INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
75-26,571

DONALDSON, John Laurence, 1950-
MINKOWSKI REDUCTION OF INTEGRAL MATRICES.

The Ohio State University, Ph.D., 1975
Mathematics

Xerox University Microfilms, Ann Arbor, Michigan 48106
MINKOWSKI REDUCTION OF INTEGRAL MATRICES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the Graduate

School of The Ohio State University

By

John Laurence Donaldson, B.S., M.S.

* * * * *

The Ohio State University

1975

Reading Committee:

Hans Zassenhaus
Bostwick Wyman
Alan Woods
Bohuslav Divis

Approved By

Hans Zassenhaus
Adviser
Department of Mathematics
ACKNOWLEDGMENTS

I wish to express my deep gratitude to Hans Zassenhaus for directing me to this problem, and for his helpful suggestions as I worked on it. I also would like to thank David Ford for making available to me his work on the subject, and Merle Owdom and the IRCC for their help. I especially want to thank Bostwick Wyman for his close attention, interest, and assistance throughout the past two years.
VITA

May 26, 1950 . . . . Born - Cleveland, Ohio

1971 . . . . . . . B.S., Case Western Reserve University, Cleveland, Ohio

1971-1975 . . . . Teaching Associate, Department of Mathematics, The Ohio State University, Columbus, Ohio

1972 . . . . . . . M.S., The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Mathematics

   Studies in Number Theory. Professor Alan Woods

   Studies in Analysis. Professor Bogdan Baishanski

   Studies in Algebra. Professor Joseph Ferrar
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. THEORETICAL BACKGROUND</td>
<td>2</td>
</tr>
<tr>
<td>Reduction</td>
<td></td>
</tr>
<tr>
<td>The Hermite normal form</td>
<td></td>
</tr>
<tr>
<td>Matrices and modules</td>
<td></td>
</tr>
<tr>
<td>Gauge functions</td>
<td></td>
</tr>
<tr>
<td>Minkowski reduction</td>
<td></td>
</tr>
<tr>
<td>The Zassenhaus-Ford conjecture</td>
<td></td>
</tr>
<tr>
<td>II. A THEOREM ON MINKOWSKI REDUCTION OF MATRICES</td>
<td>9</td>
</tr>
<tr>
<td>The probability model</td>
<td></td>
</tr>
<tr>
<td>The reduction inequalities</td>
<td></td>
</tr>
<tr>
<td>Probability distributions of vectors in $\mathbb{R}^{2n}$</td>
<td></td>
</tr>
<tr>
<td>Estimates for $H_n(t), G_n(t)$</td>
<td></td>
</tr>
<tr>
<td>The conditional probability of reduction</td>
<td></td>
</tr>
<tr>
<td>The other probability models</td>
<td></td>
</tr>
<tr>
<td>III. EXTENDED REDUCTION</td>
<td>22</td>
</tr>
<tr>
<td>Reduction by submatrices</td>
<td></td>
</tr>
<tr>
<td>The extension of the gauge function</td>
<td></td>
</tr>
<tr>
<td>The gauge function on submatrices</td>
<td></td>
</tr>
<tr>
<td>Reduction by submodules</td>
<td></td>
</tr>
<tr>
<td>$(l_1,\ldots,l_p)$ reduction of matrices</td>
<td></td>
</tr>
<tr>
<td>The $(1,m-1)$ reduction scheme</td>
<td></td>
</tr>
<tr>
<td>APPENDIX</td>
<td>34</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>35</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Percentage of reduced matrices .............. 34
INTRODUCTION

In 1905 Hermann Minkowski introduced his theory of reduction of positive definite quadratic forms. [1] This theory is one of the essential foundations of the geometry of numbers. Recently Hans J. Zassenhaus has suggested that Minkowski reduction can be applied to the problem of row reduction of matrices of integers. [2] It is the study and development of this idea that forms the basis of this dissertation; emphasis is placed particularly on the reduction algorithm, as adapted to machine computation. Chapter one is devoted to the theoretical foundations of the subject, including an outline of Minkowski's original work. Statistical methods are used in the second chapter to examine the algorithm, and the result is the main theorem of the dissertation. This theorem gives a relationship between the dimensions of a matrix and the probability that it is reduced. In chapter three, motivated by the theorem, a modified version of Minkowski reduction is defined and developed.
CHAPTER ONE
THEORETICAL BACKGROUND

1.1 Reduction

Reduction, in the sense we use here, comes from the idea of an equivalence relation on a set, partitioning it into equivalence classes. Reduction means the selection of a canonical representative of each equivalence class. For example, the numbers 0, 1, ..., n-1 form a set of representatives of the equivalence classes modulo n. Here we are concerned with integral matrices. The equivalence relation we deal with is as follows: Two matrices A and B are equivalent if and only if there exists an m x n unimodular integral matrix U such that A = UB. Our aim then is to define a canonical representative, the "reduced" matrix, in each class, and to provide an algorithm for finding the reduced matrix equivalent to a given matrix.

1.2 The Hermite normal form

The usual method of reduction of integral matrices was introduced by Hermite in 1851. [3] Among more recent accounts of his theory is the one by MacDuffee. [4] Hermite showed that every matrix is unimodularly equivalent to one in upper triangular form, such that each entry above the diagonal is bounded by one-half the magnitude of the diagonal entry directly below it. Such a matrix is said to be in Hermite normal form.

Several algorithms to compute the Hermite normal form of a given matrix have been published, all of them similar. [5, 6, 7]
Implementation of these algorithms for machine computation suffer one major drawback, that of overflow. The computer investigations by Zassenhaus and David Ford verified the fact that the numbers generated by the reduction procedure can increase rapidly in magnitude, beyond the capacity of the typical machine, even when the matrices to be reduced have as few as three or four rows and entries bounded in magnitude by 100.

1.3 Matrices and modules

Before we pursue our alternative to Hermitian reduction, we shall mention the connection between a matrix and its associated Z-module, first developed by Chatelet. Suppose \( A \) is an \( m \times n \) real matrix, with rows \( a_1, a_2, \ldots, a_m \). With addition and scalar multiplication defined component-wise, the Z-module generated by \( a_1, \ldots, a_m \) is called the row module of \( A \). \( a_1, \ldots, a_m \) form a basis for the row module if and only if \( A \) is nonsingular. If \( A \) is nonsingular, there is a one-to-one correspondence between matrices unimodularly equivalent to \( A \) (modulo the ordering of the rows) and bases of the row module. Thus reduction is equivalent to selecting a canonical basis for the row module. It is this correspondence which leads us to develop the theory of Minkowski reduction in terms of modules.

Remark The row module of \( A \) may be thought of as a lattice in real \( n \)-dimensional space. It is here that the connection with the geometry of numbers lies. Minkowski's work is done in terms of lattices.
Remark Although the primary concern is with integral matrices, the reduction theory is valid for real matrices. Most of the work following will be in this more general setting.

1.4 Gauge functions

The basic idea of Minkowski reduction is to choose a matrix from each equivalence class whose rows are as short as possible, according to some definition of length. Minkowski used the common Euclidean length; this was generalized to the gauge function by Weyl. [9] Since we can consider any \(\mathbb{Z}\)-module of row vectors as a subset of a real \(n\)-dimensional vector space, we shall define gauge functions on real vector spaces.

