, as some extra terminology is needed for the statement. It is theorem (4.3.8) below. To go from rational to integral dependencies we will use the following result. An algorithmic version of this result is given as a corollary.

Theorem 4.3.1. *Let D be a 2-regular sequence of primitive dependencies in $\mathbb{Z}^2$. Let $p_{\mathbb{Q}}$ be the grid reconstruction system over $\mathbb{Q}$ associated to $(\mathbb{Z}^2, D)$ and $p_{\mathbb{Z}}$ the corresponding system over $\mathbb{Z}$. Let $r_1, \ldots, r_n$ be as in theorem (3.5.5). Let ϕ be the linear map*

$$\mathrm{Dep}(p_{\mathbb{Q}}) \longrightarrow \mathbb{Q}^{r_1-1} \oplus \cdots \oplus \mathbb{Q}^{r_n-1}$$

sending a dependency d to

$$((c_1(1), \ldots, c_1(r_1-1)), \ldots, (c_n(1), \ldots, c_n(r_n-1))),$$

where $c_1, \ldots, c_n$ are the coefficient sequences associated to d as in remark (3.5.3). Then $\mathrm{Dep}(p_{\mathbb{Z}})$ as a subspace of $\mathrm{Dep}(p_{\mathbb{Q}})$ is given by

$$\mathrm{Dep}(p_{\mathbb{Z}}) = \phi^{-1}\Big(\mathbb{Z}^{r_1-1} \oplus \cdots \oplus \mathbb{Z}^{r_n-1}\Big).$$

Corollary 4.3.2. *There is an algorithm that on input a 2-regular sequence of primitive directions D in $\mathbb{Z}^2$ gives a basis for $\mathrm{Dep}(p_{\mathbb{Z}})$, where $p_{\mathbb{Z}}$ is the grid reconstruction system over $\mathbb{Z}$ associated to $(\mathbb{Z}^2, D)$. The runtime of this algorithm is polynomial in the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$ and the length of the input. The output is given as a sequence of vectors in*

$$\mathbb{Z}^{r_1-1} \oplus \cdots \oplus \mathbb{Z}^{r_n-1},$$

each vector representing the coefficients $c_1(1), \ldots, c_n(r_n-1)$ associated to the corresponding dependency.

Note that this is precisely the information we need in the consistency algorithm from corollary (3.5.7). The runtime of the algorithm depends on the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$ and this rank can grow exponentially in the length of the input D. Already when we have the two directions $(1, 0)$ and $(1, a)$ the length of the input is of size $\log(|a|)$, but the rank is $|1 \cdot a - 0 \cdot 1| = |a|$.

Before we can state our analogue of theorem (4.2.2), we have to look a bit closer at the set of intersection points M associated to D in theorem (4.2.6).

Lemma 4.3.3. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$ and write $d_i = (a_i, b_i)$ for $i = 1, \ldots, n$. Let S be a subset of $\{1, \ldots, n\}$ that has at least 2 elements. Put*

$$M_S = \{(x, y) \in (\mathbb{Q}/\mathbb{Z})^2 \mid \forall i \in S : a_i x + b_i y \in \mathbb{Z}\}.$$

Then M_S is a finite cyclic subgroup of $(\mathbb{Q}/\mathbb{Z})^2$.

Proof. It is clear that M_S is a subgroup of $(\mathbb{Q}/\mathbb{Z})^2$. Hence it suffices to show that if $\#S = 2$ it is finite cyclic, as any subgroup of a finite cyclic subgroup is again finite cyclic. So let $S = \{i, j\}$ with $i \neq j$.

Next, observe that $(x, y) \in (\mathbb{Q}/\mathbb{Z})^2$ satisfies $a_i x + b_i y \in \mathbb{Z}$ if and only if $(x, y) = (-b_i\lambda, a_i\lambda)$ for some $\lambda \in \mathbb{Q}$. The 'if' part is clear. For the 'only if', assume $a_i x + b_i y = \mathbb{Z}$. Choose lifts $\tilde{x}$ and $\tilde{y}$ of x, y to $\mathbb{Q}$. Then we have $a_i\tilde{x} + b_i\tilde{y} = n \in \mathbb{Z}$. Let $c_i, d_i \in \mathbb{Z}$ be such that $a_i d_i - b_i c_i = 1$.

Then $(x, y) = (\tilde{x} - nd_i, \tilde{y} + nc_i)$ holds in $(\mathbb{Q}/\mathbb{Z})^2$ and in $\mathbb{Q}$ we have

$$a_i(\tilde{x} - nd_i) + b_i(\tilde{y} + nc_i) = a_i x + b_i y - n(a_i d_i - b_i c_i) = n - n = 0.$$

It follows that $(\tilde{x} - nd_i, \tilde{y} + nc_i) = \lambda(-b_i, a_i)$ holds in $\mathbb{Q}^2$ for some $\lambda \in \mathbb{Q}$.

From the above it follows that $(x, y) \in M_S$ if and only if

$$(x, y) = (-\lambda b_i, \lambda a_i)$$

for some $\lambda \in \mathbb{Q}$ which also satisfies $a_j x - b_j y \in \mathbb{Z}$, i.e. $\lambda(-b_i a_j + a_i b_j) \in \mathbb{Z}$. It is clear that the λ's which satisfy this form a finite cyclic group. □

Lemma 4.3.4. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$ and write $d_i = (a_i, b_i)$ for $i = 1, \ldots, n$. Let $S \subset \{1, \ldots, n\}$ have $\#S \geq 2$. Then there is an $(x, y) \in M_S$ such that $S = \{i | a_i x + b_i y \in \mathbb{Z}\}$ if and only if for every T that strictly contains S, M_T is a proper subset of S.*

Proof. If such an (x, y) exists, then it clearly is not in M_T for every T that strictly contains S, so all these M_T are proper subsets of M_S. Conversely, suppose that M_T is a proper subset for all T that strictly contain S. Then it is a proper subgroup, so it doesn't contain any generator of M_S. Hence for any generator $(x, y) \in M_S$, there is no T that strictly contains S such that $(x, y) \in M_T$. Therefore $\{i | a_i x + b_i y \in \mathbb{Z}\} = M_S$. □

Definition 4.3.5. Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$ and write $d_i = (a_i, b_i)$ for $i = 1, \ldots, n$. If S satisfies the equivalent conditions of the previous lemma, we call S an *intersection type* for D.

Example 4.3.6. We consider once again the directions from example (1.4.3):

$$d_1 = (0, 1), \quad d_2 = (1, 2), \quad d_3 = (2, 1), \quad d_4 = (1, 0).$$

In example (4.2.7) we computed

$$M = \left\{(0,0), \left(0,\frac{1}{2}\right), \left(\frac{1}{2},0\right), \left(\frac{1}{3},\frac{1}{3}\right), \left(\frac{2}{3},\frac{2}{3}\right)\right\}.$$

In example (4.2.23) we computed for each element (x, y) of M the set

$$S = \{i \mid a_i x + b_i y \in \mathbb{Z}\}.$$

By definition, these sets correspond to the intersection types. There are four and they are listed in the table below.

S	M_S
$\{1,2,3,4\}$	$\{(0,0)\}$
$\{1,3\}$	$\{(0,0),(1/2,0)\}$
$\{2,3\}$	$\{(0,0),(1/3,1/3),(2/3,2/3)\}$
$\{2,4\}$	$\{(0,0),(0,1/2)\}$

Example 4.3.7. We will show that like the number of dependencies, the number of intersection types can grow exponentially in the size of D.

Let $n \geq 1$. We will produce a 2-regular sequence D of primitive directions $d_1, \ldots, d_{n+1}$ whose length is bounded by a polynomial in n, but such that there are at least 2^n intersection types.

Let $p_1, \ldots, p_n$ be the first n prime numbers. It is a well-known fact from elementary number theory that p_n is bounded above by by $an\log(n)$ for some positive constant a independent of n (see. e.g. [16, Thm. 113]). For $i = 1, \ldots, n$, let $b_i = \prod_{j \neq i} p_j$ and let $d_i = (1, b_i)$. Finally, let $d_{n+1} = (1, 0)$. Note that $\log(b_i)$ is bounded by some constant times $n\log(n)$, so that the length of D is bounded by some constant times $n^2\log(n)$.

Let $S \subset \{1, \ldots n\}$ and put $y = \prod_{i \notin S} p_i^{-1}$. Then the point $(0, y)$ lies on the line corresponding to d_{n+1} and on the line corresponding to d_i for all $i \in S$. But it does not lie on the line corresponding to d_i for any $i \notin S$. Thus $S \cup \{n+1\}$ is an intersection type, for every subset S of $\{1, \ldots, n\}$. Hence there are at least 2^n intersection types.

We can now state the theorem analogous to theorem (4.2.2) for dependencies over $\mathbb{Q}$. Note that it doesn't produce a basis for the space of dependencies,

but a generating set. We will see in the proof of corollary (4.3.2) how much larger this generating set can be compared to a basis for the space.

Theorem 4.3.8. *Let $D = (d_1, \ldots, d_n)$ be a 2-regular sequence of primitive directions in $\mathbb{Z}^2$ and let p_D be the grid reconstruction system over $\mathbb{Q}$ associated to $(\mathbb{Z}^2, D)$. Let S be an intersection type and $(m_x, m_y) \in (\mathbb{Q}/\mathbb{Z})^2$ a generator of M_S. Let $s, t \in \mathbb{Z}_{\geq 0}$, put $d = \#S - 2 - s - t$ and suppose that $d \geq 0$. Let $g \in (\mathbb{Q}/\mathbb{Z})^S$ be given by $g_i = a_i' m_x + b_i' m_y$ for $i \in S$, where $a_i', b_i' \in \mathbb{Z}$ are such that $a_i b_i' - b_i a_i' = 1$. Let $h \in \frac{1}{\#\langle g \rangle}\mathbb{Z}/\mathbb{Z}$. Let the $\mathbb{Q}$-linear map*

$$\delta : \mathbb{Q}[w, w^{-1}]^n \longrightarrow \mathbb{Q}$$

be given on the i-th component by

$$\begin{array}{rcll} w^m & \longmapsto & \dfrac{r_i(-b_i)^s a_i^t m^d}{d!} & \text{if } i \in S \text{ and } h - mg_i \in \mathbb{Z} \\ & \longmapsto & 0 & \text{otherwise,} \end{array}$$

where

$$r_i = (-1)^{\#\{j \in S \mid j \leq i\}} \prod_{j_1, j_2} (a_{j_1} b_{j_2} - a_{j_2} b_{j_1}),$$

with the product running over $j_1, j_2 \in S \setminus \{i\}$ such that $j_1 < j_2$. Then $\delta \circ p_D = 0$, so δ corresponds to a dependency in $\mathrm{Dep}(p)$. *Moreover, running over all S, s, t and h produces a generating set for this space.*

For the rest of this section, fix a 2-regular sequence of primitive directions $D = (d_1, \ldots, d_n)$ in $\mathbb{Z}^2$ and write $d_i = (a_i, b_i)$ for $i = 1, \ldots, n$. Let

$$p : A \to B$$

be as in the previous section, the reconstruction system associated to D defined in lemma (4.2.4) for $k = \overline{\mathbb{Q}}$. Let M be as in theorem (4.2.6).

Remark 4.3.9. Let $p_{\mathbb{Q}} : A_{\mathbb{Q}} \to B_{\mathbb{Q}}$ be the reconstruction system associated to D defined in lemma (4.2.4) for $k = \mathbb{Q}$. Note that p can be seen as the base change $\mathrm{BC}_{\overline{\mathbb{Q}}/\mathbb{Q}}(p_{\mathbb{Q}})$. Recall from lemma (4.2.4) that $\mathrm{cok}(p_{\mathbb{Q}})$ is the same as the cokernel of the grid reconstruction system over $\mathbb{Q}$ associated to $(\mathbb{Z}^2, D)$. Hence, by theorem (3.5.1) it is free of finite rank. The upshot of this is we can apply lemma (2.5.15) to conclude that there is a natural inclusion $\mathrm{Dep}(p_{\mathbb{Q}}) \longrightarrow \mathrm{Dep}(p)$ and a natural action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $\mathrm{Dep}(p)$ such that $\mathrm{Dep}(p_{\mathbb{Q}})$ is the space of fixed points of that action.

The strategy to compute the dependencies for $p_{\mathbb{Q}}$ is to study the Galois action on the dependencies for p. In the previous section, we decomposed p as a sum of local factors. One might hope that the Galois group acts on each factor individually, but in general, it doesn't. This is where the intersection types come in. We will show that an element of the Galois group permutes the local factors corresponding to each intersection type. So instead of looking at the local factors individually, we have to look at all the local factors corresponding to an intersection type. We can then use the explicit basis for the dependencies over $\overline{\mathbb{Q}}$ for these local factors to see how the Galois group acts and what the invariant elements are.

Example 4.3.10. Consider the reconstruction system over $\overline{\mathbb{Q}}$ from example (1.4.3). We computed in example (4.2.23) a basis for the dependencies of this system. Let $\mu = \zeta(1/3)$ and consider the dependency δ corresponding to the point $(2/3, 2/3) \in M$ as described in that example. The weight this dependency assigns to the m-th line in each direction is given in the following table.

direction	$(0,1)$	$(1,2)$	$(2,1)$	$(1,0)$
weight	0	μ^{2m}	$-\mu^m$	0

Recall that we have fixed an inclusion $\overline{\mathbb{Q}} \subset \mathbb{C}$ in order to define ζ. Let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/Q)$ be the automorphism induced by complex conjugation. Note that $\sigma(\mu) = \mu^2$. It follows as in example (2.5.17) that $\sigma\delta$ is the dependency that assigns the following weights to the m-th line in each direction.

direction	$(0,1)$	$(1,2)$	$(2,1)$	$(1,0)$
weight	0	μ^m	$-\mu^{2m}$	0

Note that this is precisely the dependency we computed in example (4.2.23) corresponding to the point $(1/3, 1/3) \in M$.

