Application of the Smith Normal Form to the Structure of Lattice Rules

J. N. Lyness* P. Keast†

Abstract

Two independent approaches to the theory of the lattice rule have been exploited at length in the literature. One is based on the generator matrix A of the lattice Λ whose elements provide the abscissas of Q. The other, based on the t-cycle form $Q(\Lambda)f$ of Sloan and Lyness, leads to a canonical form for Q. In this paper, a close connection between these approaches is demonstrated. This connection stems from the close relation between the Kronecker decomposition theorem for Abelian groups and the Smith normal form of an integer matrix. It is shown that the invariants of the canonical form of $Q(\Lambda)f$ coincide with the elements of the Smith normal form of $B = A^{T-1}$, the reciprocal lattice generator matrix. This fact may be used to provide a straightforward solution to the previously intransigent problem of identifying and removing a repetition in the general t-cycle form.

Keywords: Smith normal form, lattice rule, multidimensional quadrature, good lattice points

AMS(MOS) subject classification: 65D32

*Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439. The work of this author was supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

†Dept. of Mathematics, Statistics, and Computing Science, Dalhousie University, Halifax, Nova Scotia B3H 3J5, Canada. The work of this author was partially supported by grant number OGP0002699 from the Natural Sciences and Engineering Research Council of Canada.

1 Background and Introduction

A lattice rule is a multidimensional quadrature rule for integrating over an s-dimensional hypercube. In this section we provide a brief description of some of the theory followed by an outline of the contents of the rest of the paper. Without loss of generality we shall take the hypercube of integration to be $[0,1)^s$.

An *s-dimensional lattice* Λ is an infinite array of points. These satisfy; (a) $\mathbf{p}, \mathbf{q} \in \Lambda$ implies $\mathbf{p} - \mathbf{q} \in \Lambda$; and (b) there exists no limit point; that is, there exists a positive $\epsilon(\Lambda)$ such that $|\mathbf{p} - \mathbf{q}| \geq \epsilon(\Lambda)$ unless $\mathbf{p} = \mathbf{q}$. A *generator matrix* A of Λ is a $t \times s$ matrix, whose rows $\mathbf{a}_r$ are elements of Λ having the property that the lattice comprises all points $\mathbf{p}$ of the form

$$\mathbf{p} = \sum_{r=1}^{t} \lambda_r \mathbf{a}_r = \lambda A, \tag{1.1}$$

where λ_i are integer and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_t)$.

Of special note is the *unit lattice* Λ_0, which is generated by the $s \times s$ unit matrix I. An *integration lattice* Λ is a lattice that contains Λ_0 as a sublattice. A lattice rule $Q(\Lambda)$ has an abscissa set comprising all the points of the integration lattice Λ which lie in $[0,1)^s$. Corresponding to an s-dimensional lattice Λ is its *reciprocal lattice* $\Lambda^{\perp}$. When the generator matrix of A is nonsingular, this may be defined as the lattice having a generator matrix $B = (A^T)^{-1}$. The generator matrix of an integration matrix is nonsingular. It is relatively straightforward to show that the condition for Λ to be an integration lattice is that B, the generator matrix of $\Lambda^{\perp}$, be an integer matrix, i.e., every element of B be an integer. When Λ is an integration lattice,

$$N = |\det A|^{-1} = |\det B| \tag{1.2}$$

is an integer and coincides with the number of lattice points in the hypercube $[0,1)^s$. These N points form the abscissa set of the *lattice rule* $Q(\Lambda)$ based on the lattice Λ. This rule applies an equal weight $1/N$ to each abscissa. Moreover, every nonzero element of an integration lattice generator matrix A is a rational whose denominator is a factor of N.

A generator matrix of a lattice Λ is not unique. Different generator matrices of the same lattice may be related using unimodular matrices. A *unimodular matrix* V is a square integer matrix of determinant ± 1. The inverse of a unimodular matrix is also unimodular,

and elementary integer row (column) operations on a general matrix may be accomplished by pre- (post-) multiplication by a unimodular matrix. In particular, when A is a generator matrix of Λ, so is VA, where V is any unimodular matrix, and all other generator matrices of Λ are of this form. (But the lattice generated by AV is generally unrelated to Λ.)