Definition Let \(V\) be a real vector space. A gauge function on \(V\) is a function \(f:V \to \mathbb{R}\) such that for \(x, y \in V\) we have:

1. \(f(x) > 0\), except \(f(0) = 0\)
2. \(f(tx) = |t|f(x)\) for all \(t \in \mathbb{R}\)
3. \(f(x+y) \leq f(x)+f(y)\)

With \(f\) as the norm, \(V\) becomes a finite-dimensional normed linear space. It follows that all gauges on \(V\) generate the same topology. [10]

By means of the correspondence between matrix rows and elements of the row module we now have a means by which to measure the length of rows.

1.5 Minkowski reduction

We shall define a Minkowski reduced matrix by first considering the row module \(M\). We assume we are given a gauge
Definition A set of elements of $M$, $x_1, \ldots, x_k$, is called a primitive system if it can be extended to a basis of $M$.

Definition A basis of $M$ is called Minkowski reduced if the following properties are satisfied:

1. $x_1$ is a shortest element of $M$.

2. $x_k$ ($k=2, \ldots, m$) is a shortest element among all elements of $M$ which together with $x_1, \ldots, x_{k-1}$ form a primitive system.

This definition is equivalent to the following:

Definition A basis of $M$ is called Minkowski reduced if the following condition is satisfied:

$$f(x_k) \leq f(s_1 x_1 + s_2 x_2 + \cdots + s_m x_m)$$

for all systems $s_1, \ldots, s_m$ of integers with greatest common divisor $(s_k, \ldots, s_m)$ equal to one.

The second definition is most useful for matrices, for the inequalities (1.5.1) can be checked once we are given the matrix and the function $f$. The algorithm for reducing a matrix is based on the first so-called finiteness theorem:

Theorem When the gauge function used is the Euclidean norm $f(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2$, then we can select finitely many of the inequalities (1.5.1) from which all the rest follow.

The theorem is true only for gauge functions in the form of positive definite quadratic forms. [11]

Thus the algorithm is as follows: We are given a matrix $A$,
with rows \(a_1, \ldots, a_m\). We test all the finitely many inequalities. If all are satisfied, we are done. If not, if suppose 
\[ f(a_k) > f(s_1a_1 + \ldots + s_m a_m), \]
we may apply a unimodular transformation to replace \(a_k\) by \(s_1a_1 + \ldots + s_m a_m\) and leave \(a_1, \ldots, a_{k-1}\) fixed (That this is possible follows from the fact that \((s_1, \ldots, s_m) = 1\)). Eventually all the inequalities must be satisfied, and the matrix will be Minkowski reduced.

**Theorem** If all the inequalities (1.5.1) are strict, the Minkowski reduced basis is unique. If not, there are at most finitely many different such bases.

Since our concern is the determination of a unique representative of each class, this is important. In the case where we must select from several different reduced matrices in a class, Zassenhaus uses the lexicographic ordering. This consideration completes the algorithm.

When \(m = 2, 3,\) or 4, the necessary inequalities are those for which the coefficients \(s_1\) take on the values 0 or \pm 1. In these cases the number of tests is small, and the unimodular transformations used in the reduction steps consist merely of replacing the row \(a_k\) by the row \(s_1a_1 + \ldots + s_m a_m\).

However when \(m \geq 5\), the number of inequalities to be tested begins to increase rapidly. A table of these for \(m \leq 6\) is found in Tammela. When \(m \geq 7\), we may no longer assume \(s_k = 1\) in (1.5.1), and this is another significant complication.
1.6 The Zassenhaus–Ford conjecture

This algorithm for Minkowski reduction of matrices was programmed by David Ford, a student of Zassenhaus, for matrices of 2, 3, or 4 rows and a varying number of columns. The reduction procedure was applied to randomly generated matrices, and the average number of steps required was computed for each dimension pair \( m,n \). Ford's results produced the following unexpected phenomenon: Among matrices with a fixed number of rows, those with a larger number of columns required significantly fewer steps, on the average, to reduce. Zassenhaus and Ford conjectured that there was some provable connection between the number of columns and the average number of steps. This connection is the subject of chapter two.

The conjecture is important because of its implication for the efficiency of the reduction algorithm. It would indicate another advantage of Minkowski reduction as opposed to Hermitian reduction, at least for matrices with more columns than rows. It is clear that no similar phenomenon should occur in the Hermitian case, since only the leftmost square of the matrix is actually used to make decisions in the algorithm, with the rest of the columns merely being carried along.

It is important to note that we must investigate the algorithm from a statistical point of view. Since any matrix can be lengthened by adding columns of zeros, without affecting its
reduction, we cannot expect that any absolute measure of the speed of the algorithm would improve as the number of columns increased. We consider instead the average behavior of matrices.
CHAPTER TWO

A THEOREM ON MINKOWSKI REDUCTION OF MATRICES

In this chapter we assume that the gauge function is the Euclidean norm, $f(x_1, \ldots, x_n) = (x_1^2 + \ldots + x_n^2)^{\frac{1}{2}}$, so the reduction inequalities (1.5.1) may be written as

\[(2.0.1) \quad (s_1 a_1 + \ldots + s_m a_m)^2 \geq a_k^2 \quad \text{for all systems } s_1, \ldots, s_m \]
with $(s_k, \ldots, s_m) = 1$.

We note, in particular, that in a reduced matrix the rows must be in increasing order according to length, that is

\[(2.0.2) \quad a_1^2 \leq a_2^2 \leq \ldots \leq a_m^2 \]

We will assume here that all matrices have already been put in this form.

2.1 The probability model

We are going to investigate the probability that a real $m \times n$ matrix is reduced, and therefore we need to make the set of all such matrices into probability space. We shall consider each matrix as a point in $m n$-dimensional real space. Ideally we would consider a random variable distributed uniformly in $\mathbb{R}^{mn}$, so that probabilities of events $A$ would be defined by

\[P(X \in A) = \frac{m(A)}{m(\mathbb{R}^{mn})} \]

where $m$ denotes Lebesgue measure and $A$ is any measurable set. Obviously we cannot do this, so we have a number of alternatives:

1. Give $X$ some distribution other than uniform; say its density function is $f$. Then
(2.1.1) \[ P(x \in A) = \int_A f(x) \]

2. \( X \) can have a uniform distribution if we think of \( R^m \) as a conditional probability space. \([13]\) Then

(2.1.2) \[ P(X \in A | X \in B) = \frac{m(A \cap B)}{m(B)} \]

The disadvantage is that we can determine only conditional probabilities.

3. Since we are concerned primarily with integer matrices, we might define the density of a set of integer points in \( R^m \) as follows:

(2.1.3) \[ D(A) = \lim_{k \to \infty} \left( \frac{\# \text{ of integer points in } A \cap S_k}{\# \text{ of integer points in } S_k} \right) \]

where \( S_k \) denotes the sphere of radius \( k \) centered at the origin. This \( D \) is not, however, a probability, since it is only finitely additive.

In our computations in the following sections, we investigate directly the conditional probability \( P_{m,n,k} = P(K \cap S_k | S_k) \) as defined by (2.1.2), where \( K \) is the set of non-reduced matrices in \( R^m \). We will show later the connection between \( P_k \) and the probabilities defined in (2.1.1) and (2.1.3).

2.2 The reduction inequalities

The set of reduced matrices, the measure of which is to be computed, is determined by the inequalities (2.0.1). To simplify matters, we make some observations on these inequalities.

In our list of necessary inequalities, we may assume we never
have $s_k = 0$. For if $(s_k, \ldots, s_m) = 1$ and $s_k = 0$, then for some $t > k$, $s_t \neq 0$, $(s_t, \ldots, s_m) = 1$, and $(s_1a_1 + \ldots + s_ma_m)^2 > a_t^2 > a_k^2$.