Lemma 4.3.11. *Identify A with $\bigoplus_{(x,y)\in M} A_{(x,y)}$ as in lemma (4.2.8). Then $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on M and for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and every $(x, y) \in M$ we have $\sigma(A_{(x,y)}) = A_{\sigma(x,y)}$ as subspaces of A. Moreover, we can extend σ to an*

isomorphism of reconstruction systems $p_{(x,y)} \to p_{\sigma(x,y)}$ *such that*

$$\begin{array}{ccc} p & \twoheadrightarrow & p_{(x,y)} \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\sigma} \\ p & \twoheadrightarrow & p_{\sigma(x,y)} \end{array}$$

is a commutative diagram of reconstruction systems.

Proof. By definition M is a subset of $(\mathbb{Q}/\mathbb{Z})^2$. This space is identified with the torsion subgroup of $(\overline{\mathbb{Q}}^\times)^2$ through ζ. The Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on this torsion subgroup and hence acts on M. Furthermore, the bijection from theorem (4.2.6) identifies M with the set of maximal ideals of A. This set also has an action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and the bijection is compatible with the Galois actions.

Write ϕ for the isomorphism

$$\begin{array}{rcl} \phi : A & \longrightarrow & \bigoplus_{(x,y)\in M} A_{(x,y)} \\ a & \mapsto & (a, \ldots, a). \end{array}$$

For $(x, y) \in M$, we want to identify the $a \in A$ such that all coordinates of $\phi(a)$ are zero except for the coordinate in $A_{(x,y)}$.

Consider the set

$$B = \{(\zeta(x) - \zeta(x'), \zeta(y) - \zeta(y')) \mid (x, y), (x', y') \in M, (x, y) \neq (x', y')\}.$$

Note that B is a finite subset of $\overline{\mathbb{Q}}^2$ that does not contain $(0, 0)$. Hence there are $a, b \in \mathbb{Q}$ such that

$$\forall(\alpha, \beta) \in B : a\alpha + b\beta \neq 0.$$

For $(x, y) \in M$, let

$$g_{(x,y)} = a(u - \zeta(x)) + b(v - \zeta(y)) \in A$$

and let

$$G_{(x,y)} = \prod_{(x',y')\neq(x,y)} g_{(x',y')}.$$

Note that $g_{(x,y)}$ is in the maximal ideal corresponding to (x, y) and not in any other maximal ideal, by construction of a and b.

Hence, for every $(x, y) \in M$, $G_{(x,y)}$ is in the maximal ideal corresponding to (x', y') for every $(x', y') \neq (x, y)$, but not in the maximal ideal corresponding to (x, y). Note that for every maximal ideal $\mathfrak{m}$ of A, some power of $\mathfrak{m}$ is zero in $A_{\mathfrak{m}}$, as $A_{\mathfrak{m}}$ is a finite-dimensional $\overline{\mathbb{Q}}$-vector space. Hence any sufficiently large power of $G_{(x,y)}$ maps under ϕ to a unit in $A_{(x,y)}$ and to 0 in $A_{(x',y')}$ for all $(x', y') \neq (x, y)$. Moreover, since M is finite, we can choose a positive integer e such that this is true for $G^e_{(x,y)}$ for every (x, y).

Let $(x, y) \in M$, $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and $a \in A_{(x,y)}$. Then there is a $\tilde{a} \in A$ such that the image of $\tilde{a}G^e_{(x,y)}$ in $A_{(x,y)}$ is a, as the map $A \to A_{(x,y)}$ is surjective and $G_{(x,y)}$ maps to a unit under this map. Then $\phi(\tilde{a}G^e_{(x,y)})$ maps to the element in $\bigoplus_{(x',y') \in M} A_{(x',y')}$ that is a in $A_{(x,y)}$ and zero everywhere else, i.e., when we view $A_{(x,y)}$ as a subset of this sum, $\phi(\tilde{a}G^e_{(x,y)})$ corresponds to the element a. It follows that

$$\sigma(a) = \sigma(\phi(\tilde{a}G^e_{(x,y)})) = \phi(\sigma(\tilde{a})\sigma(G_{(x,y)})^e).$$

Note that for all $(x', y') \in M$, $\sigma(g_{(x',y')}) = g_{\sigma(x',y')}$, so $\sigma(G_{(x,y)}) = G_{\sigma(x,y)}$. We conclude that $\sigma(a)$ is in $A_{\sigma(x,y)}$ and that we have a commutative diagram

$$\begin{array}{ccc} A & \twoheadrightarrow & A_{(x,y)} \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\sigma} \\ A & \twoheadrightarrow & A_{\sigma(x,y)} \end{array}$$

The map $\sigma : A_{(x,y)} \to A_{\sigma(x,y)}$ is bijective as the map $A_{\sigma(x,y)} \to A_{(x,y)}$ induced by σ^{-1} is its inverse.

Recall that $B_{(x,y)}$ is defined as $A_{(x,y)} \otimes_A B$. Note that σ also induces an isomorphism

$$\begin{array}{ccc} A_{(x,y)} \otimes_A B & \longrightarrow & A_{\sigma(x,y)} \otimes_A B \\ a \otimes b & \mapsto & \sigma(a) \otimes \sigma(b). \end{array}$$

A straightforward verification shows that the diagram

$$\begin{array}{ccc} A_{(x,y)} & \xrightarrow{p_{(x,y)}} & A_{(x,y)} \otimes_A B \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\sigma} \\ A_{\sigma(x,y)} & \xrightarrow[p_{\sigma(x,y)}]{} & A_{\sigma(x,y)\otimes_A B} \end{array}$$

is commutative, i.e., σ induces an isomorphism of reconstruction systems $p_{(x,y)} \to p_{\sigma(x,y)}$. Another straightforward verification shows that it fits in a commutative diagram with p as claimed in the lemma. □

For an intersection type S, let D_S be the sub-sequence $(d_i | i \in S)$ of D and let $p_S : A_S \to B_S$ be the reconstruction system over $\overline{\mathbb{Q}}$ associated to D_S in lemma (4.2.4). Note that there is a natural surjective morphism $p_{\overline{\mathbb{Q}}} \to p_S$ for every S that is given by forgetting the coordinates that are not in S. Hence there is a natural inclusion $\mathrm{Dep}(p_S) \subset \mathrm{Dep}(p_{\overline{\mathbb{Q}}})$. Let

$$V_S = \bigoplus_{(x,y) \in M_S} \mathrm{Dep}(p_{S,(x,y)}) \subset \mathrm{Dep}(p_S).$$

Lemma 4.3.12. *Let S be an intersection type. Then $\mathrm{Dep}(p_S)$ and V_S, viewed as subspaces of $\mathrm{Dep}(p_{\overline{\mathbb{Q}}})$ are stable under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, that is, for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ one has $\sigma(\mathrm{Dep}(p_S)) \subset \mathrm{Dep}(p_S)$ and $\sigma(V_S) \subset V_S$.*

Proof. Let $p_{S,\mathbb{Q}}$ be the reconstruction system over $\mathbb{Q}$ associated to D_S in lemma (4.2.4). Then p_S is the base change $\mathrm{BC}_{\overline{\mathbb{Q}}/\mathbb{Q}}(p_{S,\mathbb{Q}})$. It is clear that the natural map $p \to p_S$ is the base change of the equivalent map $p_{\mathbb{Q}} \to p_{S,\mathbb{Q}}$ and that the Galois actions on p_S and p are compatible. It follows that $\mathrm{Dep}(p_S) \subset \mathrm{Dep}(p)$ is stable under the Galois action.

Let M' be the set of intersection points corresponding to D_S. Note that it is a subset of M that contains M_S, but is in general not equal to M_S. As mentioned in the proof of lemma (4.3.11), the Galois action on A_S induces an action on M'. We will now look at this action in a bit more detail. Let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

As M' is a finite set of pairs $(x, y) \in (\mathbb{Q}/\mathbb{Z})^2$, there is an element $d \in \mathbb{Q}/\mathbb{Z}$ such that for every $(x, y) \in M'$, x and y are both multiples of d. Such a d can be obtained for example by multiplying together the denominators of all x's and y's. Note that $\zeta(d)$ is a root of the polynomial $X^d - 1$, and therefore, so is $\sigma(\zeta(d))$. Hence, $\sigma(\zeta(d)) = \zeta(td)$ for some positive integer t. Moreover, it follows that for all $(x, y) \in M'$ we have $\sigma(\zeta(x)) = \zeta(tx)$ and $\sigma(\zeta(y)) = \zeta(ty)$. We conclude that σ acts on M' by sending (x, y) to (tx, ty).

In particular, the cyclic subgroup $M_S \subset M'$ is mapped to itself by every σ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. It follows then from lemma (4.3.11) that

$$q = \bigoplus_{(x,y) \in M_S} p_{S,(x,y)}$$

has a Galois action and that the natural morphism $p_S \to q$ is compatible with the Galois actions. It follows that $V_S = \mathrm{Dep}(q)$ is stable under the Galois action on $\mathrm{Dep}(p_S)$. □

Example 4.3.13. We consider once again the reconstruction system over $\overline{\mathbb{Q}}$ associated to the directions from example (1.4.3). We computed the intersection types for these directions in example (4.3.6). Consider the intersection type $S = \{2, 3\}$. Note that

$$M_S = \{(0, 0), (1/3, 1/3), (2/3, 2/3)\}$$

is in fact the full set of intersection points corresponding to the full grid reconstruction system p_S. It follows that $V_S = \mathrm{Dep}(p_S)$ in this case.

Let $\mu = \zeta(1/3)$ be a primitive third root of unity. Using the computations from example (4.2.23), we see that V_S has dimension 3 and a basis is given in the following table.

	$(0,1)$	$(1,2)$	$(2,1)$	$(1,0)$
v_1	0	1	-1	0
v_2	0	μ^m	$-\mu^{2m}$	0
v_3	0	μ^{2m}	$-\mu^m$	0

Note that any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ will send μ to another primitive third root of unity, so $\sigma(\mu) = \mu$ or $\sigma(\mu) = \mu^2$. In the former case σ acts trivially on v_1, v_2 and v_3. In the latter case, as we saw in example (4.3.10), σ maps v_3 to v_2. One computes easily that it in fact interchanges v_2 and v_3 and maps v_1 to itself. It follows that V_S is indeed invariant under the action of the Galois group.

The strategy for the proof of theorem (4.3.8) is to show that the basis for $\mathrm{Dep}(p)$ we produced in theorem (4.2.2) is completely contained in the union $\bigcup_S V_S$ for all intersection types S. Within each V_S we then still have to find the elements that are fixed by the Galois group. To that end, we have the following lemma.

Lemma 4.3.14. *Let $x = (x_1, \ldots, x_d) \in (\mathbb{Q}/\mathbb{Z})^d$ and let V be the sub-$\overline{\mathbb{Q}}$-vector space of $(\mathrm{Map}(\mathbb{Z}, \overline{\mathbb{Q}}))^d$ generated by*

$$v_s = \Big(j \mapsto \zeta(jsx_i)\Big)_{i=1}^{d} \quad \textit{for } s \in \mathbb{Z}/\#\langle x\rangle\mathbb{Z}.$$

Then V is also generated by

$$w_t = \left(j \mapsto \begin{cases} 1 & \text{if } t - jx_i \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \right)_{i=1}^{d}$$

for $t \in \frac{1}{\#\langle x \rangle}\mathbb{Z}/\mathbb{Z}$. It follows that V is stable under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and $V^{\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})}$ is the $\mathbb{Q}$-vector space generated by the w_t with $t \in \frac{1}{\#\langle x \rangle}\mathbb{Z}/\mathbb{Z}$.

Proof. We have to show that each w_t can be written as a $\overline{\mathbb{Q}}$-linear combination of the v_s, and conversely, that each v_s can be written as a $\overline{\mathbb{Q}}$-linear combination of the w_t. It follows that the v_s and the w_t span the same $\overline{\mathbb{Q}}$-vector space V. Each of the w_t is fixed by the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $(\mathrm{Map}(\mathbb{Z}, \overline{\mathbb{Q}}))^d$, so V is stable under the Galois action.

Note that given i and j, the i-th coefficient of $w_t(j)$ is non-zero for at most one t. Hence a $\overline{\mathbb{Q}}$-linear combination of the w_t's is fixed by the action of the Galois group if and only if it is actually a $\mathbb{Q}$-linear combination. In other words, $V^{\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})}$ is the $\mathbb{Q}$-span of the w_t.

Put $e = \#\langle x \rangle$. We will now verify that the following relations hold between the w_t and v_s. For all $s \in \mathbb{Z}/e\mathbb{Z}$ we have

$$v_s = \sum_{t \in \frac{1}{e}\mathbb{Z}/\mathbb{Z}} \zeta(st) w_t$$

and for all $t \in \frac{1}{e}\mathbb{Z}/\mathbb{Z}$ we have

$$w_t = \frac{1}{e} \sum_{s \in \mathbb{Z}/e\mathbb{Z}} \zeta(-st) v_s.$$

For the first equation, consider the value of the i-th component of the right-hand side at $j \in \mathbb{Z}$. Note that $ex = 0$ in $(\mathbb{Q}/\mathbb{Z})^d$, so the denominator of x_i is a divisor of e. It follows that there is precisely one $t \in \frac{1}{e}\mathbb{Z}/\mathbb{Z}$ such that $t - jx_i$ is an integer. Hence the sum on the right-hand side of the first equation collapses to the single value $\zeta(st)$. Moreover, since $t - jx_i$ is an integer, $\zeta(st) = \zeta(jsx_i)$, which is precisely the value of the i-th component of v_s at j. Hence the first equation holds.