The approach to the theory of lattice rules based on lattice generator matrices is described in more detail in Lyness 1989. There the connection between the reciprocal lattice and the accuracy of the rule is described. This has led to results about the number of distinct lattice rules and the structure of embedded lattice rules, and has provided much of the underlying theory needed for a complete search for good (cost-effective) lattice rules.

However, the original approach to this theory (see e.g. Sloan 1985 and Sloan and Lyness 1989) is quite different, and uses a notation which appears to be entirely independent. This is a development of a standard form for the *number theoretic rule*

$$Qf = \frac{1}{N}\sum_{j=1}^{N} \bar{f}\left(\frac{j\mathbf{z}}{N}\right) \qquad \mathbf{z} \in \Lambda_0. \tag{1.3}$$

Here $\bar{f}(\mathbf{x})$ is a unit periodic continuation of $f(\mathbf{x})$ that coincides with $f(\mathbf{x})$ in the hypercube $[0,1)^s$. We note that

$$\bar{f}(\mathbf{x}) = f(\{\mathbf{x}\}), \tag{1.4}$$

where $\{\mathbf{x}\}$ has its conventional meaning as the vector whose components are the fractional parts of those of $\mathbf{x}$. Specifically,

$$\{\mathbf{x}\} \in [0,1)^s; \quad \{\mathbf{x}\} - \mathbf{x} \in \Lambda_0. \tag{1.5}$$

Number theoretic rules have been the subject of continuous and thorough investigation since their introduction in Korobov 1959. A recent survey of this work appears in Niederreiter 1988.

The number theoretic rule is itself a lattice rule. The key to understanding form (1.3) is to note the use of $\bar{f}$ in place of f. It appears that this has the effect of taking a set of N points arranged at equal intervals along a line in R^s, and translating each point individually so as to end up with a set of points which are distributed on an s-dimensional integration lattice Λ and comprise all the points on Λ which lie in $[0,1)^s$. One motivation for this paper is to illuminate the connection between this lattice Λ and any of its generator matrices A.

In the rest of this paper, t is a positive integer, D is a diagonal $t \times t$ matrix whose elements d_i are positive integers and Z is a $t \times s$ integer matrix whose rows are the vectors $\mathbf{z}_i$. What we term a *t-cycle D-Z form* of an s-dimensional lattice rule is an expression of the form

$$Qf = \frac{1}{d_1 d_2 \ldots d_t} \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} \ldots \sum_{j_t=1}^{d_t} \bar{f}\left(\frac{j_1 \mathbf{z}_1}{d_1} + \frac{j_2 \mathbf{z}_2}{d_2} + \ldots + \frac{j_t \mathbf{z}_t}{d_t}\right), \tag{1.6}$$

It is shown in theorem 2.1 of Sloan and Lyness 1989 that, so long as $t \geq 1$ and D is not singular, this form represents a lattice rule. This may is done by showing that all points lie on a lattice, that all points are in $[0,1)^s$, and that each point is assigned equal weight. The first two items are trivial. The third is straightforward, but leads us to one of the problems associated with this form. It may well happen that the same point occurs more than once in the summation. However, if this happens, then every point is repeated the same number of times. The form is termed *k-repetitive*, or simply repetitive, when each point is repeated $k > 1$ times, in which case $\nu(Q)$, the number of abscissas required by the rule, is given by $d_1 d_2 \ldots d_t / k$ which is of course an integer. Unfortunately, the proof given in that paper is not constructive. No immediate way of determining k from the elements of D and Z was then available.

The same rule may be expressed in a D-Z form in many different ways. The rule Q is defined to be of *rank* $r = r(Q)$ if it can be expressed in an r-cycle D-Z form, but not in an $(r-1)$-cycle D-Z form. An r-cycle form of Q is termed a *minimal* form. It is possible to express a rule of rank r in a non-repetitive minimal form in such a way that $d_{i+1} | d_i$ $i = 1, 2, \ldots, r-1$. When this is the case, the $\mathbf{z}_i$ are linearly independent, and the elements d_i are known as *invariants*. These are unique; that is, every lattice rule Q has a unique rank and unique set of invariants.