We may write the inequality (2.0.1) as

$$(2.2.1) \quad \sum_{i=1}^{m} \sum_{j=1}^{m} s_i s_j (a_i a_j) \geq a_k^2$$

or

$$(2.2.2) \quad \sum_{i=1}^{m} s_i^2 a_i^2 - a_k^2 \geq \sum_{i=1}^{m} \sum_{j=1}^{m} s_i s_j (a_i a_j)$$

Inequality (2.2.2) will hold if

$$(2.2.3) \quad \sum_{i=1}^{m} s_i^2 a_i^2 - a_k^2 \geq \sum_{i=1}^{m} \sum_{j=1}^{m} |s_i s_j| \frac{|a_i a_j|}{m s_j^2}$$

To further simplify (2.2.3) we use the following lemma:

**Lemma** If $\frac{a_i a_j}{a_i^2} < \frac{|s_i|}{m |s_j|}$ for all $1 \leq i, j \leq m$, $s_i, s_j \neq 0$, then (2.2.3) is satisfied.

**Proof** Since we have observed that in (2.2.3) $s_k \neq 0$, we may assume without loss of generality that $s_i \neq 0$ for $i = 1, \ldots, m$.

Now, if necessary, relabel so that $a_m^2 > a_i^2$ for $i = 1, \ldots, m-1$, and further that $s_m^2 a_m^2 > \ldots > s_2 a_2^2 > s_1 a_1^2$. Then

$$\sum_{i=1}^{m} \sum_{j=1}^{m} |s_i s_j| \frac{|a_i a_j|}{s_j^2} = 2 \sum_{i=1}^{m} \sum_{j=i+1}^{m} |s_i s_j| \frac{|a_i a_j|}{s_j^2}$$

$$\leq 2 \sum_{i=1}^{m} \sum_{j=i+1}^{m} \frac{s_i a_i^2}{m}$$
\[ \begin{align*}
= & \ 2 \sum_{i=1}^{m-1} \frac{m-1}{m} s_i a_i^2 \\
= & \ \sum_{i=1}^{m-1} \frac{m + (m-2i)}{m} s_i a_i^2 \\
= & \ \sum_{i=1}^{m-1} s_i a_i^2 + \sum_{i=1}^{m-1} \frac{m-2i}{m} s_i a_i^2 \\
= & \ \sum_{i=1}^{m-1} s_i a_i^2 + \sum_{1 \leq i < \frac{m}{2}} \frac{m-2i}{m} s_{i+m} a_i^2 + \sum_{\frac{m}{2} \leq i < m-1} \frac{m-2i}{m} s_{i+m} a_i^2 \\
\end{align*} \]

Letting \( j = m-1 \) in the third summation, the right side becomes

\[ \sum_{i=1}^{m-1} s_i a_i^2 + \sum_{1 \leq i < \frac{m}{2}} \frac{m-2i}{m} s_i a_i^2 + \sum_{\frac{m}{2} \leq i < m-1} \frac{2i-m}{m} s_{m-j} a_{m-j} \]

\[ \sum_{i=1}^{m-1} s_i a_i^2 + \sum_{1 \leq i < \frac{m}{2}} \frac{m-2i}{m} (s_i a_i^2 - s_{i+m} a_i^2) \]

The second summation here is less than or equal to 0, so we have

\[ \sum_{i=1}^{m} \sum_{j=1 \atop j \neq i}^{m} |s_i s_j| |a_i a_j| \leq \sum_{i=1}^{m-1} s_i a_i^2 \]

\[ \leq \sum_{i=1}^{m-1} s_i a_i^2 + (s_{m-m} a_{m-k}) \]

\[ = \sum_{i=1}^{m} s_i a_i^2 - a_{m-k}^2 \]

This is (2.2.3), so the lemma is proved.

**Corollary** If \( \frac{|a_i a_j|}{a_i^2} \leq \frac{|s_i|}{m |s_j|} \) for all \( 1 \leq i, j \leq m \) with \( s_i, s_j \neq 0 \),
for each of the inequalities (2.0.1), then \( A \) is Minkowski reduced.

Let \( M \) be a bound for the coefficients \( a_i \) in the inequalities (2.0.1). Then \( M \) depends on \( m \) but not \( n \).

**Corollary** If \( \frac{|a_i \cdot a_j|}{a_i^2 + a_j^2} \leq \frac{1}{2(mM+1)} \) and \( \frac{a_i^2}{a_i^2 + a_j^2} \geq \frac{mM}{2(mM+1)} \) for all \( 1 \leq i \leq j \leq m \), \( A \) is Minkowski reduced.

**Proof** Multiplying the inequalities gives

\[
\frac{|a_i \cdot a_j|}{a_i^2 + a_j^2} \leq \frac{1}{2(mM+1)} \leq \frac{a_i^2}{a_i^2 + a_j^2} \leq \frac{mM}{2(mM+1)}
\]

or

\[
\frac{|a_i \cdot a_j|}{a_i^2} \leq \frac{1}{mM} \leq \frac{|s_i|}{m|s_j|}
\]

### 2.3 Probability distributions of vectors in \( R^{2n} \)

The last corollary suggests that we examine the probability distributions of the functions \( \frac{|x \cdot y|}{x^2 + y^2} \) and \( \frac{x^2}{x^2 + y^2} \) where \( x \) and \( y \) are row vectors of \( R^n \). According to the discussion in section 2.1, we let

\( (2.3.1) \quad H_n(t) = P\left( \frac{x^2}{x^2 + y^2} \leq t \big| x^2 + y^2 \leq k \right) \)

\[
m(\{ (x,y) : x^2 + y^2 \leq k, \frac{x^2}{x^2 + y^2} \leq t \} )
\]

\[
m(\{ (x,y) : x^2 + y^2 \leq k \})
\]

\( (0 \leq t \leq 1) \)

and
(2.3.2) \( G_n(t) = P\left( t \leq \frac{|x \cdot y|}{x^2 + y^2} \mid x^2 + y^2 \leq k \right) \)

\[
m(\xi(x, y) : x^2 + y^2 \leq k, \ t \leq \frac{|x \cdot y|}{x^2 + y^2})
= \frac{m(\xi(x, y) : x^2 + y^2 \leq k)}{m(\xi(x, y) : x^2 + y^2 \leq k)} \quad (0 \leq t \leq \frac{1}{2})
\]

To compute the resulting multiple integrals we use the following integration formula:

**Lemma** Let \( f \) be an integrable function defined on \((a, b)\). \((0 \leq a \leq b)\)

Then

\[
\iiint_{x_1^2 + \ldots + x_n^2 \leq b} f(x_1^2 + \ldots + x_n^2) \, dx_1 \ldots dx_n
= \int_a^b \int (\frac{1}{w})^n \int f(w) w^{3n-1} \, dw
\]

**Proof** On the left, replace \( x_n \) by \( w \) according to the substitution

\[
w = (x_1^2 + \ldots + x_n^2)(\text{sgn} \ x_n)
\]

Then

\[
dw = 2x_n \text{ sgn} \ x_n \, dx_n
= 2 \left| x_n \right| \, dx_n
\]

So

\[
dx_n = \frac{dw}{2 \left| w \right| - (x_1^2 + \ldots + x_{n-1}^2)}
\]