For the second equation, we also consider the value of the i-th component of the right-hand side at $j \in \mathbb{Z}$. It is given by

$$\frac{1}{e} \sum_{s \in \mathbb{Z}/e\mathbb{Z}} \zeta(-st)\zeta(jsx_i) = \frac{1}{e} \sum_{s \in \mathbb{Z}/e\mathbb{Z}} \zeta((t - jx_i)s)$$

There are now two possibilities.
If $t - jx_i$ is an integer, each of the summands is 1 and so the value of the right-hand side is $\frac{1}{e} \cdot e = 1$, which is also the value of the i-th component of w_t at j.
If $t - jx_i$ is not an integer, let f be its denominator. Note that f is a divisor of e, so every f-th root of unity comes up $\frac{e}{f}$ times in the sum. But

$$1 + \zeta(\frac{1}{f}) + \zeta(\frac{2}{f}) + \cdots + \zeta(\frac{f-1}{f}) = 0,$$

so the value of the right-hand side is 0, which is the value of the i-th component of w_t at j.
We conclude that the second equation holds as well. □

Proof of (4.3.8). Let D and p_D be as in the theorem. Let p, $p_{\mathbb{Q}}$, etc. be as in the rest of the section. From lemma (4.2.4) we know that $\mathrm{cok}(p_D)$ and $\mathrm{cok}(p_{\mathbb{Q}})$ are equal. So to compute the dependencies for p_D, we can also look at $p_{\mathbb{Q}}$.

Let S be an intersection type. We will show below that running over all admissible s, t and h as in the theorem produces a $\overline{\mathbb{Q}}$-basis of V_S consisting of elements that are fixed under the Galois action. We first show why running over all intersection types S, this will produce a generating set for the rational dependencies. It is clear that it produces a set of rational dependencies, so it suffices to show that this set contains a basis over $\overline{\mathbb{Q}}$ of the space.

Note that for every (m_x, m_y) in M, (m_x, m_y) is a generator of M_S, where S is the intersection type $S = \{i \mid a_i m_x + b_i m_y \in \mathbb{Z}\}$. So we have

$$p_{(m_x,m_y)} = p_{S,(m_x,m_y)}$$

by lemma (4.2.10). Hence $\mathrm{Dep}(p_{(m_x,m_y)})$ is contained in V_S. In particular, if we run over all intersection types, the union $\bigcup_S V_S$ will contain all the dependencies from theorem (4.2.2), which form a basis for $\mathrm{Dep}(p)$.

Hence we need to show that for an intersection type S, running over all admissible s, t and h as in the theorem produces a $\overline{\mathbb{Q}}$-basis of V_S consisting of elements that are fixed under the Galois action. Let (m_x, m_y) be a generator of M_S and for $i \in S$ let a'_i, b'_i and g_i be as in the statement of the theorem.

First we show that $\#\langle g \rangle$ is equal to $\#M_S$. We can write (m_x, m_y) as $(X/Z, Y/Z)$ where X, Y and Z are integers and the greatest common divisor of $\{X, Y, Z\}$ is 1. Note that we then have $\#M_S = Z$. Moreover, for every i in S we have

$$g_i = a'_i x + b'_i y = \frac{a'_i X + b'_i Y}{Z}.$$

Suppose that $p|a_i'X + b_i'Y$ and $p|Z$. As $a_ix + b_iy \in \mathbb{Z}$, Z is a divisor of $a_iX + b_iY$. It follows that p is a divisor of

$$b_i'(a_iX + b_iY) - b_i(a_i'X + b_i'Y) = (a_ib_i' - b_ia_i')X = X$$

and also of

$$-a_i'(a_iX + b_iY) + a_i(a_i'X + b_i'Y) = (a_ib_i' - b_ia_i')Y = Y,$$

but that means that p divides the gcd of $\{X, Y, Z\}$, so $p = \pm 1$. Hence the fraction

$$g_i = \frac{a_i'X + b_i'Y}{Z}$$

is reduced to its lowest terms, so the order of g in $(\mathbb{Q}/\mathbb{Z})^{\#S}$ is $Z = \#M_S$.

Let s and t such that $d = \#S - 2 - s - t$ is non-negative. For every $e \in \mathbb{Z}/\#\langle g\rangle\mathbb{Z}$ let δ_e be the dependency for p_S from theorem (4.2.2) applied to D_S corresponding to the point (em_x, em_y) and the given s and t. For $i \in S$ the weight functions of this dependency is given by

$$c_e = \left(m \mapsto \frac{r_i(-b_i)^s a_i^t}{d!}\zeta_{e,i}^m m^d\right)_{i\in S},$$

where $\zeta_{e,i} = \zeta(a_i'(ex_i)+b_i'(ey_i)) = \zeta(eg_i)$. Running over all s, t and e produces a basis for $p_{S,(x',y')}$ for all $(x', y') \in M_S$, i.e., a basis for V_S.

Note that c_e can be written as the product of the functions

$$v_e = \left(m \mapsto \zeta(meg_i)\right)_{i\in S}$$

with the functions

$$\lambda = \left(m \mapsto \frac{r_i(-b_i)^s a_i^t}{d!}m^d\right)_{i\in S}.$$

Note that λ does not depend on e and takes rational values everywhere. Thus the linear combinations of the c_e that are fixed under the Galois action are precisely the linear combinations of the v_e that are fixed under the Galois action. Note that we can apply lemma (4.3.14) to the v_e. We conclude that the span of the c_e is the same as the span of the functions $\lambda \cdot w_h$ where for $h \in \frac{1}{\#\langle g\rangle}\mathbb{Z}/\mathbb{Z}$, w_h is given by

$$w_h = \left(m \mapsto \begin{cases} 1 & \text{if } h - mg_i \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases}\right)_{i\in S}.$$

Note that $\lambda \cdot w_h$ is precisely the dependency associated to S, s, t and h in the theorem. We conclude that running over all s, t and h produces a $\overline{\mathbb{Q}}$-basis of V_S consisting of elements fixed by the Galois action. Running over all S produces a set of dependencies that are fixed by the Galois action such that their $\overline{\mathbb{Q}}$-span contains a basis for $\mathrm{Dep}(p)$, hence is all of $\mathrm{Dep}(p)$. Hence this is a generating set for $\mathrm{Dep}(p)^{\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})} = \mathrm{Dep}(p_{\mathbb{Q}})$. By lemma (4.2.4), $\mathrm{Dep}(p_{\mathbb{Q}}) = \mathrm{Dep}(p_D)$. □

Proof of (4.3.1). Let $d \in \mathrm{Dep}(p_{\mathbb{Q}})$ and let $c_1, \ldots, c_n$ be the associated coefficient functions from (2.1.17). It is clear from this lemma that d is in $\mathrm{Dep}(p_{\mathbb{Z}})$ if and only if the c_i only take integer values. From this it follows at once that $\phi(\mathrm{Dep}(p_{\mathbb{Z}}))$ is contained in

$$\mathbb{Z}^{r_1-1} \oplus \cdots \oplus \mathbb{Z}^{r_n-1}.$$

Conversely, suppose we have a dependency d such that $\phi(d)$ has all integer coefficients. By theorem (3.5.5), the coefficient functions c_i each satisfy a recurrence of the form

$$\forall m \in \mathbb{Z} : a_{i,1}c_i(m+1) + \cdots + a_{i,r_i}c_i(m+r_i) = 0$$

with all the $a_{i,j} \in \mathbb{Z}$, $a_{i,1} = \pm 1$ and $a_{i,r_i} = \pm 1$. By assumption $c_i(m) \in \mathbb{Z}$ for $m \in \{1, \ldots, r_i - 1\}$. Applying the recurrence for $m = 0$, we conclude that $\pm c_i(r_i)$ is equal to a $\mathbb{Z}$-linear combination of integers, hence it is an integer. We can proceed in this way by induction and show that $c_i(m) \in \mathbb{Z}$ for all $m \in \mathbb{Z}$. Hence d is in $\mathrm{Dep}(p_{\mathbb{Z}})$, proving the other inclusion. □

Proof of (4.3.2). We assume the input D is given as a sequence of n pairs $d_i = (a_i, b_i)$. It is easy to check in polynomial time in the input length that these indeed form a 2-regular sequence of primitive directions.

Let $p_{\mathbb{Q}}$ be the grid reconstruction system over $\mathbb{Q}$ associated to $(\mathbb{Z}^2, D)$. As in theorem (4.3.1), let ϕ be the linear map

$$\mathrm{Dep}(p_{\mathbb{Q}}) \longrightarrow \mathbb{Q}^{r_1-1} \oplus \cdots \oplus \mathbb{Q}^{r_n-1}$$

sending a dependency to

$$((c_1(1), \ldots, c_1(r_1 - 1)), \ldots, (c_n(1), \ldots, c_n(r_n - 1))),$$

where the c_i are its coefficient functions as defined in lemma (2.1.17). The algorithm to compute the integral dependencies now has the following steps:

1. determine all intersection types for D;
2. for each intersection type S and every admissible s, t and h as in theorem (4.3.8), let d be the associated dependency and compute the vector $\phi(d)$;
3. determine a $\mathbb{Z}$-basis for the intersection of the span of these vectors with the integer lattice.

By theorem (4.3.1), this $\mathbb{Z}$-basis is the required output. It remains to be shown that each of these steps can be executed in time polynomial in the input length and the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$.

For the first step, note that we can enumerate M by computing for each $1 \leq i < j \leq n$ the set $M_{\{i,j\}}$. Note that $M_{\{i,j\}}$ has $|a_i b_j - a_j b_i|$ elements, so the number of points we generate in this way is precisely the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$. In general there will be duplicate points, so M has at most this many elements. For each element $(x, y) \in M$ we can then compute the set $S = \{i \mid a_i x + b_i y \in \mathbb{Z}\}$. From lemma (4.3.4) we know that S is an intersection type and that we obtain every intersection type in this manner. Thus the number of intersection types is at most the number of elements of M and we can enumerate them in time polynomial in the input and the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$.

For the second step, the number of admissible s and t is bounded by n and the number of admissible h's is bounded by $\#M$, which, as we have seen, is bounded by the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$. From the construction of the r_i in theorem (3.5.5) it follows that $(r_1 - 1) + \cdots + (r_n - 1)$ is precisely two times the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$ as computed in theorem (3.5.1). This is the number of coefficients we have to compute for each dependency. Computing each coefficient is easy, so it follows that this step can also be done in time polynomial in the input length and the rank of $\mathrm{Dep}(p_{\mathbb{Z}})$.

The last step is a linear algebraic step. It can be done in the required amount of time using well-known methods from linear algebra over the integers (see e.g. [6, Ch. 2]). □

5 – Periodic grids

5.1 Introduction

As the title suggests, in this chapter we will explore grid reconstruction systems whose grid is 'periodic', which we will define in a moment. We will derive some important characteristics of their kernels and cokernels, and give algorithms to compute them over the rationals and the integers. Some of the finer points concerning their structure remain open and we conclude with a few avenues for possible future research.

Definition 5.1.1. Let $A \subset \mathbb{Z}^r$ be a subset and let $\Lambda \subset \mathbb{Z}^r$ be a finite index subgroup. Then A is Λ-*periodic* if it satisfies

$$\forall a \in A : \forall \lambda \in \Lambda : a + \lambda \in A.$$

Example 5.1.2. The grid from example (1.4.4), given by

$$A = \mathbb{Z}^2 \setminus 3\mathbb{Z}^2 = \{(x, y) \in \mathbb{Z}^2 \mid (x, y) \not\equiv (0, 0) \bmod 3\},$$

is periodic for $\Lambda = 3\mathbb{Z}^2$. Note that any element of A can be written in a unique way as $\lambda + x$ with $\lambda \in \Lambda$ and x from the set

$$X = \{0, 1, 2\} \times \{0, 1, 2\} \setminus \{(0, 0)\}.$$

This means we can view the set A as a union of copies of X translated by elements of Λ, or conversely as a union of 8 translated copies of Λ, indexed by the elements of X.

In section 3.1 and chapter 4 we looked at grid reconstruction systems where the grid is $\mathbb{Z}^r$. The key ingredient in a lot of this work is that the reconstruction system is a map between rings and that it respects the ring structures. For periodic grids this is not always the case, but we do have a similar structure that we can employ to our advantage, as is explained in the following lemma.

Lemma 5.1.3. *Let $A \subset \mathbb{Z}^r$ be Λ-periodic and let D be a sequence of directions in $\mathbb{Z}^r$. Let k be a commutative ring and let $p : T \to P$ be the grid reconstruction system over k associated to (A, D). Then there are natural $k[\Lambda]$-module structures on T and P, and p is a $k[\Lambda]$-linear map.*

Proof. The inclusion $\Lambda \subset \mathbb{Z}^r$ gives rise to an inclusion $k[\Lambda] \subset k[\mathbb{Z}^r]$, which is a ring homomorphism, so we can view $k[\mathbb{Z}^r]$ as a $k[\Lambda]$-module. We claim $k[A]$ viewed as a subset of $k[\mathbb{Z}^r]$ is a sub-$k[\Lambda]$-module. It is clearly a sub-k-module, so it suffices to show it is closed under multiplication by elements of $k[\Lambda]$. By linearity, we only have to check this for k-bases of $k[\Lambda]$ and $k[A]$ Let $\lambda \in \Lambda$ and $a \in A$. Let u^λ and u^a be the corresponding elements of $k[\mathbb{Z}^r]$. Then we have $u^\lambda \cdot u^a = u^{\lambda+a}$ and since A is Λ-periodic, $\lambda + a$ is in A. Hence the multiplication by elements of $k[\Lambda]$ maps $k[A]$ into itself, i.e. $k[A]$ is a sub-$k[\Lambda]$-module of $k[\mathbb{Z}^r]$ and in particular a $k[\Lambda]$-module.

Let $D = (d_1, \ldots, d_n)$ and let $i \in \{1, \ldots, n\}$. The quotient map

$$p_{d_i} : k[\mathbb{Z}^r] \to k[\mathbb{Z}^r/d_i]$$

is a ring homomorphism. Its restriction to $k[\Lambda]$ therefore is one as well, giving a $k[\Lambda]$-module structure on $k[\mathbb{Z}^r/d_i]$, defined by

$$\begin{array}{ccc} k[\Lambda] \times k[\mathbb{Z}^r/d_i] & \longrightarrow & k[\mathbb{Z}^r/d_i] \\ (x, y) & \mapsto & p_{d_i}(x)y. \end{array}$$

We show that $k[A_i]$ viewed as a subset of $k[\mathbb{Z}^r/d_i]$ is in fact a sub-$k[\Lambda]$-module. Recall that A_i is the image of A in the quotient $\mathbb{Z}^r/d_i$, so any element x of $k[A_i]$ can be written as $p_{d_i}(y)$ with $y \in k[A]$. For $z \in k[\Lambda]$ we then have

$$zx = p_{d_i}(z)p_{d_i}(y) = p_{d_i}(zy)$$

and since $k[A]$ is a sub-$k[\Lambda]$-module of $k[\mathbb{Z}^r]$, the element zy is in $k[A]$ and so $p_{d_i}(zy)$ is in $k[A_i]$. Hence $k[A_i]$, viewed as a subset of $k[\mathbb{Z}^r/d_i]$ is closed under multiplication by elements of $k[\Lambda]$. So it is a submodule, and in particular a $k[\Lambda]$-module.