This nomenclature is taken from Group Theory. The abscissa set of a Lattice rule forms an Abelian group under addition modulo 1. The t-cycle D-Z form (1.6) corresponds to an expression of this group as a direct sum of cyclic groups, in accordance with the famous theorem of Kronecker 1877. This is known as the decomposition theorem. See also Hartley and Hawkes 1970, pp 153 et.seq. Many of the results of Sloan and Lyness are obtained as applications of the group theory based on this theorem.

Any minimal D-Z form of Q in which the nonzero elements of D are the invariants is

known as a *canonical form*. The vectors $\mathbf{z}_i$ are not unique, but are linearly independent.

In Section 4 we employ a trivial modification of this definition of a canonical form. It is clear from definition (1.6) that when any $d_j = 1$, the corresponding sum (over one element) may be omitted, whatever the corresponding vector $\mathbf{z}_j$ may be. In the sequel, on occasion, there will arise naturally what we term an s-cycle canonical form, where the rank of the rule is $r \leq s$. Such a canonical form has $d_j = 1$, $j \in [r+1, s]$; we may obtain a standard r-cycle canonical form by removing the final $s - r$ rows of D and Z and the final $s - r$ columns of D.

In this section, we have described very briefly the two principal approaches to lattice rule theory. These are through the lattice generator matrix A and through the t-cycle D-Z form respectively. This description is mainly to provide a proper background and a coherent list of definitions.

In Section 2 we outline the theory as it exists for relating one approach to the other. This is not a long section, as this problem has not been treated seriously before. Sections 3 and 4 contain new results, based on the Smith normal form of a matrix having rational elements. In Section 3, we show how to obtain a canonical form directly from B, a generator matrix of the reciprocal lattice. In Section 4 we describe a shorter calculation to obtain a canonical form from a possibly repetitive form that bypasses the explicit calculation of either B or A.

2 Some Relations between the Two Approaches

In the preceding section, we described two distinct ways of specifying a lattice rule. One requires a single generator matrix A. The other a pair of matrices, D and Z. Sections 3 and 4 are concerned with developing an elegant connection between these different specifications of the same rule, and between different D-Z specifications of the same rule. This section is devoted to the somewhat pedestrian methods currently available.

We note first that the lattice Λ formed by Qf in (1.6) includes all t points $\mathbf{z}_i/d_i$ $i = 1, 2, \ldots, t$ together with all points generated by them. These t points by themselves may not happen to generate an integration lattice. However, the expression (1.6) uses $\bar{f}$ in place of f. This implies that the fractional part of any of these t points also lies on the lattice Λ.

The effect of this is that the lattice has to contain the points of Λ_0, and so in total includes all points of the form

$$\mathbf{p} = \sum_{i=1}^{t} j_i \mathbf{z}_i / d_i + \sum_{i=1}^{s} k_i \mathbf{e}_i, \tag{2.1}$$

and any point expressible in this form is a member of Λ. (Here, as is conventional, $\mathbf{e}_i$ is the i-th unit s-vector.) In other words, the lattice Λ is generated by the rows of the $(t+s) \times s$ matrix

$$A^* = \begin{pmatrix} D^{-1}Z \\ I \end{pmatrix}. \tag{2.2}$$

However, as mentioned above, the same lattice is generated by any matrix obtainable from A^* by using elementary integer row operations. These have the same effect as premultiplying A^* by a unimodular $(t+s) \times (t+s)$ matrix V. Thus Λ is generated by the rows of any $s \times s$ matrix A satisfying

$$\begin{pmatrix} A \\ 0 \end{pmatrix} = V \begin{pmatrix} D^{-1}Z \\ I \end{pmatrix}. \tag{2.3}$$

A natural approach is to put A^* in upper triangular form, but any construction that results in t zero rows is sufficient.

Occasionally, one can pick out s rows from (2.2) by inspection. The following lemma may justify such a result.