The left-hand integral becomes

\[
\iiint_{|w| \leq b} f(|w|) \frac{dx_1 \ldots dx_n dw}{2 \left| w \right| - (x_1^2 + \ldots + x_{n-1}^2)}
\]
and by symmetry,

\[
\int_a^b \int_{x_1^2 + \ldots + x_{n-1}^2 \leq w} f(w) dx_1 \cdots dx_{n-1} dw
\]

\[
\sqrt{w - (x_1^2 + \ldots + x_{n-1}^2)}
\]

\[
\int_a^b \int_{x_1^2 + \ldots + x_{n-1}^2 \leq w} dx_1 \cdots dx_{n-1} \frac{1}{w - (x_1^2 + \ldots + x_{n-2}) - x_{n-1}^2} dw
\]

Integrating with respect to \( x_{n-1} \), we get

\[
\int_a^b \int_{x_1^2 + \ldots + x_{n-2}^2 \leq w} f(w) \sin^{-1} \left( \frac{x_{n-1}}{\sqrt{w - (x_1^2 + \ldots + x_{n-2}^2)}} \right) \frac{dx_1 \cdots dx_{n-2}}{\sqrt{w - (x_1^2 + \ldots + x_{n-2}^2)}} dw
\]

\[
= \int_a^b f(w) \int_{x_1^2 + \ldots + x_{n-2}^2 \leq w} \pi dx_1 \cdots dx_{n-2} dw
\]

The multiple integral is now just the volume of an \((n-2)\)-dimensional sphere of radius \( \frac{1}{w} \), so we have

\[
\pi \int_a^b f(w) \left[ \frac{\Gamma(\frac{1}{2})^{n-2}}{\Gamma(\frac{3}{2}n)} \frac{(\frac{1}{w})^{n-2}}{w^{\frac{3}{2}}} \right] dw
\]

\[
= \frac{\Gamma(\frac{1}{2})^n}{\Gamma(\frac{3}{2}n)} \int_a^b f(w) \frac{1}{w^{\frac{3}{2}n-1}} dw
\]

and the lemma is proved.
Returning to the computation of $H_n(t)$ we have from its definition that

$$H_n(t) = \iint \cdots \int dx_1 \cdots dy_n$$

$$\text{if} \quad x_1^2 + \ldots + x_n^2 \leq t(x_1^2 + \ldots + y_n^2)$$

Using the formula for volume of a sphere to integrate $y_1, \ldots, y_n$ in the numerator and $x_1, \ldots, y_n$ in the denominator, we get

$$\frac{\Gamma'(n+1)}{\Gamma\left(\frac{1}{2}n\right)\Gamma\left(\frac{1}{2}(n+1)\right)k^n} \iint \cdots \int \left[ (k-x_1^2-\ldots-x_n^2)^{\frac{3}{2}n} - \frac{1-t}{t}(x_1^2+\ldots+x_n^2)^{\frac{3}{2}n} \right] dx_1 \cdots dx_n$$

Applying the lemma,

$$H_n(t) = \frac{\Gamma'(n+1)}{\Gamma\left(\frac{1}{2}n\right)\Gamma\left(\frac{1}{2}(n+1)\right)k^n} \int_0^t \left[ (k-w)^{\frac{3}{2}n} - \frac{1-t}{t} w^{\frac{3}{2}n} \right] w^{\frac{3}{2}n-1} dw$$

Substituting $w=zk$ gives

$$(2.3.3) \quad H_n(t) = \frac{\Gamma'(n+1)}{\Gamma\left(\frac{1}{2}n\right)\Gamma\left(\frac{1}{2}(n+1)\right)} \left[ \int_0^t (1-z)^{\frac{3}{2}n} z^{\frac{3}{2}n-1} dz - \frac{1}{n} (1-t)^{\frac{3}{2}n} t^{\frac{3}{2}n} \right]$$

In the same fashion we write
In the numerator we make the change of variables

\[ x_i = (u_i + v_i)^{\frac{1}{2}} \]

\[ y_i = (u_i - v_i)^{\frac{1}{2}} \]

Its determinant is 1, so

\[
G_n(t) = \iint \ldots \int \frac{dx_1 \ldots dy_n}{x_1^2 + \ldots + y_n^2 \leq k} \\
\frac{t(x_1^2 + \ldots + y_n^2) \leq |x_1y_1 + \ldots + x_ny_n|}{\iint \ldots \int dx_1 \ldots dy_n} \\
\frac{x_1^2 + \ldots + y_n^2 \leq k}{\iint \ldots \int dx_1 \ldots dy_n}
\]

\[
G_n(t) = \frac{2^n \Gamma(n+1)}{\Gamma(\frac{1}{2}n) \Gamma(\frac{3}{2}n)} \iint \ldots \int \frac{du_1 \ldots dv_n}{u_1^2 + \ldots + v_n^2 \leq k} \\
\frac{2t(u_1^2 + \ldots + v_n^2) \leq |(u_1^2 + \ldots + u_n^2) - (v_1^2 + \ldots + v_n^2)|}{\Gamma(\frac{1}{2}n) \Gamma(n+1) \Gamma(\frac{3}{2}n)} \cdot k^n
\]

\[
= \frac{2^n \Gamma(n+1)}{\Gamma(\frac{1}{2}n) \Gamma(\frac{3}{2}n)} \iint \ldots \int \frac{du_1 \ldots dv_n}{u_1^2 + \ldots + v_n^2 \leq k} \\
\frac{2t(u_1^2 + \ldots + v_n^2) \leq (u_1^2 + \ldots + u_n^2) - (v_1^2 + \ldots + v_n^2)}{\Gamma(\frac{1}{2}n) \Gamma(n+1) \Gamma(\frac{3}{2}n)} \cdot k^n
\]

\[
= \frac{2^n \Gamma(n+1)}{\Gamma(\frac{1}{2}n) \Gamma(\frac{3}{2}n)} \iint \ldots \int \frac{du_1 \ldots dv_n}{\frac{1+2t}{1-2t}(v_1^2 + \ldots + v_n^2) \leq u_1^2 + \ldots + u_n^2 \leq k-(v_1^2 + \ldots + v_n^2)} \\
\frac{k-(v_1^2 + \ldots + v_n^2)}{\frac{1+2t}{1-2t}(v_1^2 + \ldots + v_n^2) \leq k-(v_1^2 + \ldots + v_n^2)} \cdot k^n \cdot \Gamma(\frac{1}{2}n) \Gamma(n+1) \Gamma(\frac{3}{2}n)
\]

\[
= \frac{2^n \Gamma(n+1)}{\Gamma(\frac{1}{2}n) \Gamma(\frac{3}{2}n)} \iint \ldots \int \frac{du_1 \ldots dv_n}{\frac{1+2t}{1-2t}(v_1^2 + \ldots + v_n^2) \leq k-(v_1^2 + \ldots + v_n^2)} \\
\frac{k-(v_1^2 + \ldots + v_n^2)}{\frac{1+2t}{1-2t}(v_1^2 + \ldots + v_n^2) \leq k-(v_1^2 + \ldots + v_n^2)} \cdot k^n \cdot \Gamma(\frac{1}{2}n) \Gamma(n+1) \Gamma(\frac{3}{2}n)
\]

\[
= \frac{2^n \Gamma(n+1)}{\Gamma(\frac{1}{2}n) \Gamma(\frac{3}{2}n)} \iint \ldots \int \frac{du_1 \ldots dv_n}{\frac{1+2t}{1-2t}(v_1^2 + \ldots + v_n^2) \leq k-(v_1^2 + \ldots + v_n^2)} \\
\frac{k-(v_1^2 + \ldots + v_n^2)}{\frac{1+2t}{1-2t}(v_1^2 + \ldots + v_n^2) \leq k-(v_1^2 + \ldots + v_n^2)} \cdot k^n \cdot \Gamma(\frac{1}{2}n) \Gamma(n+1) \Gamma(\frac{3}{2}n)
\]
Applying the lemma, 

\[ G_n(t) = \frac{2\Gamma(n+1)}{k^n\Gamma(\frac{1}{2n+1})\Gamma(\frac{1}{2n})} \int_0^k \frac{1}{2n+1} \left[ (k-w)^{\frac{1}{2n}} \left( \frac{1+2t}{1-2t} \right)^{\frac{1}{2n}} \right] w^{n-1} dw \]