Moreover, the projection map p_{d_i} that maps $k[A]$ into $k[A_i]$ is $k[\Lambda]$-linear more or less by construction. It follows that the reconstruction system

$$p = (p_{d_1}, \ldots, p_{d_n})$$

is a $k[\Lambda]$-linear map $k[A] \to \bigoplus_{i=1}^n k[A_i]$. □

We conclude this section with some results on the kernels and cokernels of grid reconstruction systems with periodic grids. These will be proved in the

next two sections. The first result is analogous to theorem (3.1.2) for full grids.

Theorem 5.1.4. *Let $A \subset \mathbb{Z}^r$ be Λ-periodic and D a 2-regular sequence of directions in $\mathbb{Z}^r$. Let k be a field. Let p be the grid reconstruction system over k associated to (A, D). Then there are $f_1, \ldots, f_m \in k[A]$ such that $\ker(p)$ is generated by the f_i as a $k[\Lambda]$-module.*

To give an algorithm for computing these kernels, we need to make some additional assumption on the field k, for example that it is a prime field, i.e. $k = \mathbb{Q}$ or $k = \mathbb{F}_p$ for some prime p.

Theorem 5.1.5. *There exists an algorithm that given a Λ-periodic grid $A \subset \mathbb{Z}^r$, a 2-regular sequence D of directions in $\mathbb{Z}^r$ and a prime field k, computes m and $f_1, \ldots, f_m$ from theorem* (5.1.4).

For the cokernel, the results are not nearly as nice as theorem (3.1.6) for full grids. In the two-dimensional case we do still get a finitely generated module, but it will not always be free.

Theorem 5.1.6. *Let $A \subset \mathbb{Z}^2$ be Λ-periodic and D a 2-regular sequence of directions in $\mathbb{Z}^2$. Let k be a commutative ring. Let p be the grid reconstruction system over k associated to (A, D). Then $\operatorname{cok}(p)$ is a finitely generated k-module.*

When k is a field, k-modules are always free, so we can apply corollary (2.1.19) and conclude the dependencies give a full answer to the consistency problem. However, for $k = \mathbb{Z}$ this does not work. Giving generators for the cokernel as elements in the codomain is possible, but not very useful. The following theorem shows that we can compute an explicit realisation of the cokernel as a quotient of some finitely generated free abelian group.

Theorem 5.1.7. *There exists an algorithm that given a Λ-periodic grid $A \subset \mathbb{Z}^2$ and a 2-regular sequence D of primitive directions in $\mathbb{Z}^2$ produces an integer N, a linear map*

$$\phi : \bigoplus_{i=1}^{n} \mathbb{Z}[A_i] \longrightarrow \mathbb{Z}^N$$

and generators of a subgroup $M \subset \mathbb{Z}^N$ such that

$$\bar{\phi} : \bigoplus_{i=1}^{n} \mathbb{Z}[A_i] \longrightarrow \mathbb{Z}^N/M$$

is onto and $\ker(\bar{\phi}) = \mathrm{im}(p)$*, where p is the grid reconstruction system over $\mathbb{Z}$ associated to (A, D).*

Having a presentation of a finitely generated abelian group as a quotient $\mathbb{Z}^N/M$ as in the theorem above makes it possible to compute things like the torsion group and the rank of the free part using standard algorithms for linear algebra over $\mathbb{Z}$.

It is a priori not at all clear how the map ϕ from the theorem can be represented in a finite way. This is a similar problem to the representation of dependencies for full grid reconstruction systems that was discussed in remark (3.5.3). The solution is similar too and will be discussed in detail in the proof of theorem (5.1.7). One thing that seems reasonable and turns out to be true is that we can evaluate ϕ on any given element of $\bigoplus_{i=1}^{n} \mathbb{Z}[A_i]$ efficiently. The upshot of this is that there is a consistency algorithm as described in the following corollary.

Corollary 5.1.8. *Let $A \subset \mathbb{Z}^2$ be Λ-periodic and D a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. Let $p : T \to P$ be the grid reconstruction system over $\mathbb{Z}$ associated to (A, D). Then there is an algorithm that given $x \in P$ decides if $x \in \mathrm{im}(p)$ in time polynomial in the length of the input.*

5.2 Computing the kernel

The aim of this section is to prove theorems (5.1.4) and (5.1.5). We derive the first of these as consequence of a ring-theoretic generality.

Proof of (5.1.4). When k is a field, it is in particular a Noetherian ring (see e.g. chapters 6 and 7 of [1]) and so $k[\Lambda]$ is Noetherian as it is a finitely generated k-algebra [1, Cor. 7.7]. Note that $k[A]$ is a finitely generated $k[\Lambda]$-module as Λ has finite index in $\mathbb{Z}^r$ and so any sub-$k[\Lambda]$-module of $k[A]$ is finitely generated [1, Prop. 6.5] as a $k[\Lambda]$-module. Hence there are elements $f_1, \ldots, f_m$ that generate $\ker(p)$ as a sub-$k[\Lambda]$-module of $k[A]$. □

Sadly, this proof is entirely non-constructive, so it is of no help when trying to compute m and $f_1, \ldots, f_m$. To find these, we use the same technique as for reconstruction systems with finite convex grids, we compare with the full grid system using lemma (3.2.2).

Lemma 5.2.1. *Let $A \subset \mathbb{Z}^r$ be Λ-periodic, let D be a 2-regular sequence of directions in $\mathbb{Z}^r$ and let k be a commutative ring. Let p and p_f be the grid reconstruction systems over k associated to (A, D) and $(\mathbb{Z}^r, D)$ respectively. Then $\ker(p_f)$ and $k[\mathbb{Z}^r \setminus A]$ are free sub-$k[\Lambda]$-modules of $k[\mathbb{Z}^r]$ of finite rank, the natural map*

$$\phi : \ker(p_f) \longrightarrow k[\mathbb{Z}^r \setminus A]$$

is $k[\Lambda]$-linear and $\ker(p) = \ker(\phi)$.

Proof. The action of Λ on $\mathbb{Z}^r$ is free and as Λ is of finite index, a system of representatives for $\mathbb{Z}^r/\Lambda$ is finite, so $k[\mathbb{Z}^r]$ is a free $k[\Lambda]$-module of finite rank. By theorem (3.1.2), the image of the map

$$k[\mathbb{Z}^r] \longrightarrow k[\mathbb{Z}^r] \qquad x \mapsto xF_D$$

is $\ker(p_f)$. This map is $k[\mathbb{Z}^r]$-linear, so it certainly is $k[\Lambda]$-linear and thus gives an isomorphism of $k[\Lambda]$-modules $k[\mathbb{Z}^r] \cong \ker(p_f)$. Hence $\ker(p_f)$ is a free $k[\Lambda]$-module of finite rank.

Note that $\mathbb{Z}^r \setminus A$ is also Λ-periodic, and so Λ acts freely on this set. Hence $k[\mathbb{Z}^r \setminus A]$ is also a free $k[\Lambda]$-module of finite rank. The map ϕ is induced by the natural map $k[\mathbb{Z}^r] \to k[\mathbb{Z}^r \setminus A]$ that restricts a function $f : \mathbb{Z}^r \to k$ to $\mathbb{Z}^r \setminus A$. As Λ acts by translation on both sets, it is clear that this map, and therefore ϕ, is $k[\Lambda]$-linear.

Applying lemma (3.2.2) to $A \subset \mathbb{Z}^r$ we see that $\ker(p)$ is given by

$$\ker(p) = \{f \in \ker(p_f) \mid \operatorname{supp}(f) \subset A\}.$$

For $f \in \ker(p_f)$, we have $\operatorname{supp}(f) \subset A$ if and only if f maps every point in $\mathbb{Z}^r \setminus A$ to zero, i.e., if and only if $\phi(f) = 0$. Hence $\ker(p) = \ker(\phi)$ holds as claimed. □

Example 5.2.2. Let A be given by $A = \mathbb{Z}^2 \setminus 3\mathbb{Z}^2$ as in example (1.4.4) and let D be the sequence of four directions from that example. As we saw in example (5.1.2), A is Λ-periodic with $\Lambda = 3\mathbb{Z}^2$.

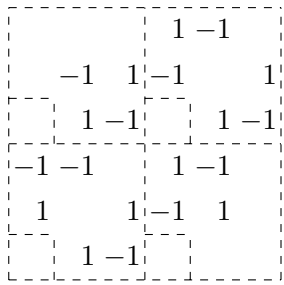
Identify $k[\mathbb{Z}^2]$ with the Laurent polynomial ring $k[u, u^{-1}, v, v^{-1}]$ as usual. Then $k[\Lambda]$ is identified with the subring generated by $U = u^3$ and $V = v^3$. A basis for $k[u, u^{-1}, v, v^{-1}]$ as a $k[U, U^{-1}, V, V^{-1}]$-module is given by the monomials

$$1, v, v^2, u, uv, uv^2, u^2, u^2v, u^2v^2.$$

The generator of the kernel of the full grid reconstruction system associated with these directions was given as a polynomial in $k[u, u^{-1}, v, v^{-1}]$ in example (3.1.4). We write F_D for this polynomial. To give a matrix for ϕ, we give for each element in the $k[U, U^{-1}, V, V^{-1}]$ basis of $k[u, u^{-1}, v, v^{-1}]$ above the product of that element with F_D and restrict it to $k[\mathbb{Z}^2 \setminus A] = k[\Lambda]$. The results are in the table below.

x	restriction of xF_D to $k[\Lambda]$
1	$1 + UV$
v	$-UV$
v^2	$-V + UV - UV^2$
u	$-UV$
uv	$2UV$
uv^2	$-UV$
u^2	$-U + UV - U^2V$
u^2v	$-UV$
u^2v^2	$UV + U^2V^2$

Various elements of $\ker(\phi)$ can be derived easily from this table, such as $u - v$. The element $(u - v)F_D$ of $\ker(p_f)$ is pictured below. Its support is indeed contained in A, thus it belongs to $\ker(p)$.



The upshot of lemma (5.2.1) is that computing the kernel of a periodic grid reconstruction system is reduced to a computation of syzygies in a free $k[\Lambda]$-module of finite rank. Fix a basis $g_1, \ldots, g_s$ of $\ker(p_f)$. What we want to

know is the $k[\Lambda]$-linear relations (i.e. the syzygies) that hold between the elements $\phi(g_1), \ldots, \phi(g_s)$ in $k[\mathbb{Z}^r \setminus A]$.

The computation of syzygies is a fundamental problem in computational commutative algebra. For modules over a polynomial ring $S = k[x_1, \ldots, x_n]$ over a field k the theory of Gröbner bases deals with this and related problems. An introduction to this topic can be found in [8, Ch. 15]. The basic algorithm for computing Gröbner bases, Buchberger's algorithm [8, Alg. 15.9] also produces generators for the syzygies (cf. [8, Thm. 15.10]) for an S-linear map $S^a \to S^b$.

Of course, the ring $k[\Lambda]$ we are dealing with is not a (multivariate) polynomial ring, but it's 'almost' one. The reduction to the polynomial case is carried out in the proof below.

Proof of (5.1.5). We suppose that a basis $\lambda_1, \ldots, \lambda_r$ of Λ is given. Identify $S = k[x_1, \ldots, x_r]$ with a subring of $k[\Lambda]$ by mapping x_i to λ_i.

Compute a system of representatives for $\mathbb{Z}^r/\Lambda$ and divide it into sets X and Y consisting of the classes in A and $\mathbb{Z}^r \setminus A$ respectively.

Let p_f be the grid reconstruction system over k corresponding to $(\mathbb{Z}^r, D)$ and compute F_D, a generator of $\ker(p_f)$ as a $k[\mathbb{Z}^r]$-module as in theorem (3.1.2).

For $a, b \in X \cup Y$ compute the elements $f_{a,b} \in k[\Lambda]$ defined by the following equations. For every $a \in X \cup Y$ we have

$$u^a F_D = \sum_{b \in X \cup Y} u^b f_{a,b}$$

in $k[\mathbb{Z}^r]$. (We only need the $f_{a,b}$ for $a \in X \cup Y$ and $b \in Y$, but it is more convenient to define them for $b \in X$ as well.)

For every $a \in X \cup Y$ pick a $\lambda_a \in \Lambda$ such that $s_a = u^{\lambda_a}$ is in S and $g_{a,b} = s_a f_{a,b}$ if in S for all $b \in Y$. Note that the $g_{a,b}$ represent a matrix for the map $\phi : \ker(p_f) \to k[\mathbb{Z}^r \setminus A]$ from lemma (5.2.1) relative to the $k[\Lambda]$-basis $(u^{a+\lambda_a} F_D)_{a \in X \cup Y}$ of $\ker(p_f)$ and $(u^b)_{b \in Y}$ for $k[\mathbb{Z}^r \setminus A]$.

Moreover, since all the $g_{a,b}$ are in S, the same matrix gives rise to a map

$$\phi_S : S[X \cup Y] \longrightarrow S[Y].$$

This is an S-linear map between finitely generated free modules over the multivariate polynomial ring S, and so we can use Gröbner basis methods to compute a generating set for its kernel. As mentioned before, one way to do this would be to apply Buchberger's algorithm, see [8, Thm. 15.10].

Note that any element in $(k[\Lambda])[X \cup Y]$ can be multiplied by a unit in $k[\Lambda]$ so that it is in $S[X \cup Y]$. Hence the elements that generate $\ker(\phi_S)$ as an S-module, will generate $\ker(\phi)$ as a $k[\Lambda]$-module.