Lemma 2.4. *Let $Q(\Lambda)$ be given by an s-cycle D-Z representation, and set $\tilde{A} = D^{-1}Z$; then if $\Lambda(\tilde{A})$ is an integration lattice, it is the integration lattice of Q.*

Proof. $\Lambda(\tilde{A})$ is generated by the rows of $\tilde{A}$. Thus it includes all points of the form

$$\mathbf{p} = \sum_{i=1}^{s} j_i \mathbf{z}_i / d_i \qquad \mathbf{j} \in \Lambda_0. \tag{2.4}$$

Since $\tilde{A}$ is an integration lattice, it includes all points $\mathbf{e}_i$ $i = 1, 2, \ldots, s$. Thus, specification (2.4) coincides with specification (2.1) of the lattice Λ. □

For example, if Z is known to be unimodular, the following theorem allows us to write down a generator matrix directly.

Theorem 2.5. *Let $Q(\Lambda)$ be given in an s-cycle D-Z representation with Z unimodular. Then this representation is nonrepetitive, and $A = D^{-1}Z$ is a generator matrix of Λ.*

Proof. Clearly,

$$B = (A^T)^{-1} = DZ^{T-1},$$

being the product of two integer matrices, is an integer matrix. Thus, A generates an integration lattice and, in view of the previous lemma, is the generator matrix of Λ. Moreover, since $|\det Z| = 1$, we find

$$d_1 d_2 \ldots d_s = \det D = |\det A|^{-1} = N,$$

where N is the number of distinct abscissas used by $Q(\Lambda)$. Thus, the D-Z form is not repetitive. □

The reverse process, that of obtaining an s-cycle D-Z form of $Q(\Lambda)$ from a given generator matrix A of the integration lattice Λ is also straightforward. Let $\mathbf{a}_r$ be a row of A and d_r be the smallest integer (or any integer) for which $\mathbf{z}_r = d_r \mathbf{a}_r \in \Lambda_0$. Then an s-cycle D-Z specification is given by the $s \times s$ matrix Z whose rows are $\mathbf{z}_r$ and $D = \mathrm{diag}\{d_1, d_2, \ldots, d_s\}$. (Unfortunately, this simple approach gives, in general, a highly repetitive D-Z form.)

We believe that Theorem 2.5 is new and in simple examples may be helpful in recognising a non-repetitive form. But what is particularly noticable in the results of this section is the absence of any *general* procedure for avoiding or recognising a repetitive form, or for producing a canonical form. A new way of carrying out these tasks, which leads directly to a canonical form, is given in the next section.

3 Reduction of B to D-Z Form

We noted earlier that the same lattice may have many different generator matrices. These are related by elementary integer *row* operations. That is, $B' = VB$ and B generate the same lattice when V is any unimodular matrix. Successive row operations may be used to put B into *upper triangular* lattice form (utlf), in which all elements are nonnegative, and the largest element in any column lies on the diagonal. This is essentially the Hermite normal form. It has been exploited in previous papers to count the number of lattice rules, to obtain information about sublattices and superlattices, and to form the basis of a search program for good lattice rules (see, e.g., Lyness, Sorevik, and Keast 1991).

As mentioned earlier, integer column operations applied to B (or *post*multiplication by unimodular matrices) result in a matrix that represents a different lattice. Nevertheless, if one allows column operations as well as row operations, one may put B into diagonal form. There are generally several ways of doing this though, of course, any such form has the same determinant (or product of nonzero diagonal elements). This procedure is significantly more involved than the procedure for the Hermite normal form but is reasonably straightforward. Since elementary operations may be used to interchange rows and columns, it is apparent that we may rearrange the order of these diagonal elements. However, there are, in addition, generally different possibilities for the set of diagonal elements. For example, the matrix given by

$$B = \begin{pmatrix} 7 & 14 & 21 \\ 35 & 73 & 117 \\ 7 & 20 & 66 \end{pmatrix} \tag{3.1}$$

can be reduced to diagonal form in many ways by using unimodular matrices. Two ways are as follows:

$$\begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 9 & -2 & 1 \end{pmatrix} \begin{pmatrix} 7 & 14 & 21 \\ 35 & 73 & 117 \\ 7 & 20 & 66 \end{pmatrix} \begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 21 \end{pmatrix}, \tag{3.2}$$