Substituting \( w = k(\frac{1}{2}-z) \),

\[ G_n(t) = \frac{2\Gamma(n+1)}{\Gamma(\frac{2n+1}{2})\Gamma(\frac{1}{2n})} \int_0^t \left[ (z+2)^{\frac{1}{2n}} (1-z)^{\frac{1}{2n}} \right] dz \frac{1}{n}(\frac{1}{2} + t)^{\frac{3}{2n}}(\frac{1}{2} - t)^{\frac{3}{2n}} \]

2.4 Estimates for \( H_n(t), G_n(t) \)

We estimate the integrals in the formulas for \( H_n(t) \) and \( G_n(t) \) as follows:

If \( t \leq \frac{1}{2} \),

\[ \int_0^t (1-z)^{\frac{3}{2n}} z^{n-1} dz = \int_0^t (z^2)^{\frac{3}{2n}} (1-z) dz \]

\[ \leq \frac{1-t}{1-2t} \int_0^t (z^2)^{\frac{3}{2n}} (1-2z) dz \]

\[ = \frac{1-t}{1-2t} \frac{2}{n} (z^2)^{\frac{3}{2n}} \bigg|_0^t \]

\[ = \frac{2}{n} (t-t^2)^{\frac{3}{2n}} \frac{1-t}{1-2t} \]

So

\[ H_n(t) \leq \frac{\Gamma(n+1)}{\Gamma(\frac{2n+1}{2})\Gamma(\frac{1}{2n})} \frac{1}{n} (t-t^2)^{\frac{3}{2n}} \left[ 2\left( \frac{1-t}{1-2t} \right) - 1 \right] \]

(2.4.1) \[ H_n(t) \leq \frac{\Gamma(n+1)}{\Gamma(\frac{2n+1}{2})^2} \frac{(t-t^2)^{\frac{3}{2n}}}{2(1-2t)} \]
Also,
\[
\int_t \left( \frac{1}{\tau^2} \right) |z|^{n-1} |dz| = \int_t \left( \frac{1}{\tau^2} \right) |z|^{n-1} \left( \frac{1}{\tau^2} \right) |dz|
\]
\[
\leq - \frac{1}{2t} \int_t \left( \frac{1}{\tau^2} \right) |z|^{n} (-2z) |dz|
\]
\[
= - \frac{1}{2t} \frac{2\left( \frac{1}{\tau^2} \right) |z|^{n}}{n}
\]
\[
= \frac{1}{2t} \frac{2\left( \frac{1}{\tau^2} \right) |z|^{n}}{n}
\]

So,
\[
G_n(t) \leq \frac{2\left( \Gamma(n+1) \right)}{\Gamma\left( \frac{1}{2} n+1 \right) \Gamma\left( \frac{1}{2} n \right) \Gamma\left( \frac{1}{2} n+1 \right)} \frac{1}{n} \left( \frac{1}{\tau^2} \right)^{n} \left[ \frac{1}{t} \right]^{n-1}
\]

(2.4.2) \[
G_n(t) \leq \frac{\Gamma(n+1)}{\Gamma\left( \frac{1}{2} n+1 \right)^2} \frac{\left( \frac{1}{\tau^2} \right)^{n}}{2t}
\]

2.5 The conditional probability of reduction

Let \( K \) be the set of non-reduced matrices in \( \mathbb{R}^{mn} \). Let \( B_{i,j} \) be the set of matrices for which \( \left| \frac{a_i}{a_i + a_j} \right| > \frac{1}{2(mM+1)} \) and \( C_{i,j} \) (\( i < j \)) be the set of matrices for which \( \frac{a_i^2}{a_i^2 + a_j^2} < \frac{mM}{2(mM+1)} \). Then by the corollary of section 2.2,

\[
K \cap S_K \subseteq ( \bigcup_{i,j} B_{i,j} \cup \bigcup_{i < j} C_{i,j} ) \cap S_K
\]
So
\[ P(K \cap S_k | S_k) \leq \sum_{i,j} P(B_{i,j} \cap S_k | S_k) + \sum_{i<j} P(C_{i,j} \cap S_k | S_k) \]
\[ = \sum_{i,j} H_n\left(\frac{mM}{2(mM+1)^2}\right) + \sum_{i<j} G_n\left(\frac{1}{2(mM+1)}\right) \]
\[ = m(m-1)H_n\left(\frac{mM}{2}\right) + \frac{m(m-1)}{2} G_n\left(\frac{1}{2(mM+1)}\right) \]

So
\[ (2.5.1) \quad P_{m,n,k} \leq m(m-1) \frac{\Gamma(n+1)}{\Gamma(\frac{3}{2}n+1)^2} \left(\frac{m^2M + 2mM}{4(mM+1)^2}\right)^{\frac{3}{2}n}\frac{1}{(mM+1)^2} \]

Applying Stirling's formula to estimate the Gamma function, we get
\[ P_{m,n,k} \leq m(m-1) \frac{\sqrt{2\pi n^\frac{1}{2}}}{\Gamma\left(\frac{3}{2}n+1\right)} \left(\frac{m^2M + 2mM}{4(mM+1)^2}\right)^{\frac{3}{2}n}\frac{1}{(mM+1)^2} \]

(2.5.2) \[ P_{m,n,k} \leq m(m-1) \left(\frac{mM+1}{2}\right)^{\frac{1}{2}} \left(1 + \frac{1}{12n-1}\right) \left(\frac{m^2M + 2mM}{(mM+1)^2}\right)^{\frac{1}{2}n} \]

Since \[ \frac{m^2M + 2mM}{(mM+1)^2} < 1, \] we have

**Theorem** For fixed \( m, \) as \( n \rightarrow \infty, \) \( P_{m,n,k} \rightarrow 0. \)

**Remark** Since \( A \) is reduced if and only if any scalar multiple of \( A \) is reduced, \( P_{m,n,k} \) does not actually depend on \( k. \) Therefore from now on we shall simply write \( P_{m,n}. \)
2.6 The other probability models

We now consider the other situations discussed in section 2.1. Suppose \( X \), the random variable representing a \( m \times n \) matrix, has density function \( f \) on \( \mathbb{R}^{mn} \). If \( f \) is of a certain type, then

\[
P(K) = \int_K f(x) = P_{m,n}.
\]

Suppose \( f \) depends only on \( |x| \), the Euclidean length of \( X \).