Computing the images of these generators in $\ker(p_f)$ we obtain generators for $\ker(p)$ as a sub-$k[\Lambda]$-module of $\ker(p_f)$. □

The computation of Gröbner bases, syzygies and related algorithms are implemented natively in many computer algebra systems, such as Magma (see `http://magma.maths.usyd.edu.au/magma/` or [4]) and the open-source system Sage (see `http://sagemath.org/`). They often use algorithms that perform better than Buchberger's cited above. If we want to compute examples, we can simply ask the system to produce the required syzygies and not worry about the implementation.

Example 5.2.3. We will now give generators for the kernel of the grid reconstruction system described in example (1.4.4) as a $k[\Lambda]$-module, where $\Lambda = 3\mathbb{Z}^2$. In example (5.2.2) we computed the map

$$\phi : k[U, U^{-1}, V, V^{-1}]^9 \longrightarrow k[U, U^{-1}, V, V^{-1}]$$

from lemma (5.2.1). A generating set for the kernel of this map can be computed using the algorithm from the proof of theorem (5.1.5). One such set is given by

$$\begin{aligned}
&u^2v(UV+1) + u^2v^2,\\
&U + u^2 + u^2v,\\
&Uv^2 + Uu^2v + u^2v^2,\\
&V + v^2 + u^2v,\\
&Vu^2 + Vu^2v + u^2v^2,\\
&v - u^2v,\\
&u - u^2v,\\
&uv - 2u^2v,\\
&uv^2 - u^2v.
\end{aligned}$$

It was computed using Magma. Multiplying these with F_D from example (5.2.2) and recalling that $U = u^3$ and $V = v^3$ one obtains generators for $\ker(p)$ as a subspace of $\ker(p_f)$.

5.3 Computing the cokernel

We now turn our attention to the cokernels of grid reconstruction systems with periodic grid. Continuing with the approach to these systems we used in the previous section, we have the following sufficient condition for when the cokernel of the periodic grid system is the same as that of the full grid system.

Lemma 5.3.1. *Let $A \subset \mathbb{Z}^r$ be Λ-periodic and let D be a sequence of directions in $\mathbb{Z}^r$. Let k be a commutative ring and let p and p_f be the grid reconstruction systems over k associated to (A, D) and $(\mathbb{Z}^r, D)$ respectively. Let f_D be the generator of* $\ker(p_f)$ *from theorem* (3.1.2)*. Suppose that*

1. *all lines in directions from D intersect A;*
2. *for every $x \in \mathbb{Z}^r \setminus A$ there is a $y \in \mathbb{Z}^r$ such that*

$$\operatorname{supp}(u^y f_D) \cap (\mathbb{Z}^r \setminus A) = \{x\}$$

and the coefficient of u^x in $u^y f_D$ is ± 1;

then the natural map $\operatorname{cok}(p) \to \operatorname{cok}(p_f)$ *is an isomorphism.*

Proof. Let $p_{\mathbb{Z}^r/A}$ be the reconstruction system defined in lemma (2.4.4) for the inclusion $A \subset \mathbb{Z}^r$. From that lemma, there is an exact sequence of reconstruction systems

$$0 \to p \to p_f \to p_{\mathbb{Z}^r/A} \to 0.$$

Note that by point 1 from the lemma, $p_{\mathbb{Z}^r/A}$ is the map

$$k[\mathbb{Z}^r \setminus A] \longrightarrow 0,$$

so we have $\operatorname{cok}(p_{\mathbb{Z}^r/A}) = 0$ and $\ker(p_{\mathbb{Z}^r/A}) = k[\mathbb{Z}^r \setminus A]$. Hence by lemma (2.4.7), the map $\operatorname{cok}(p) \to \operatorname{cok}(p_f)$ is surjective.

Note that by point 2 from the lemma, the natural map

$$\ker(p_f) \to k[\mathbb{Z}^r \setminus A]$$

is onto: this point states that for every $x \in \mathbb{Z}^r \setminus A$, the element $u^y f_D$ in $\ker(p_f)$ maps to $\pm u^x$. Hence, by corollary (2.4.13), the map $\operatorname{cok}(p) \to \operatorname{cok}(p_f)$ is injective. □

The conditions from the lemma may appear rather specific, but they are often both satisfied when $\mathbb{Z}^r \setminus A$ is a relatively sparse set. Moreover, they are relatively easy to check, both by hand and by computer. In the following example, we see that they apply to the running example from the previous section.

Example 5.3.2. Let A be given by $A = \mathbb{Z}^2 \setminus 3\mathbb{Z}^2$ as in example (1.4.4) and let D be the sequence of four directions from that example. As we saw in example (5.1.2), A is Λ-periodic with $\Lambda = 3\mathbb{Z}^2$.

Let k be a commutative ring and let p and p_f be the grid reconstruction systems over k associated to (A, D) and $(\mathbb{Z}^2, D)$ respectively. Let F_D be the generator of $\ker(p_f)$ as a $k[\mathbb{Z}^2]$ module, as in example (3.1.4). Identify $k[\mathbb{Z}^2]$ with the Laurent polynomial ring $k[u, u^{-1}, v, v^{-1}]$ as usual. Then $k[\Lambda]$ is identified with the subring generated by $U = u^3$ and $V = v^3$.

Note that any line in a direction from D goes through a point in A, hence the first condition of lemma (5.3.1) is satisfied.

We computed in example (5.2.2) that the restriction of uF_D to $k[\Lambda]$ is $-UV$. It follows that

$$\operatorname{supp}(uF_D) \cap (\mathbb{Z}^r \setminus A)$$

consists of a single point and the coefficient at that point is -1. As Λ acts transitively on $\mathbb{Z}^r \setminus A$, it follows that the second condition of lemma (5.3.1) is also satisfied. Hence we conclude that $\operatorname{cok}(p) \cong \operatorname{cok}(p_f)$ by the lemma.

Lemma (5.3.1) does not apply to all grid reconstruction systems for periodic grids. In the rest of this section, we give a different approach to understanding these systems, which lends itself better to computations for the cokernel. Rather than look at the inclusion $A \subset \mathbb{Z}^r$, we consider a modification of the set of lines that makes the problem easier.

Throughout the section, fix a finite index subgroup $\Lambda \subset \mathbb{Z}^r$, a Λ-periodic grid $A \subset \mathbb{Z}^r$ and a 2-regular sequence $D = (d_1, \ldots, d_n)$ of directions in $\mathbb{Z}^r$. Fix a commutative ring k and let p be the grid reconstruction system over k associated to (A, D).

For $i = 1, \ldots, n$ let d_i' be a generator of $\langle d_i \rangle \cap \Lambda$ and let $A_{s,i}$ be the image of A in $\mathbb{Z}^r / d_i'$. Let Λ_i be the quotient $\Lambda / \langle d_i' \rangle$, and note that this is also the image of Λ in $\mathbb{Z}^r / d_i$.

Let $D' = (d'_1, \ldots, d'_n)$ and let p_s be the grid reconstruction system

$$p_s : k[A] \longrightarrow \bigoplus_{i=1}^{n} k[A_{s,i}]$$

associated to (A, D').

Lemma 5.3.3. *For $i = 1, \ldots, n$ there is a natural surjective k-linear map*

$$s_i : k[A_{s,i}] \longrightarrow k[A_i].$$

Let $s = (s_1, \ldots, s_n)$, then the pair $(\mathrm{id}_{k[A]}, s)$ is a surjective morphism of reconstruction systems $p_s \to p$.

Proof. As d'_i is a multiple of d_i, the quotient map $\mathbb{Z}^r \to \mathbb{Z}^r/d_i$ factors naturally as

$$\mathbb{Z}^r \longrightarrow \mathbb{Z}^r/d'_i \longrightarrow \mathbb{Z}^r/d_i.$$

Note that A_{s_i} is the image of A under the first map, and A_i is the image of A under the composition. It follows that the second map is a surjection $A_{s_i} \to A_i$. This map extends linearly to a surjective k-linear map

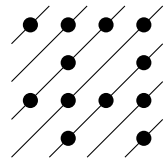
$$s_i : k[A_{s,i}] \longrightarrow k[A_i]$$

that fits into the commutative diagram below.

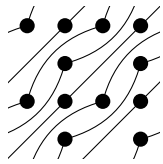
$$\begin{array}{ccc} k[A] & \xrightarrow{p_{d'_i}} & k[A_{s,i}] \\ & \searrow^{p_{d_i}} & \downarrow{\scriptstyle s_i} \\ & & k[A_i] \end{array}$$

Hence we have $p = s \circ p_s$ and so $(\mathrm{id}_{k[A]}, s)$ is a morphism of reconstruction systems, which is surjective since both component maps are. □

Example 5.3.4. Let $A = \mathbb{Z}^2 \setminus 2\mathbb{Z}^2$ and $\Lambda = 2\mathbb{Z}^2$. Note that A is Λ-periodic. Let $D = (d_1, d_2, d_3)$ with $d_1 = (0, 1)$, $d_2 = (1, 0)$ and $d_3 = (1, 1)$. The following picture shows part of the set A together with the lines in direction d_3 going through it.



The direction d_3' is a generator $\langle d_3\rangle \cap 2\mathbb{Z}^2$, so we can take $d_3' = 2d_3 = (2,2)$. The lines in direction d_3' through the same subset of A are shown in the following picture.



In this second picture, some of the lines from the first picture have been split in two. The result is that while some of the lines in the first picture go through several equivalence classes in $A/2\mathbb{Z}^2$, each line in the second picture goes through points from exactly one such class. As a result, we can 'pull apart' p_s into three copies of a full grid reconstruction system. This is explained in the next lemma.

Lemma 5.3.5. *Let ϕ be an isomorphism $\Lambda \to \mathbb{Z}^r$. For $i = 1, \ldots, n$, let $e_i = \phi(d_i')$. Let $E = (e_1, \ldots, e_n)$. Let q be the grid reconstruction system over k associated to $(\mathbb{Z}^r, E)$. Then p_s is isomorphic to the direct sum of $\#(A/\Lambda)$ copies of q.*

Proof. As $d_1', \ldots, d_n'$ are in Λ we can apply lemma (2.3.10) and conclude that p_s can be written as a direct sum of q_x for $x \in \mathbb{Z}^r/\Lambda$, where q_x is the grid reconstruction system associated to (A_x, E) for some $A_x \subset \mathbb{Z}^r$ and E as in the lemma.

As A is Λ-periodic, for every $x \in \mathbb{Z}^r/\Lambda$, the set $A \cap (x+\Lambda)$ is either empty (if $x \notin A/\Lambda$) or equal to $x + \Lambda$ (if $x \in A/\Lambda$). From the proof of (2.3.10), we therefore conclude that if $x \notin A/\Lambda$ we have $q_x = 0$; and if $x \in A/\Lambda$ we have $A_x = \mathbb{Z}^r$ and so $q_x = q$. The result follows. □

Specialising to the case where $r = 2$ and D is primitive, we can now prove the structure result for the cokernel announced in section 5.1.

Proof of (5.1.6). Suppose $r = 2$ and $D = (d_1, \ldots, d_n)$ is primitive. Fix some isomorphism $\phi : \Lambda \to \mathbb{Z}^2$ and let $E = (e_1, \ldots, e_n)$ and q be as in lemma (5.3.5). We want to show that E is again primitive. Suppose it is not, then for some $i \in \{1, \ldots, n\}$ we have $e_i = mv$ for some integer $m > 1$ and some $v \in \mathbb{Z}^2$. Let $w = \phi^{-1}(v)$. Then $mw = \phi^{-1}(e_i) = d_i'$ is a multiple of d_i, so w is a multiple of d_i. Hence $w \in \langle d_i\rangle \cap \Lambda$ and $mw = d_i'$, contradicting the fact that d_i' is a generator of $\langle d_i\rangle \cap \Lambda$.

Hence we can apply theorem (3.5.1) to q, as E is a 2-regular sequence of primitive directions in $\mathbb{Z}^2$. We conclude that $\mathrm{cok}(q)$ is a free k-module of finite rank. By lemma (5.3.5), p_s is a direct sum of finitely many copies of q, hence $\mathrm{cok}(p_s)$ is also a free k-module of finite rank.

By lemma (5.3.3) there is a surjective map $p_s \to p$ and so $\mathrm{cok}(p_s)$ maps onto $\mathrm{cok}(p)$. Hence $\mathrm{cok}(p)$ is finitely generated. □

To get to the algorithmic result for the cokernel announced in section 5.1, we have to describe the kernel of the map $\mathrm{cok}(p_s) \to \mathrm{cok}(p)$. Fortunately, the kernel of each $s_i : k[A_{s,i}] \to k[A]$ can be described easily, as is shown in the following lemmas. Note that these lemmas hold for all r and k.

Lemma 5.3.6. *Let $i \in \{1, \ldots, n\}$. Let X be a system of representatives for A/Λ. Then the image of X in $A_{s,i}$ is a system of representatives for the Λ_i action on $A_{s,i}$.*

Proof. Write π_i for the quotient map $A \to A_{s_i}$. For $a \in A_{s,i}$, there is some $\tilde{a} \in A$ such that $\pi(\tilde{a}) = a$. As X is a system of representatives for the Λ action on A, $\tilde{a}$ can be written as $x + \lambda$ for some $x \in X$ and $\lambda \in \Lambda$. It follows that $a = \pi(x) + \bar{\lambda}$, where $\bar{\lambda}$ is the image of λ in Λ_i.

Suppose x and x' are in X and that $\pi(x) = \pi(x') + \lambda$ for some $\lambda \in \Lambda_i$. Let $\tilde{\lambda}$ be any lift of λ to Λ. Then $x'' = x' + \tilde{\lambda}$ differs from x by a multiple of d_i', as they have the same image under π. By construction, $d_i' \in \Lambda$, so that x'' and x differ by an element of Λ. As x'' and x' also differ by an element of Λ, x' and x must be in the same class in A/Λ. Since they are both in X, it follows that $x' = x$.