and:

$$\begin{pmatrix} 11 & -2 & 0 \\ -38 & 7 & 0 \\ 9 & -2 & 1 \end{pmatrix} \begin{pmatrix} 7 & 14 & 21 \\ 35 & 73 & 117 \\ 7 & 20 & 66 \end{pmatrix} \begin{pmatrix} -1 & -8 & 5 \\ 1 & 7 & -4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 21 & 0 \\ 0 & 0 & 21 \end{pmatrix}, \tag{3.3}$$

where the pre- and post-multiplying matrices are unimodular.

Apart from sign changes and from reordering the diagonal elements, these are, in fact, the only possibilities for diagonalizing this particular matrix B by unimodular transformations. This may be shown from theory developed in the nineteenth century. The *Smith*

normal form of B, denoted by snf(B), is a diagonalization of B using integer elementary row and column operations in which the nonzero diagonal entries satisfy $d_{j,j}/d_{i,i}$ = integer for all $j \geq i$. If the restriction that the diagonal entries be in non-decreasing order is removed, then the diagonal form is not unique. However, any ordering can be achieved by pre- and post-multiplication by permutation matrices, which are unimodular.

Theorem 3.4. (Smith, 1861) *Given a $t \times s$ matrix $\tilde{A}$ whose elements are rational numbers, there exist unimodular matrices V and U of sizes $t \times t$ and $s \times s$, respectively, such that*

$$\delta = snf(\tilde{A}) = V\tilde{A}U \tag{3.4}$$

is a $t \times s$ diagonal matrix having $\bar{t}$ non-zero elements which are rationals satisfying

$$\delta_{i+1,i+1}/\delta_{i,i} = \text{integer} \qquad i = 1, 2, \ldots, \bar{t} - 1, \tag{3.5}$$

The matrix δ is unique and is known as the Smith normal form of $\tilde{A}$. (But the matrices V and U are not unique.)

A convenient, accessible modern reference to this theory, which contains a brief proof of this theorem, is Schrijver 1986, pp. 40 et seq. A deeper treatment, set in the appropriate number theory context, appears in Newman 1972. Algorithms to obtain the Hermite normal form and the Smith normal form have been published; see, for example, Bradley 1971 and Kannan and Bachem 1979; in addition, Maple 1991 contains a procedure for finding D for integer matrices, (but not U or V).

The key theorem of this paper, which is a simple application of the theorem defining the Smith normal form, follows.

Theorem 3.5 *Let $Q(\Lambda)$ be an s-dimensional lattice rule, and let B be a generator matrix of the reciprocal lattice $\Lambda^{\perp}$. Then an s-cycle canonical form of $Q(\Lambda)$ is given by Z and D, where*

$$D = snf(B) = VBU \text{ and } Z = U^T, \tag{3.6}$$

U and V being unimodular.

Proof. Note that since B is an integer matrix, the elements of D are integers. Let us consider the lattice rule $Q(\Lambda')$ whose D-Z form comprises these particular matrices D and

Z. Since Z is unimodular, we may invoke Theorem 2.6 to establish that $D^{-1}Z$ is a generator matrix of Λ'. This being so, since V^T is unimodular $A = V^T D^{-1} Z$ is also a generator matrix of Λ' and so, by elementary manipulation B is a generator matrix of $\Lambda'^{\perp}$. Since the lattice generated by B is unique and its reciprocal is unique, Λ' coincides with Λ in the theorem. To establish the theorem, we note that, in view of Theorem 3.4, the elements of D have the divisibility property required for a canonical form. □

Corollary 3.7. *Every lattice rule $Q(\Lambda)$ has an s-cycle canonical form with Z a unimodular matrix.*

Corollary 3.8. *The invariants (and rank) of $Q(\Lambda)$ coincide with the non unit elements (and their number) of the Smith normal form of B, the generator matrix of the reciprocal lattice of Λ.*

Let us now return to the numerical example. The lattice rule $Q(\Lambda)$, whose reciprocal lattice is generated by B in (3.1), is of rank 2, has invariants $n_1 = n_2 = 21$ and may be expressed in canonical D-Z form with $\mathbf{z}_1 = (-8, 7, 0)$ and $\mathbf{z}_2 = (5, -4, 1)$.