Then

\[
P(K) = \lim_{k \to \infty} P(X \in K \cap S_K)
\]

\[
= \lim_{k \to \infty} P(\frac{X}{|X|} \in K \text{ and } |X| \leq k)
\]

\[
= \lim_{k \to \infty} P(\frac{X}{|X|} \in K)P(|X| \leq k)
\]

\[
= P(X \in K \mid |X| = 1) \lim_{k \to \infty} P(|X| \leq k)
\]

\[
= P_{m,n}
\]

Now we consider the density of the integer points of \( K \).

\[
D(K) = \lim_{k \to \infty} \frac{\text{(# of integer points in } K \cap S_K\text{)}}{\text{(# of integer points in } S_K\text{)}}
\]

\[
= \lim_{k \to \infty} \frac{\text{(# of integer points in } k(K \cap S_1)\text{)}}{\text{(# of integer points in } kS_1\text{)}}
\]

\[
= \lim_{k \to \infty} \frac{k^{mn} m(K \cap S_1)}{k^{mn} m(S_1)}
\]

\[
= \frac{m(K \cap S_1)}{m(S_1)}
\]

\[
= P_{m,n}
\]
3.1 Reduction by submatrices

The purpose of this chapter is to investigate Minkowski reduction, with the aim of making it more suitable for computation. We recall that the main difficulty in Minkowski reduction is the large number of inequalities to be tested. The theorem of chapter two suggests that some type of blockwise reduction might be advantageous; since long thin matrices are easy to reduce, perhaps we can develop some connection between reduction of a matrix and its submatrices. For example, given a 24 x 24 matrix A, we break it up into six 4 x 24 matrices $A_1, \ldots, A_6$. By reducing $A_1, \ldots, A_6$, we hope to be able to find the reduced form of A.

Unfortunately, even if all submatrices of a matrix are reduced, there is still a possibility that a linear combination of all its rows may give a further reduction. In fact, the inequalities to be checked involving all the rows will be the most numerous and difficult to use.

To circumvent this difficulty, we must be able to make final the choice of the submatrices $A_1, \ldots, A_k$ from the row module; that is, no more operations involving rows of different submatrices should be necessary, once $A_1, \ldots, A_k$ are chosen. Also, further reduction of the individual submatrices should be possible without altering this choice.

Minkowski's definition of reduction will not permit this type
of scheme; we need a new definition. To do so, we need a way of choosing submatrices of a matrix which generalizes Minkowski's method of choosing rows, which is based on the gauge function. We will obtain a gauge function on matrices by first looking at the Grassmann algebra of a matrix.

3.2 The extension of the gauge function

Our gauge functions were originally defined on a real vector space $V$, and we shall extend them to the Grassmann algebra $G$ of that vector space. We recall that the Grassmann algebra is an anticommutative associative algebra with vector space basis
\[ \left\{ e_1 \wedge e_2 \wedge \ldots \wedge e_m : r_i = 0 \text{ or } 1 \right\}, \]
where $e_1, \ldots, e_m$ is a basis of $V$. We can write $G = G_0 \oplus G_1 \oplus \ldots \oplus G_m$, where $G_i$ is the set of homogeneous elements of degree $i$. In particular $G_1 \cong V$.

**Definition** Let $f$ be a gauge function on $V$. Then $\overline{f} : G \to \mathbb{R}$ is called an extension of $f$ if:

1. $\overline{f} \mid_{G_1} = f$
2. $\overline{f}$ is a gauge function on $G$, considered as a vector space.
3. $\wedge : G \times G \to G$ is continuous in the topology defined by $\overline{f}$.

In fact, condition 3 is not necessary. For

**Proposition** Any bilinear mapping of finite dimensional normed linear spaces is continuous.

**Proof** Let $B : C \times D \to E$ be a bilinear mapping of finite dimensional
normed linear spaces. Let \( c_1, c_2, \ldots, c_t; d_1, \ldots, d_u; e_1, \ldots, e_v \) be bases for \( C, D, \) and \( E \) respectively. Because all norms on a finite dimensional space give the same topology, we may assume without loss of generality that the norms are:

\[
\begin{align*}
N_C(\sum x_i c_i) &= \sum |x_i| \\
N_D(\sum x_i d_i) &= \sum |x_i| \\
N_E(\sum x_i e_i) &= \sum |x_i|
\end{align*}
\]

Then

\[
N_E(\sum x_i c_i, \sum y_i d_j) = N_E(\sum x_i y_j B(c_i, d_j)) \leq \sum |x_i| |y_j| N_E(B(c_i, d_j)) \leq N_C(\sum x_i c_i) N_D(\sum y_j d_j) \max N_E(B(c_i, d_j))
\]

Thus there is a constant \( k \) such that

\[
N_E(B(c, d)) \leq k N_C(c) N_D(d)
\]

for all \( c \in C, d \in D \).

Now let \((x_n, y_n) \rightarrow (x, y)\) in \( C \times D \). Then \( x_n \rightarrow x \) and \( y_n \rightarrow y \).

\[
N_E(B(x_n, y_n) - B(x, y)) = N_E(B(x_n - x, y_n - y) + B(x_n - x, y) + B(x, y_n - y)) \leq N_E(B(x_n - x, y_n - y)) + N_E(B(x_n - x, y)) + N_E(B(x, y_n - y)) \leq k N_C(x_n - x) N_D(y_n - y) + k N_C(x - x) N_D(y) + k N_C(x) N_D(y_n - y)
\]

which tends to \( 0 \) as \( n \rightarrow \infty \). So \( B(x_n, y_n) \rightarrow B(x, y) \).

3.3 The gauge function on submatrices

We construct a map \( q \) from the set of matrices into the Grassmann algebra as follows: Let \( A \) be a matrix, made up of the row vectors \( a_1, \ldots, a_m \). Then \( q(A) = a_1 \wedge a_2 \wedge \ldots \wedge a_m \). If \( f \) is a
gauge function on a vector space containing $a_1, \ldots, a_m$, and if $\overline{f}$ is its extension to the Grassmann algebra, then $\overline{f} \circ q$ will be the gauge by which we measure the size of submatrices of a matrix. To complete the connection we note that

**Proposition** If $L$ is any linear transformation of $V$, and $a_1, \ldots, a_m$ is a basis of $V$ then

$$L(a_1) \wedge \cdots \wedge L(a_m) = (\det L) a_1 \wedge \cdots \wedge a_m \quad [17]$$

**Corollary** If $B$ is any $m \times m$ matrix, then $f(BA) = |\det B| f(A)$.

**Corollary** If $U$ is any $m \times m$ unimodular matrix, then $f(UA) = f(A)$.

**Corollary** $A$ is row-dependent if and only if $f(A) = 0$.

These are the important properties we need for our applications to matrices. Corollary two allows us to consider $f$ as a function on submodules of the row module, as follows: Let $N$ be a submodule. Let $b_1, \ldots, b_k$ be a basis of $N$. Define $f(N) = f(b_1 \wedge \cdots \wedge b_k)$. Since any other basis of $N$ is obtained from $b_1, \ldots, b_k$ by a unimodular transformation, $f$ is well-defined.

**Example** The function $f(A) = \sqrt{\det(AA^T)}$ is a gauge on matrices.

It corresponds to the function $f$ on the Grassmann algebra defined by $f(\sum x_i e_i) = \sqrt{\sum x_i^2}$, where $e_1, \ldots, e_k$ is the standard basis of $G$. In this case we have $f(AB) \leq f(A)f(B)$; that is, we can take the constant $k$ of the proposition of section 3.2 to be 1.

### 3.4 Reduction by submodules

In Minkowski reduction, we selected a basis for the row module $M$ of a matrix by choosing basis vectors as short as
possible, according to the gauge function $f$. We will now define a reduction procedure which selects submodules of $M$ as small as possible according to the extended gauge function, which we shall also call $f$.

**Definition** Let $l_1, l_2, \ldots, l_k$ be positive integers such that $l_1 + l_2 + \ldots + l_k \leq m$. Let $M$ be an $m$-dimensional free $\mathbb{Z}$-module. A system of submodules $N_1, \ldots, N_k$ of $M$ is called $(l_1, \ldots, l_k)$ primitive if

1. $\dim N_i = l_i$, $i = 1, \ldots, k$
2. $N_1, \ldots, N_k$ are linearly independent.
3. If $r_1 a_1 + \ldots + r_k a_k \in A_1 + \ldots + A_k$ ($r_i$ real, $a_i \in A_i$), then $r_1, \ldots, r_k \in \mathbb{Z}$.