We conclude that $\pi(X)$ is a system of representatives for the Λ_i action on $A_{s,i}$. □

Lemma 5.3.7. *Let $i \in \{1, \ldots, n\}$. Let X be a system of representatives for A/Λ and let Y be a system of representatives for the Λ_i action on A_i. For $x \in X$, let $y_x \in Y$ and $\lambda_x \in \Lambda_i$ be such that the image of x in A_i is $y_x + \lambda_x$. Let π be the quotient map $A \to A_{s,i}$. Then the kernel of $s_i : k[A_{s,i}] \to k[A_i]$ is generated as a $k[\Lambda_i]$ module by*

$$u^{\lambda_{x'}} u^{\pi}(x) - u^{\lambda_x} u^{\pi}(x')$$

for all $x, x' \in X$ such that $y_{x'} = y_x$.

Proof. Note that every $f \in k[A_{s_i}]$ can be written as

$$f = \sum_{x \in X} f_x u^{\pi}(x)$$

with $f_x \in \Lambda_i$ for all $x \in X$. Similarly, every $g \in k[A_i]$ can be written as

$$g = \sum_{y \in Y} g_y u^y$$

with $g_y \in \Lambda_i$ for all $y \in Y$. As the actions of Λ_i on $\mathbb{Z}^r/d_i$ and $\mathbb{Z}^r/d_i'$ are free, these representations are unique and the modules can be viewed as free $k[\Lambda_i]$-modules with bases X and Y respectively.

Moreover, the quotient map $A_{s,i} \to A_i$ is compatible with the Λ_i-actions, so $s_i : k[A_{s,i}] \to k[A_i]$ is $k[\Lambda_i]$-linear. Hence for $f \in k[A_{s_i}]$ as before, we have

$$\begin{aligned} s_i(f) &= s_i(\textstyle\sum_{x \in X} f_x u^{\pi}(x)) \\ &= \textstyle\sum_{x \in X} f_x s_i(u^{\pi}(x)) \\ &= \textstyle\sum_{x \in X} f_x u^{y_x + \lambda_x} \\ &= \textstyle\sum_{y \in Y} \Big(\sum\limits_{\substack{x \in X \\ y_x = y}} f_x u^{\lambda_x} \Big) u^y. \end{aligned}$$

As the $y \in Y$ are a system of representatives for that Λ_i-action on A_i, this last sum will be 0 if and only if each summand is 0. In other words, we have $f \in \ker(s_i)$ if and only if for all $y \in Y$ we have

$$\sum_{\substack{x \in X \\ y_x = y}} f_x u^{\lambda_x} = 0.$$

The generators of $\ker(s_i)$ given in the lemma are deduced easily from these conditions. First of all, note that for $x, x' \in X$ with $y_x = y_{x'}$ the element

$$z_{x,x'} = u^{\lambda_{x'}} u^{\pi}(x) - u^{\lambda_x} u^{\pi}(x')$$

is indeed in $\ker(s_i)$. Now pick for each element $y \in Y$ an element $x_y \in X$ such that x_y maps to y under the quotient map $A \to A_i$ and let $\lambda_y = \lambda_{x_y}$ for all $y \in Y$.

Let V be the submodule of $k[A_{s,i}]$ of elements of the form

$$v = \sum_{y \in Y} v_y \pi(x_y).$$

For $v \in V$ we have

$$s_i(v) = \sum_{y \in Y} (v_y u^{\lambda_y}) u^y$$

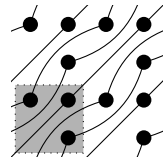
and so $s_i(v) = 0$ if and only if $v_y = 0$ for all $y \in Y$, i.e., if and only if $v = 0$.

Let $f = \sum_{x \in X} f_x u^{\pi(x)}$ be in $\ker(s_i)$. Let $x \in X$ and let $y = y_x$. If $x = x_y$, then $v_x = f_x u^{\pi(x)}$ is in V. If $x \neq x_y$, put $v_x = (f_x u^{\lambda_x - \lambda_y}) u^{\pi(x_y)}$. Then v_x is in V and we have

$$\begin{aligned} f_x u^{\pi(x)} &= (f_x u^{-\lambda_y})(u^{\lambda_y} u^{\pi(x)} - u^{\lambda_x} y^{\pi(x_y)}) + (f_x u^{\lambda_x - \lambda_y}) u^{\pi(x_y)} \\ &= (f_x u^{-\lambda_y}) z_{x,x_y} + v_x. \end{aligned}$$

Let $v = \sum_{x \in X} v_x$, then we see that $f - v$ is in the sub-$k[\Lambda_i]$-module of $k[A_{s,i}]$ generated by the z's, so in particular, we have $s_i(f - v) = 0$, and so v is in $\ker(s_i)$. But then $v = 0$, so f is in the submodule generated by the z's. Since $f \in \ker(s_i)$ was arbitrary, we conclude that the z's are indeed generators of $\ker(s_i)$. □

Example 5.3.8. We continue with the setup from example (5.3.4) and consider s_3. A system of representatives for the Λ-action on A is given by $X = \{(0,1), (1,0), (1,1)\}$. Looking at the picture with the split lines we see a $1-1$ correspondence between the point of X and the lines in direction d_3 up to Λ-translation.



The straight lines correspond to $(1,1)$, the lines that bend upward correspond to $(0,1)$ and the lines that bend downward to $(1,0)$. This is in keeping with the result from lemma (5.3.6).

When we go back to the original reconstruction system, the upward and downward bending lines are added together to create straight lines that go through points from both classes $(0,1)$ and $(1,0)$. It is important to note that what we should add together is the upward bending line through $(0,1)$ and the downward bending line through $(1,2)$, not the one through $(1,0)$. This shift by $(2,0)$ is kept track of by the $u^{\lambda_{x'}}$ and u^{λ_x} in lemma (5.3.7). Applied to this example, the result of that lemma is that $\ker(s_3)$ is generated by the vector in $k[A_{s,3}]$ that assigns 1 to the line through $(0,1)$ and -1 to the line through $(1,2)$ (i.e. the translate of the line $(1,0)$ by $(2,0) \in \Lambda$), together with all its translates by Λ (or Λ / d_3' as translation by d_3' maps all lines to themselves).

The description of p_s from lemma (5.3.5) and the description of $\ker(s)$ from lemma (5.3.7) are the ingredients we need to prove the algorithmic result for the cokernel. This theorem holds for $k = \mathbb{Z}$ and D primitive.

Proof of (5.1.7). Consider the commutative diagram

$$\begin{array}{ccc} \ker(s) & & \\ \downarrow & \searrow & \\ \bigoplus_i \mathbb{Z}[A_{s,i}] & \twoheadrightarrow & \operatorname{cok}(p_s) \\ {\scriptstyle s}\downarrow & & \downarrow \\ \bigoplus_i \mathbb{Z}[A_i] & \twoheadrightarrow & \operatorname{cok}(p) \end{array}$$

From it, we see that $\operatorname{cok}(p)$ is the quotient of $\operatorname{cok}(p_s)$ by the image of $\ker(s)$ in that space. We know from lemma (5.3.5) that $\operatorname{cok}(p_s)$ is a free $\mathbb{Z}$-module of finite rank, as it is a direct sum of copies of $\operatorname{cok}(q)$, which is free of finite rank by theorem (3.5.1). This will be the module $\mathbb{Z}^N$ from the theorem.

Our algorithm has three steps. We state them first, then describe each in detail below.

1. Describe an explicit map

$$\psi : \bigoplus_i \mathbb{Z}[A_{s,i}] \longrightarrow \mathbb{Z}^N$$

   that factors via $\operatorname{cok}(p_s)$ and gives an isomorphism $\operatorname{cok}(p_s) \cong \mathbb{Z}^N$.

2. Give an explicit section

$$t : \bigoplus_i \mathbb{Z}[A_i] \longrightarrow \bigoplus_i \mathbb{Z}[A_{s,i}]$$

   of s and use it to describe $\phi = \psi \circ t$.

3. Give generators for $M = \psi(\ker(s))$ as a subgroup of $\mathbb{Z}^N$.

For the first step, note that the proof of lemma (2.3.10) gives an explicit isomorphism of p_s with $\bigoplus_{x \in A/\Lambda} q_x$ and each q_x is the grid reconstruction system over $\mathbb{Z}$ corresponding to $(\mathbb{Z}^2, E)$. On the space $\bigoplus_i \mathbb{Z}[A_{s,i}]$ the isomorphism comes down to deciding for each line in which orbit in $A_{s,i}/\Lambda_i$ it falls. Thus

there is an algorithm that takes as its input an $f \in \bigoplus_i \mathbb{Z}[A_{s_i}]$ and outputs the corresponding $f_x \in \bigoplus_i \mathbb{Z}[\mathbb{Z}^2/e_i]$ for $x \in A/\Lambda$.

It follows from corollary (4.3.2) and the discussion on evaluating dependencies in the proof of corollary (3.5.7) that we can write down in an explicit, finite way a map $\bigoplus_i \mathbb{Z}[\mathbb{Z}^2/e_i] \longrightarrow \mathbb{Z}^{\mathrm{rk}(\mathrm{cok}(q))}$ that yields an isomorphism $\mathrm{cok}(q) \cong \mathbb{Z}^{\mathrm{rk}(\mathrm{cok}(q))}$ and that this map can be evaluated algorithmically. Combined with the decomposition of p_s as a direct sum of copies of q we see that there is an explicit map

$$\psi : \bigoplus_i \mathbb{Z}[A_{s,i}] \longrightarrow \mathbb{Z}^N$$

where $N = \#(A/\Lambda)\mathrm{rk}(\mathrm{cok}(q))$ that induces an isomorphism $\mathrm{cok}(p_s) \cong \mathbb{Z}^N$. Moreover, there is an algorithm to evaluate this map on any element of $\bigoplus_i \mathbb{Z}[A_{s,i}]$.

For the second step, we construct the section on each component. Let i be in $\{1, \ldots, n\}$ and pick a system of representatives Y for the Λ_i action on A_i. For each $y \in Y$ pick an $x_y \in A_{s,i}$ that maps to y. Every $a \in A_i$ can be written in a unique way as $y + \lambda$ with $y \in Y$ and $\lambda \in \Lambda_i$. A lift of a to $A_{s,i}$ is then given by $x_y + \lambda$. In this way we obtain a section $\tau : A_i \to A_{s,i}$ of the quotient map and there is an algorithm that on input a computes $\tau(a)$. Let t_i be the extension of τ to $\mathbb{Z}[A_i] \to \mathbb{Z}[A_{s,i}]$ and let $t = (t_1, \ldots, t_n)$. Then t is a section of s and there is an algorithm that evaluates this map.

It is easily verified that for fixed A and D, after computing the explicit presentation of the generators of $\mathrm{Dep}(q)$, the algorithms to evaluate $\phi = \psi \circ t$ run in polynomial time in the length of the input.

For the final step, we need an algorithm that gives generators for

$$M = \psi(\ker(s))$$

as a subgroup of $\mathbb{Z}^N$. Lemma (5.3.7) gives a set of generators of $\ker(s_i)$ as a $\mathbb{Z}[\Lambda_i]$-module for every i. Therefore, translating these generators by all elements of Λ_i, we obtain an infinite set of generators for $\ker(s_i) \subset \mathbb{Z}[A_i]$. For each of these we can compute its image under ψ as noted before, so we can produce an infinite set of generators for M. What remains to be shown is that there is an explicit finite number of elements of Λ_i such that they already generate the image of $\ker(s_i)$ under ψ.

For this we consider the image of $\ker(s_i)$ in

$$\bigoplus_{x \in A/\Lambda} \mathbb{Z}[\mathbb{Z}^2/e_i]$$

using the decomposition $p_s = \bigoplus_{x \in A/\Lambda} q$ as before. The generators of $\ker(s_i)$ as a $k[\Lambda_i]$-module map to generators of the image as a $\mathbb{Z}[\mathbb{Z}^2/e_i]$-module. Note that $\mathbb{Z}^2/e_i$ is isomorphic to $\mathbb{Z}$ as each e_i is primitive, and so $\mathbb{Z}[\mathbb{Z}^2/e_i]$ can be seen as a Laurent polynomial ring in one variable, $\mathbb{Z}[w, w^{-1}]$.

By theorem (3.1.5) we know that $\mathrm{im}(q)$ contains the $\mathbb{Z}[\mathbb{Z}^2/e_i]$-ideal generated by $g_i = p_{e_i}(F_i)$ from that theorem. This polynomial g_i is perfectly explicit and can be computed easily. Moreover, it was shown in the proof of theorem (3.5.5) that g_i has leading and trailing coefficients equal to ± 1. Hence every element of $\mathbb{Z}[w, w^{-1}]$ can be written as a multiple of g_i plus an element whose support is contained in $\{0, \ldots, r_i - 1\}$, where r_i is as in theorem (3.5.5).

The upshot of this is that when we compute the image of $\ker(s_i)$ inside each copy of $\mathrm{cok}(q)$, we only have to multiply the generators as a $k[w, w^{-1}]$-module with powers of w from 0 up to $r_i - 1$. Any other element will differ by a multiple of g_i from a linear combination of these and therefore will have the same image in $\mathrm{cok}(q)$.

This completes our proof for the final step: for every $i \in \{1, \ldots, n\}$, there is an algorithm to compute finitely many generators of $\ker(s_i)$ as a $\mathbb{Z}[\Lambda_i]$-module. Moreover, there exists an explicit finite subset of Λ_i such that the translates of the generators by elements of this subset will generate the image of $\ker(s_i)$ in $\mathrm{cok}(p_s)$ as an abelian group. For each of these finitely many elements of $\ker(s_i)$ we can compute their image under ψ as described before, yielding a finite set of generators for M. □

Example 5.3.9. Continue with $A = \mathbb{Z}^2 \setminus 2\mathbb{Z}^2$, $\Lambda = 2\mathbb{Z}^2$ and directions $d_1 = (0, 1)$, $d_2 = (1, 0)$ and $d_3 = (1, 1)$ as in example (5.3.4). An isomorphism $\Lambda \to \mathbb{Z}^2$ is given by sending (x, y) to $(x/2, y/2)$. Conveniently, using this isomorphism we have $e_i = d_i$ for all $i \in \{1, 2, 3\}$.