In particular, a $(1_1)$ primitive system is called a primitive submodule. A primitive $(1, 1, \ldots, 1)$ system is called a primitive system of vectors. If $l_1 + \ldots + l_k = m$, a primitive $(l_1, \ldots, l_k)$ system is called a basic $(l_1, \ldots, l_k)$ system.

**Theorem** If $l_1 + \ldots + l_p = m = \dim M$, and $1 \leq k \leq p$, then any primitive $(l_1, \ldots, l_k)$ system of submodules of $M$ can be extended to a basic $(l_1, \ldots, l_p)$ system of $M$.

**Proof** First we show that any $1_1$-dimensional primitive submodule $N_1$ of $M$ can be extended to a basic $(1_1, \ldots, l_p)$ system. Let $b_1, \ldots, b_{1_1}$ be a basis of the submodule $N_1$. We show that $b_1, \ldots, b_{1_1}$ is a primitive system of vectors.

Suppose $r_1 b_1 + \ldots + r_{1_1} b_{1_1} \in N_1$ ($r_i$ real). Since
\( N_1 = Zb_1^1 + \ldots + Zb_1 \), \( r_1 b_1^1 + \ldots + r_1 b_1 \) can be expressed as \( z_1 b_1^1 + \ldots + z_1 b_1 \), with integers \( z_1, \ldots, z_1 \). Since \( b_1, \ldots, b_1 \) are linearly independent, \( r_1 = z_1 \), so \( r_1 \in \mathbb{Z} \) for \( i = 1, \ldots, l_1 \).

Since \( b_1, \ldots, b_1 \) is a primitive system of vectors, it can be extended to a basis of \( N \). \( \{14\} \) If we group these basis vectors in groups of \( l_2, \ldots, l_p \) in any fashion we obtain the desired basic \( l_1, \ldots, l_p \) system for \( M \).

Now suppose we start with a primitive \( (l_1, \ldots, l_k) \) system \( N_1, \ldots, N_k \). Then \( N_1 + \ldots + N_k \) is a primitive submodule, for if \( r(a_1^1 + \ldots + a_k^1) \in N_1 + \ldots + N_k \), where \( a_i \in N_i \), then \( ra_1 + \ldots + ra_k \in N_1 + \ldots + N_k \), and so \( r \in \mathbb{Z} \). We can extend \( N_1 + \ldots + N_k \) to a basic \( (l_1 + \ldots + l_k^1, l_k^1 + 1, \ldots, l_p^1) \) system \( N_1 + \ldots + N_k, N_{k+1}, \ldots, N_p \), and \( N_1, \ldots, N_p \) is a basic \( (l_1, \ldots, l_p) \) system. This completes the proof.

We assume now that we have some gauge function \( f \) defined on submodules of \( M \). If \( f(N_1) \leq f(N_2) \) we say \( N_1 \) is smaller than \( N_2 \).

**Lemma** Let \( M \) be a \( \mathbb{Z} \)-module with basis \( a_1, \ldots, a_m \), \( q \) the mapping of submodules into the Grassmann algebra \( G \). Then \( \{q(N) : N \text{ is a submodule of } M\} \) is a discrete set in \( G \).

**Proof** Suppose \( q(N_n) \) converges to \( x \in G \). Let \( \{b_{n1}, \ldots, b_{nk_n}\} \) be a basis of \( N_n \). Let \( b_{ni} = \sum_{j=1}^{m} c_{ni} a_j \). Then

\[
q(N_n) = \left( \sum_{j=1}^{m} c_{ni} a_j \right) \wedge \ldots \wedge \left( \sum_{j=1}^{m} c_{nk_n} a_j \right) = \sum_{s=1}^{t} z_{ns} g_s,
\]

where \( z_{ns} \) are
integers and $g_s$ are the elements of the form $a_1^{d_1} \Lambda \cdots \Lambda a_m^{d_m}$, $d_i = 0$ or 1. Since $\{g_s\}$ is linearly independent, we must have $\{z_{ns}\}$ converges for each $s$. Therefore $z_{ns} = z_s$ for some integer $z_s$ for every $s$ for sufficiently large $n$. Therefore $q(N_n) = \sum z_s g_s$ for sufficiently large $n$.

**Theorem** In every nonempty set of submodules there is a smallest one.

**Proof** By the lemma, $\{f(N) : N \text{ is a submodule}\}$ is a discrete set of nonnegative real numbers, and therefore every subset has a least element $r$. Given any set of submodules $S$, any $N \in S$ with $f(N) = r$ has the required property.

**Definition** A basic $(l_1, \ldots, l_p)$ system $N_1, \ldots, N_p$ of submodules of $M$ is called reduced with respect to the extended gauge function $f$ if the following conditions are fulfilled:

1. $N_1$ is the smallest $l_1$-dimensional submodule of $M$.
2. For $k = 2, \ldots, p$, $N_k$ is the smallest $l_k$-dimensional submodule of $M$ which, together with $N_1, \ldots, N_{k-1}$, forms a primitive $(l_1, \ldots, l_k)$ system of submodules.

In particular, a basic $(l, l, \ldots, l)$ system is a Minkowski reduced basis for $M$.

**Theorem** Every $1 + \cdots + 1_p$ dimensional $\mathbb{Z}$-module $M$ has a reduced basic $(1, 1, \ldots, 1_p)$ system.

**Proof** We construct the system inductively. The second theorem says we can choose $N_1$ to satisfy condition 1. The first theorem says that the set of submodules which together with $N_1, \ldots, N_{k-1}$
form a primitive system is nonempty; the second allows us to select a smallest one.

\((l_1, \ldots, l_p)\) reductions may be used to select a canonical basis for \(M\), by further reducing each submodule \(N_i\) once it has been chosen. By combining \((l_1, \ldots, l_p)\) reductions we have many different reduction schemes, each of which yields a basis of \(M\).

The consideration of the uniqueness of the basic system is similar to that of Minkowski reduction. Because of the discreteness, every choice made in the proof of theorem 3 is from a finite number of submodules. Therefore a module can have only finitely many reduced basic systems.

3.5 \((l_1, \ldots, l_p)\) reduction of matrices

A matrix \(A\) with rows \(a_1, \ldots, a_m\) may be thought of as consisting of \(p\) submatrices \(A_1, \ldots, A_p\), such that \(A_1\) consists of the rows \(a_1, \ldots, a_1\), \(A_2\) consists of the rows \(a_1, a_1, \ldots, a_1\), and so on. \(A\) will be called \((l_1, \ldots, l_p)\) reduced if the submodules \(N_1, \ldots, N_p\) associated with \(A_1, \ldots, A_p\) form a basic \((l_1, \ldots, l_p)\) system in the row module \(M\). Again, by combining reductions we can construct a reduction scheme which selects a particular matrix from each equivalence class. As an example, we shall investigate the computation of the reduced form of a matrix according to a particular scheme, namely successive \((1, m-1)\) reductions.

3.6 The \((1, m-1)\) reduction scheme

The particular reduction scheme will be defined as follows: Denote by \(A^{(k)}\) the matrix formed by the last \(k\) rows of \(A\). Then \(A\)
is reduced if and only if for each $k=2,...,m$ we have that $A^{(k)}$ is a $(1,k-1)$ reduced matrix. The procedure used for finding the reduced matrix according to this scheme would be simply to first $(1,m-1)$ reduce $A$, then $(1,m-2)$ reduce $A^{(m-1)}$, then $(1,m-3)$ reduce $A^{(m-2)}$, and so on.