Let q be the grid reconstruction system over $\mathbb{Z}$ associated to $(\mathbb{Z}^2, D)$. From theorem (3.5.1) we conclude that $\mathrm{Dep}(q)$ is a free $\mathbb{Z}$-module of rank 3. Using the methods from chapter 4, we can compute a basis for this space.

One such basis, for example is given by the following 3 dependencies.

$$\begin{array}{ll} \alpha: & \sum_{i\in\mathbb{Z}} h_i - \sum_{i\in\mathbb{Z}} v_i \\ \beta: & \sum_{i\in\mathbb{Z}} h_i - \sum_{i\in\mathbb{Z}} t_i \\ \gamma: & \sum_{i\in\mathbb{Z}} ih_i - \sum_{i\in\mathbb{Z}} iv_i + \sum_{i\in\mathbb{Z}} it_i \end{array}$$

In these formulas the v_i represent the vertical line sums (corresponding to d_1), the h_i the horizontal line sums, and the t_i the anti-diagonal ones. The numbering is such that the point $(x, y) \in \mathbb{Z}^2$ lies on the x-th vertical line, the y-th horizontal line and the $(y - x)$-th anti-diagonal line.

For $i \in \mathbb{Z}$, let V_i, H_i and T_i be the elements of $\bigoplus_{j=1}^{3} \mathbb{Z}[\mathbb{Z}^2/d_j]$ that respectively send the i-th line in directions d_1, d_2 and d_3 to 1 and all others to 0. Together these form a $\mathbb{Z}$-basis for $\bigoplus_{j=1}^{3} \mathbb{Z}[\mathbb{Z}^2/d_j]$.

The grid reconstruction system p_S over $\mathbb{Z}$ associated to (A, D) is the sum of three copies of q. We label these copies by the elements of

$$X = \{(0, 1), (1, 0), (1, 1)\},$$

which is a system of representatives for the Λ-action on A. For $x \in X$, let q_x be the copy of q where the grid consists of the points $x + \Lambda$, which we identify with $\mathbb{Z}^2$ by putting the origin at x.

The space of dependencies $\mathrm{Dep}(p_s)$ is free of rank $3 \cdot 3 = 9$, a basis for this space is given by α_x, β_x and γ_x for $x \in X$. Together, these give an explicit map

$$\phi : \bigoplus_{i=1}^{3} \mathbb{Z}[A_{s,i}] \longrightarrow \mathbb{Z}^9.$$

The next step is to compute the module $M = \phi(\ker(s))$. Using lemma (5.3.7) we conclude that $\ker(s)$ is generated as a $\mathbb{Z}$ module by

$$\begin{array}{c} V_{(1,0),i} - V_{(1,1),i} \\ H_{(0,1),i} - H_{(1,1),i} \\ T_{(0,1),i} - T_{(1,0),i+1} \end{array}$$

for $i \in \mathbb{Z}$. Their images under ϕ are given in the following table.

	$V_{(1,0),i} - V_{(1,1),i}$	$H_{(0,1),i} - H_{(1,1),i}$	$T_{(0,1),i} - T_{(1,0),i+1}$
$\alpha_{(0,1)}$	0	1	0
$\beta_{(0,1)}$	0	1	-1
$\gamma_{(0,1)}$	0	i	i
$\alpha_{(1,0)}$	-1	0	0
$\beta_{(1,0)}$	0	0	-1
$\gamma_{(1,0)}$	$-i$	0	$-(i+1)$
$\alpha_{(1,1)}$	1	-1	0
$\beta_{(1,1)}$	0	-1	0
$\gamma_{(1,1)}$	i	$-i$	0

Using this, one easily sees that M is a free submodule of rank 5 of $\mathbb{Z}^9$ and that the quotient $\mathbb{Z}^9/M$ is free of rank 4. The following linear combinations of our basis of $\mathrm{Dep}(p_s)$ represent a basis for $\mathrm{Dep}(p)$.

$$\begin{array}{ll} \alpha_{(1,0)} + \alpha_{(1,1)} - \beta_{(1,1)} & \gamma_{(0,1)} - \beta_{(1,0)} + \gamma_{(1,0)} + \gamma_{1,1} \\ \beta_{(0,1)} - \beta_{(1,0)} + \beta_{(1,1)} & \alpha_{(0,1)} + \beta_{(1,1)} \end{array}$$

One verifies easily that these indeed map all the generators of $\ker(s)$ to 0 and therefore assign well-defined weights to the lines in A_1, A_2 and A_3.

Proof of (5.1.8). Suppose we are given input $f \in \bigoplus_i \mathbb{Z}[A_i]$. As mentioned in the proof of theorem (5.1.7), the map

$$\phi : \bigoplus_i \mathbb{Z}[A_i] \longrightarrow \mathbb{Z}^N$$

from that theorem can be evaluated in polynomial time in the input length, once the pre-computation of a basis for $\mathrm{Dep}(q)$ has been done. This can be done beforehand, since A and D are not part of the input of the algorithm we are currently describing. Hence we can compute $v = \phi(f) \in \mathbb{Z}^N$ in polynomial time in the input length. In particular the length of v is polynomial in the input length. This is the first step of our algorithm.

Like the pre-computations for ϕ, the computation of a basis for M can be done beforehand and is not part of the current algorithm. What we have left to do is decide if $v \in M$. Given v and a basis for M this can be done in polynomial time in the input length using standard algorithms for linear algebra over $\mathbb{Z}$. □

5.4 Some open questions

The theory presented in this chapter is far from complete. In this section, we briefly consider some questions that are unanswered and some strategies to possibly tackle them.

For the purpose of the discussion we carry over the notation from the rest of the chapter. Let A be a Λ-periodic subset of $\mathbb{Z}^r$ and let D be a 2-regular sequence of directions in $\mathbb{Z}^r$. Let k be a commutative ring and let p be the grid reconstruction system over k associated to (A, D).

The method for computing $\ker(p)$ is strongly tied to Gröbner basis algorithms to compute syzygies. For simplicity, a result over fields was presented. However, Gröbner bases and syzygy algorithms for polynomial rings over more general rings k have been studied extensively. See, e.g. [18] for an overview of some of the early work on this subject and an algorithm for computing syzygies for more general k. It would therefore seem that the algorithm from theorem (5.1.5) can easily be extended to more general rings.

Within this context of computing syzygies, the requirement that k be a Noetherian ring seems to come up naturally. It is the theoretical condition that guarantees the modules we consider will be finitely generated. Yet, we have not seen such a requirement when looking at other kinds of grid reconstruction systems. Is it possible that also in in this case, the kernels are finitely generated over all k? Or is this not true, and can one show an example of A, D and k for which $\ker(p)$ is not finitely generated?

For the grid reconstruction systems considered in chapter 3, the $\ker(p)$ over k was shown to be the base change of the kernel over $\mathbb{Z}$ for all k. For finite grids, this was related to the freeness of the cokernel by theorem (2.5.5). However, we cannot prove that $\mathrm{cok}(p)$ is always free for these periodic maps, and it may well not be. So one might expect the kernel to be different over certain rings k. Restricting ourselves to fields $\mathbb{F}_q$ for q prime, can we say something about when $\ker(p)$ for $k = \mathbb{F}_q$ is not the reduction mod q of the kernel over $\mathbb{Z}$? It is possible that these q's can be related to certain invariants of the $k[\Lambda]$-linear map ϕ from lemma (5.2.1). As these q's are related to the torsion in $\mathrm{cok}(p)$, it is interesting to bound them, or be able to compute them, or show that they do not exist in certain cases.

We saw in lemma (5.3.1) that the technique we used to compute the kernel can sometimes be used to show that the cokernel for a periodic grid set is the same as that for the full grid. It may well be possible to give more general criteria for when this is the case. Perhaps an algorithm can be found

that on input A, Λ and D decides if the morphism $\mathrm{cok}(p) \to \mathrm{cok}(p_f)$ is an isomorphism.

The description of $\mathrm{cok}(p)$ from theorem (5.1.7) is not very precise. For example, the bound on the rank of $\mathrm{cok}(p)$ is very large if $\mathbb{Z}^2 \setminus A$ is very sparse, while we expect from lemma (5.3.1) that it is often simply the rank of the cokernel of the corresponding full grid system.

Using the output of theorem (5.1.7) we can compute the torsion subgroup of $\mathrm{cok}(p)$. Here too, it would be nice to have some theoretical results constraining the outcome. For example, it would be nice to have conditions under which there is no torsion. One might also look for bounds on the primes at which torsion can occur and the exponent of the torsion subgroup.

Since p_s is a direct sum of copies of the full grid reconstruction system q, it might be worthwhile to look at the reconstruction system corresponding to q as defined in lemma (4.2.4). A certain direct sum of copies of this system will have the same cokernel as p_s, but is of finite dimension over k. This system will split up into local components as described in chapter 4. Is the image of $\ker(s)$ inside this system compatible with this decomposition? And if so, can the linearisation technique used in chapter 4 shed some light on the structure of $\ker(s)$?

Bibliography

[1] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra.* Addison-Wesley Publishing Co., 1969.

[2] K. J. Batenburg and A. Stolk. An algebraic framework for discrete tomography: Revealing the structure of dependencies. *SIAM J. Discrete Math.*, 24(3):1056–1079, 2010. Available on arXiv via `http://arxiv.org/abs/0906.0711`.

[3] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming, Theory and algorithms.* Wiley-Interscience, 2006.

[4] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997.

[5] H. Cartan and S. Eilenberg. *Homological algebra.* Princeton University Press, 1999.

[6] H. Cohen. *A course in computational algebraic number theory.* Springer-Verlag, Berlin, 1993.

[7] B. van Dalen. Dependencies between line sums. Master's thesis, Leiden University, The Netherlands, 2007, `http://www.math.leidenuniv.nl/scripties/DalenMaster.pdf`.

[8] D. Eisenbud. *Commutative algebra, with a view toward algebraic geometry.* Springer-Verlag, New York, 1995.

[9] D Gale. A theorem on flows in networks. *Pacific J. Math.*, 7:1073–1082, 1957.

[10] R. J. Gardner, P. Gritzmann, and D. Prangenberg. On the computational complexity of reconstructing lattice sets from their X-rays. *Discrete Math.*, 202(1-3):45–71, 1999.

[11] S. I. Gelfand and Y. I. Manin. *Methods of homological algebra.* Springer-Verlag, Berlin, 2003.

[12] L. Hajdu and R. Tijdeman. Algebraic aspects of discrete tomography. *J. Reine Angew. Math.*, 534:119–128, 2001.

[13] G. T. Herman and A. Kuba, editors. *Discrete Tomography: Foundations, Algorithms and Applications*. Birkhäuser, Boston, 1999.

[14] P. J. Hilton and U. Stammbach. *A course in homological algebra*. Springer-Verlag, New York, 1997.

[15] D. E. Knuth. *The art of computer programming. Volume 3, Sorting and searching*. Addison-Wesley Publishing Co., 1973.

[16] E. Landau. *Elementary number theory*. Chelsea Publishing Co., New York, N.Y., 1958.

[17] S. Lang. *Algebra*. Springer-Verlag, New York, 2002.

[18] H. M. Möller. On the construction of Gröbner bases using syzygies. *J. Symbolic Comput.*, 6(2-3):345–359, 1988.

[19] D. G. Northcott. *An introduction to homological algebra*. Cambridge University Press, 2008.

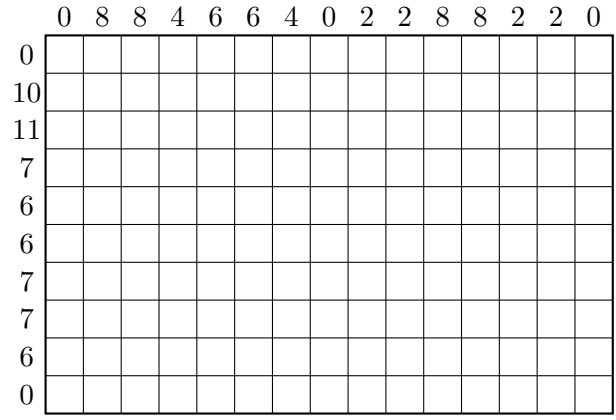
[20] D. S. Passman. *The algebraic structure of group rings*. Wiley-Interscience, New York, 1977.

[21] R. T. Rockafellar. *Convex analysis*. Princeton University Press, 1970.

[22] H. J. Ryser. Combinatorial properties of matrices of zeros and ones. *Canadian J. Math.*, 9:371–377, 1957.

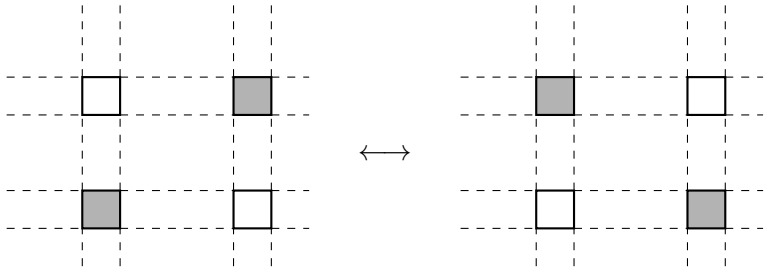
Samenvatting

De figuur hieronder is een logische puzzel. Elk vakje in het rooster moet wit of zwart gemaakt worden. De getallen langs de rand geven aan hoeveel zwarte vakjes in de betreffende rij of kolom moeten staan. Wie bekend is met Japanse puzzels (ook wel beeldzoekers genoemd) ziet waarschijnlijk enige gelijkenis, hoewel men bij zulke puzzels meer informatie krijgt dan alleen het aantal zwarte vakjes in elke rij en kolom.



Het is moeilijk om de bovenstaande puzzel aan te pakken met de oplostechnieken die meestal worden gebruikt voor beeldzoekers of andere logische puzzels. Deze methodes zijn typisch bedoeld voor puzzels met precies één oplossing. Men vult vakjes in waarvan zeker is welke kleur ze moeten krijgen en probeert op deze manier de hele puzzel in te vullen. Puzzels zoals die hierboven hebben vaak meerdere oplossingen, en dan werken deze methodes dus niet. Er zijn immers vakjes die niet in elke oplossing zwart of in elke oplossing wit zijn.

Hoe zien verschillende oplossingen van dezelfde puzzel eruit? Een eenvoudig voorbeeld is te zien in de figuur op de volgende pagina. We zien een rechthoek waarbij twee diagonaal tegenoverliggende hoekpunten zwart zijn en de twee andere wit. In zo'n rechthoek kunnen we de twee diagonalen omwisselen, zonder het aantal zwarte vakjes in de betrokken rijen en kolommen te veranderen.



Verrassend genoeg is dit eigenlijk alles wat er kan gebeuren. Een stelling van H. J. Ryser uit 1957 zegt dat elk tweetal oplossingen van dezelfde puzzel door een aantal van dit soort verwisselingen in elkaar overgevoerd kan worden. Voor wiskundige details en bibliografische verwijzingen, zie sectie 1.1 van dit proefschrift.

Doordat deze puzzels meerdere oplossingen kunnen hebben, moet het oplossen ervan anders worden aangepakt dan men gewoonlijk doet met logische puzzels. Hoe moet het dan wel? Hiervoor gaan we iets doen wat menig puzzelaar verafschuwt: zomaar iets proberen en kijken hoe ver we komen. We gebruiken onderstaande puzzel, die iets handzamer is dan de grote puzzel waarmee we begonnen zijn.

	2	1	3	3	1
1					
4					
4					
1					

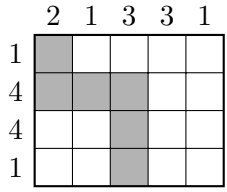
We gaan deze puzzel kolom voor kolom proberen op te lossen.

	2	1	3	3	1
1	■				
4	■				
4					
1					

In de eerste kolom moeten twee zwarte vakjes. Laten we deze in de bovenste twee vakjes zetten.

	2	1	3	3	1
1	■				
4	■	■			
4					
1					

In de tweede kolom moet één zwart vakje. Dit mag niet in de bovenste rij, want het enige zwarte vakje daarin hebben we reeds in de eerste kolom geplaatst. Laten we het dus in de tweede rij plaatsen.

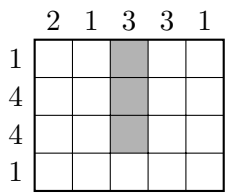


Voor de derde kolom is nu geen keus: er moeten drie zwarte vakjes in, en er zijn nog drie rijen over waar ze in kunnen.

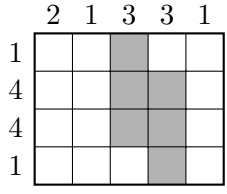
Nu hebben we een probleem. In de vierde kolom moeten drie zwarte vakjes, echter, er mag geen zwart vakje meer in de eerste of laatste rij, dus er is nog plaats voor maximaal twee zwarte vakjes. Wie nog iets beter naar de bovenstaande stappen kijkt, ziet dat er bij de tweede stap al een probleem ontstaat. In de derde rij moeten vier zwarte vakjes komen te staan, maar de eerste twee zijn wit, zodat er maximaal drie in passen.

Voor beide problemen helpt het om iets slimmer te kiezen welke vakjes we zwart maken in een kolom. Kiezen we hiervoor namelijk de vakjes in de rijen waarin nog zoveel mogelijk zwarte vakjes moeten komen, dan slaan we twee vliegen in één klap. Er worden vaak zwarte vakjes geplaatst in die rijen waarin nog veel van zulke vakjes moeten, en het laatste zwarte vakje in een rij wordt alleen opgebruikt wanneer het echt niet anders kan. Verder is er ook nog iets te winnen in de volgorde waarin we de kolommen invullen. Kolommen waarin veel zwarte vakjes moeten, worden steeds lastiger om in te vullen. We beginnen dus met de kolommen waarin de meeste zwarte vakjes moeten en werken zo op volgorde naar de kolommen met de minste zwarte vakjes toe.

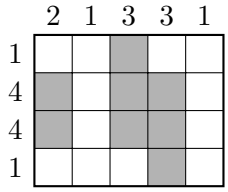
Met deze aangepaste methode lukt het wel om de puzzel op te lossen.



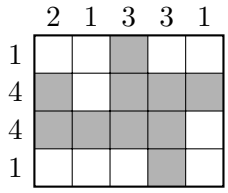
Ditmaal beginnen we met de derde kolom, want daarin moeten de meeste zwarte vakjes. We vullen in elk geval het vakje in de tweede en derde rij, want in deze rijen moeten vier zwarte vakjes worden geplaatst. Het laatste zwarte vakje plaatsen we in de eerste rij.



In de vierde kolom moeten drie vakjes worden gekleurd en er zijn nog drie rijen waarin dat kan, dus de keuze van de vakjes ligt vast.



Ook in de eerste kolom is er geen keus voor welke twee vakjes moeten worden gekleurd.



Er zijn nog twee manieren om de laatste twee kolommen op te vullen. Hier staat er één afgebeeld.

Dat het oplossen met de aangepaste methode wel lukt, is geen toeval. In 1957 hebben de eerder genoemde H. J. Ryser en onafhankelijk van hem ook D. Gale laten zien dat deze methode altijd een oplossing vindt, als er een oplossing bestaat. Wiskundige en bibliografische details zijn wederom te vinden in sectie 1.1.

In feite bewijzen zowel Ryser als Gale nog iets sterkers dan dit. Ze geven elk een aantal numerieke voorwaarden op de getallen langs de rand zodat

- getallen die horen bij een oplosbare puzzel voldoen aan deze voorwaarden; en
- de omschreven oplosmethode werkt als de getallen voldoen aan de voorwaarden.

Deze voorwaarden kunnen worden gecontroleerd zonder de oplosmethode helemaal uit te voeren. Zodoende kunnen we dus controleren of er een oplossing bestaat, zonder zo'n oplossing uit te rekenen.

Een deel van de voorwaarden is eenvoudig zelf te verzinnen. Zo zal elk getal minstens 0 zijn en hoogstens het aantal vakjes in de bijbehorende rij of kolom. Anders is er natuurlijk geen oplossing. Iets subtieler is de volgende observatie. Wanneer we alle getallen die bij de kolommen horen bij elkaar optellen, is de uitkomst het aantal zwarte vakjes in de figuur. Hetzelfde geldt voor de getallen die horen bij de rijen. Deze twee sommen moeten dus hetzelfde zijn, anders kan er geen oplossing bestaan. Er zijn nog meer voorwaarden, maar die zijn te ingewikkeld om hier te bespreken. We verwijzen de geïnteresseerde lezer naar stelling (1.1.13).

Puzzels van het soort waar we nu naar gekeken hebben, zijn fundamentele objecten die worden bestudeerd in de *discrete tomografie*. Dit is het vakgebied

waarbinnen dit proefschrift ook valt. Drie centrale vragen bij het bestuderen van dit soort problemen, zijn hierboven ook al aan bod gekomen:

- **Welke figuren geven dezelfde getallen langs de rand?** Anders gezegd, is de oplossing uniek bepaald, of zijn er meerdere oplossingen? We noemen dit ook wel het probleem van *uniciteit*.

- **Voor welke getallen langs de rand bestaat er een oplossing?** We vermeldden eerder al dat er een aantal numerieke voorwaarden op de getallen langs de rand zijn, die garanderen dat er een oplossing bestaat. Dit vraagstuk noemen we ook wel het probleem van *consistentie*.

- **Hoe vinden we een oplossing van een puzzel?** We hebben hiervoor een oplosmethode (ook wel algoritme genoemd) beschreven. Dit vraagstuk noemen we ook wel het probleem van *reconstructie*.

In de literatuur worden deze vragen en variaties daarop bestudeerd voor de puzzels zoals hierboven beschreven en vergelijkbare puzzels. Twee variaties die het meest relevant zijn voor het werk in dit proefschrift zijn het aanpassen van de basisvorm en het tellen van zwarte vakjes in meer of andere richtingen.

Met de basisvorm bedoelen we de rechthoek van vakjes die ingevuld moeten worden. Men kan deze bijvoorbeeld vervangen door een drie- of zeshoek, of een benadering van een ronde vorm. Ook kan worden gedacht aan een vorm in drie of meer dimensies opgebouwd uit eenheids(hyper)kubusjes in plaats van vierkante vakjes.

Het idee van tellen in meer of andere richtingen spreekt redelijk voor zich. In plaats van het aantal zwarte vakjes in een rij of kolom te tellen, kan men bijvoorbeeld ook kijken naar diagonale lijnen, of lijnen die de richting van een paardensprong (bijvoorbeeld 1 naar rechts, 2 omhoog) volgen. Hierbij is het handiger om aan de puzzel te denken als een verzameling roosterpunten die zwart of wit gekleurd zijn.

Een meer fundamentele aanpassing van het probleem wordt beschreven door L. Hajdu en R. Tijdeman in het artikel [12] uit 2001. In plaats van de vakjes wit of zwart te kleuren, moet er in deze variant van de puzzel in elk vakje een (geheel) getal worden geplaatst. De getallen langs de rand stellen nu de *som* van de getallen in de betreffende rij of kolom voor. Je kan dit zien als een uitbreiding van het originele probleem door de witte vakjes te interpreteren als een 0 en de zwarte als een 1. Het aantal zwarte vakjes in een rij of kolom is dan inderdaad gelijk aan de som van de getallen in die rij.

Dezelfde vragen waarmee we bij het orginele probleem bezig waren, kunnen we ook in deze situatie stellen. Als eenvoudig voorbeeld kijken we hieronder naar een rechthoekige basisvorm waarbij de rij- en kolomsommen gegeven worden. Deze situatie lijkt het meest op de puzzels waarmee we begonnen zijn. In de onderstaande puzzel is ook te zien dat negatieve getallen ook zijn toegestaan, ze zijn zelfs nodig om de derde rij op te kunnen lossen.

	2	5	1	2	4
0					
9					
−1					
6					

Het oplossen van deze puzzel gebeurt in een aantal stappen.

	2	5	1	2	4
0					
9					
−1					
6					

In de vakjes die hiernaast zijn aangegeven kunnen we alles invullen wat we willen.

	2	5	1	2	4
0	−5	4	−2	1	
9	1	−2	7	3	
−1	2	0	−4	2	
6					

Voor de vakjes die hiernaast zijn aangegeven is er precies één manier om ze in te vullen, de som van de getallen in elke rij en kolom ligt immers al vast.

	2	5	1	2	4
0	−5	4	−2	1	2
9	1	−2	7	3	0
−1	2	0	−4	2	−1
6	4	3	0	−4	

Nu is er nog één vakje over om in te vullen. Dit vakje is lastiger dan de anderen: er zijn twee voorwaarden waaraan voldaan moet worden: zowel de laatste rijsom als de laatste kolomsom moeten kloppen. Het blijkt in dit geval dat we door 3 in te vullen aan beide voorwaarden voldoen.

Vullen we het laatste getal in, dan krijgen we de oplossing van de puzzel die hieronder staat. Deze oplossing is duidelijk niet uniek: er was in de eerste stap van de oplosmethode hierboven een heleboel vrijheid.

	2	5	1	2	4
0	−5	4	−2	1	2
9	1	−2	7	3	0
−1	2	0	−4	2	−1
6	4	3	0	−4	3

Wanneer lukt het om met deze stappen een puzzel op te lossen? Alleen in de laatste stap was er mogelijk een probleem, waardoor er misschien geen oplossing zou bestaan. Verrassend genoeg hangt het bestaan van de oplossing niet af van de getallen die we bij de eerste stap hebben gekozen. Het gaat of voor alle keuzes goed, of voor alle keuzes fout.

Een berekening laat zien dat de bovenstaande methode tot een oplossing leidt mits de som van de getallen die bij de rijen staan, gelijk is aan de som van de getallen die bij de kolommen staan. Het is gemakkelijk in te zien dat dit ook een *noodzakelijke* voorwaarde is, dat wil zeggen, dat elke puzzel met een oplossing aan deze voorwaarde moet voldoen. Wanneer we namelijk de som van de getallen bij de rijen nemen, is dit de som van de sommen van de rijen van een oplossing, en dus is het precies de som van alle getallen in zo'n oplossing. Hetzelfde is waar voor de som van de getallen bij de kolommen. Deze sommen zijn dus inderdaad aan elkaar gelijk.

Hiermee hebben we het antwoord gegeven op één van de drie centrale vragen die eerder werden genoemd. Een puzzel is consistent (dat wil zeggen, heeft een oplossing) precies wanneer voldaan is aan de gelijkheid uit de vorige alinea: de som van de getallen bij de rijen is gelijk aan de som van de getallen bij de kolommen.

De methode die we hebben beschreven voor het oplossen van de puzzels is een antwoord op de reconstructie-vraag. Deze methode geeft immers voor elke oplosbare puzzel een oplossing.

Ook de vraag over uniciteit is eenvoudig te beantwoorden. In de eerste stap van de oplosmethode kon een gedeelte van de tabel vrij gekozen worden. Daarna lagen de waardes van de andere getallen uniek vast. Zoals reeds opgemerkt hangt het bestaan van de oplossing niet af van de gemaakte keuzes, dus voor elke keuze is er precies één oplossing.

In dit proefschrift worden varianten bestudeerd van het type puzzel dat we zojuist hebben besproken. Net als bij de puzzels die we eerder zagen kan worden afgeweken van de basisvorm (de rechthoekige tabel) en kunnen we sommen nemen langs lijnen die niet horizontaal of verticaal lopen.

Curriculum Vitae

Personalia

Naam	Arjen Pieter Stolk
Geboren	17 december 1982, te 's-Gravendeel

Opleidingen

2006 – 2010	Promovendus by prof. dr. S.J. Edixhoven, Mathematisch Instituut, Universiteit Leiden
2001 – 2006	Doctoraal wiskunde, cum laude, Universiteit Leiden
2001 – 2002	Propedeuse informatica, Universiteit Leiden
1995 – 2001	VWO (Gymnasium), Johan de Witt Gymnasium, Dordrecht

Huidige werkgever

TOPdesk Nederland
Martinus Nijhofflaan 2, 2624 ES Delft