The method for the actual $(1,k-1)$ reduction of each matrix is given by the following:

**Proposition** Let $A$ be a $m \times n$ matrix with rows $a_1,...,a_m$. Then $A$ is $(1,m-1)$ reduced if and only if $f(s_1 a_1 + \ldots + s_m a_m) \leq f(s_1 a_1 + \ldots + s_m a_m)$ for any relatively prime integers $s_1,...,s_m$, and

$$
\begin{pmatrix}
a_2 \\
\vdots \\
a_m
\end{pmatrix} \leq
\begin{pmatrix}
a_2 + t_2 a_1 \\
\vdots \\
a_m + t_m a_1
\end{pmatrix}
$$

for any system of integers $t_2,...,t_m$.

Here \( \begin{pmatrix} a_2 \\ \vdots \\ a_m \end{pmatrix} \) and \( \begin{pmatrix} a_2 + t_2 a_1 \\ \vdots \\ a_m + t_m a_1 \end{pmatrix} \) denote the matrices whose rows are $a_2,...,a_m$ and $a_2 + t_2 a_1,...,a_m + t_m a_1$ respectively.

**Proof** First we prove the sufficiency. If $f(s_1 a_1 + \ldots + s_m a_m) < f(a_1)$, then the first condition of the definition of reduction is not satisfied; if $f\left(\begin{pmatrix} a_2 \\ \vdots \\ a_m \end{pmatrix}\right) > f\left(\begin{pmatrix} a_2 + t_2 a_1 \\ \vdots \\ a_m + t_m a_1 \end{pmatrix}\right)$ then the second condition
is not satisfied.

Next the necessity. Suppose $A$ is not $(l,m-l)$ reduced. If $a_1$ is not the shortest vector in the row module of $A$, then $f(a_1) > f(s_1a_1 + \ldots + s_m a_m)$ for some integers $s_1, \ldots, s_m$. If $d = \gcd(s_1, \ldots, s_m)$, then

$$f\left(\frac{s_1}{d} a_1 + \ldots + \frac{s_m}{d} a_m\right) = \frac{1}{d} f(s_1a_1 + \ldots + s_m a_m) < f(a_1)$$

which contradicts the first hypothesis.

If \(
\begin{pmatrix} a_2 \\ \vdots \\ a_m \end{pmatrix}
\)

is not the smallest $(m-1)$ dimensional submodule of the row module which along with $a_1$ forms a primitive system, then

$$f\left(\begin{pmatrix} b_2 \\ \vdots \\ b_m \end{pmatrix}\right) < f\left(\begin{pmatrix} a_2 \\ \vdots \\ a_m \end{pmatrix}\right)$$

where $a_1, b_2, \ldots, b_m$ is a primitive system, and therefore a basis of the row module. We then have that

$$\begin{pmatrix} a_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = U \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

for some unimodular $U$ in the form
Let \( V = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ \end{pmatrix} \)

Then \( V \begin{pmatrix} a_1 \\ b_2 \\ \vdots \\ b_m \\ \end{pmatrix} = VU \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \\ \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ t_2 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_m & \cdots & \cdots & 0 & 1 \\ \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \\ \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 + t_2 a_1 \\ \vdots \\ a_m + t_m a_1 \\ \end{pmatrix}

So \( (U^*)^{-1} \begin{pmatrix} b_2 \\ \vdots \\ b_m \\ \end{pmatrix} = \begin{pmatrix} a_2 + t_2 a_1 \\ \vdots \\ a_m + t_m a_1 \\ \end{pmatrix} \)

Thus \( f \begin{pmatrix} a_2 + t_2 a_1 \\ \vdots \\ a_m + t_m a_1 \\ \end{pmatrix} = f (U^*)^{-1} \begin{pmatrix} b_2 \\ \vdots \\ b_m \\ \end{pmatrix} = f \begin{pmatrix} b_2 \\ \vdots \\ b_m \\ \end{pmatrix} \neq f \begin{pmatrix} a_2 \\ \vdots \\ a_m \\ \end{pmatrix} \)

which contradicts the second hypothesis.

To obtain the reduced matrix, we must therefore be able to minimize the functions
\[ F(s_1, \ldots, s_m) = f(s_1 a_1 + \ldots + s_m a_m) \]

and

\[ D(t_2, \ldots, t_m) = f\left( \begin{array}{c} a_2 + t_2 a_1 \\ \vdots \\ a_m + t_m a_1 \end{array} \right) \]

over \( \mathbb{Z} \). We can simplify

\[ D(t_2, \ldots, t_m) = f((a_2 + t_2 a_1) \Lambda \ldots \Lambda (a_m + t_m a_1)) \]

\[ = f(a_2 \Lambda \ldots \Lambda a_m + t^2 A(2) + t^3 A(3) + \ldots + t^m A(m)), \]

where \( A(i) = a_2 \Lambda \ldots \Lambda a_{i-1} \Lambda a_i \Lambda a_{i+1} \Lambda \ldots \Lambda a_m \).

In the case where the gauge function is the familiar Euclidean norm, \( F \) is just a positive definite quadratic form in \( s_1, \ldots, s_m \) and \( D \) is a positive definite quadratic polynomial.

The improvement therefore over the standard Minkowski reduction is that instead of finding the successive minima of a module or form, we need only find the absolute minimum. In addition, we can expect on the basis of the theorem of chapter 2 that the task of \((1,k-1)\) reducing each submatrix will become easier as we proceed, since the number of columns will remain fixed as the number of rows decreases.
APPENDIX

COMPUTER RESULTS

To verify the theorem of chapter two, a computer program was written by Merle Owdom of Ohio State’s Instruction and Research Computer Center. The program generated matrices using the system’s random number generator, then tested whether or not they were Minkowski reduced. The matrix entries were integers bounded in magnitude by 100, and were assumed to be distributed uniformly. For each dimension pair, 1000 matrices were tested, using the reduction inequalities. The results follow:

<table>
<thead>
<tr>
<th>m</th>
<th>n</th>
<th>% reduced</th>
<th>m</th>
<th>n</th>
<th>% reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>47.4</td>
<td>3</td>
<td>4</td>
<td>13.7</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>73.4</td>
<td>3</td>
<td>8</td>
<td>44.9</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>84.3</td>
<td>3</td>
<td>12</td>
<td>65.2</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>93.1</td>
<td>3</td>
<td>16</td>
<td>77.6</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>95.1</td>
<td>3</td>
<td>20</td>
<td>87.9</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>97.1</td>
<td>3</td>
<td>24</td>
<td>94.0</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>98.5</td>
<td>3</td>
<td>28</td>
<td>96.0</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>99.2</td>
<td>3</td>
<td>32</td>
<td>98.2</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>99.1</td>
<td>3</td>
<td>36</td>
<td>98.1</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>99.8</td>
<td>3</td>
<td>40</td>
<td>99.3</td>
</tr>
<tr>
<td>2</td>
<td>44</td>
<td>99.9</td>
<td>3</td>
<td>44</td>
<td>99.6</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>99.9</td>
<td>3</td>
<td>48</td>
<td>99.7</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>100.0</td>
<td>3</td>
<td>52</td>
<td>99.9</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>100.0</td>
<td>3</td>
<td>56</td>
<td>100.0</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>100.0</td>
<td>3</td>
<td>60</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 1. Percentage of Reduced Matrices

34
BIBLIOGRAPHY