Note that by means of permutation matrices (which are unimodular), we can rearrange the order of the diagonal elements in the Smith normal form. And, if we abandon the divisibility property, we can usually find other sets of diagonal elements. In example (3.1), only the two possibilities arise, because these two diagonal matrices are the only ones with determinant 441 that have the correct Smith normal form. The diagonal matrix diag{3,3,49} for example, cannot be obtained from B since it has Smith normal form diag{1,3,147}. This fact is worth mentioning, because when $D = VBU$ and D is diagonal but not necessarily in Smith Normal form, the rule $Q(\Lambda)$ is also defined by D and $Z = U^T$. This is also an s-cycle nonrepetitive form but is not necessarily canonical.

It is almost self evident that the Smith normal form of the reciprocal of any non-singular square matrix M is the reciprocal of the Smith normal form of M. It follows that the Smith normal form of the generator matrix $A = (B^T)^{-1}$ is the reciprocal of D in (3.6). One can write down immediately the correspondents of Theorem 3.5 and Corollary 3.8. These are:

Theorem 3.9. *Let A be a generator matrix of Λ, and let $\delta = snf(A) = VAU$, V and U being unimodular, then an s-cycle canonical form of $Q(\Lambda)$ is given by $D = inv(\delta)$ and $Z = U^{-1}$.*

Naturally, the unimodular matrices V and U occurring here are the transposes of the inverses of those in equation (3.6).

Corollary 3.10. *The invariants (and rank) of $Q(\Lambda)$ coincide with the inverses of the non-unit elements (and their number) of the Smith normal form of A, the generator matrix of Λ.*

Theorems 3.5 and 3.9 were discovered independently by Langtry 1992.

In this section we have provided a general method for obtaining a D-Z form from the generator matrix A. Unlike the cumbersome method described in section 2, this method provides a canonical form, specifying the rank and invariants of the lattice rule. In fact these quantities are provided without recourse to their possible application in constructing the summation in (1.6) for the D-Z form. They could have been defined through the Smith normal form, and their application noted afterwards. We shall return to this point in our concluding remarks.

4 Elimination of Repetition in D-Z Form

Application of the theory of the preceding section allows us to find a canonical D-Z form of $Q(\Lambda)$ directly from any generator matrix A of Λ or B of its reciprocal lattice using only the Smith normal form decomposition. This leads naturally to a solution to the problem of identifying and reducing a repetitive D-Z form. One could, in principle, follow the thrust of Section 2 and form $\tilde{A} = D^{-1}Z$, find A an upper triangular matrix by triangularizing $\begin{pmatrix} \tilde{A} \\ I \end{pmatrix}$, and use the Smith normal form to construct $D = VBU$. Each of these processes involves integer row and column operations, each set having a different immediate objective. In the rest of this section we show how these two sets of operations may be condensed into one set. To this end we describe a process which leads from a $t \times s$ matrix $\tilde{A} = D^{-1}Z$ to D_c and Z_c which specify an s-cycle canonical form of Q with Z_c unimodular.

Following Theorem 3.4, we set

$$\delta = snf(\tilde{A}) = V\tilde{A}U. \tag{4.1}$$

We recall that δ is a $t \times s$ diagonal matrix whose only nonzero elements are rationals satisfying

$$\delta_{i+1,i+1}/\delta_{i,i} = \text{integer} \qquad i = 1, 2, \ldots, \bar{t} - 1, \tag{4.2}$$

where $\bar{t} \leq \min(s, t)$ is the number of non-zero elements of δ.

Lemma 4.3. *Let the nonzero diagonal elements of δ in (4.1), expressed in their lowest terms, be*

$$\delta_{i,i} = m_i/n_i \qquad i = 1, 2, \ldots, \bar{t}. \tag{4.3}$$

Then

$$n_{i+1} | n_i \qquad i = 1, 2, \ldots, \bar{t} - 1. \tag{4.4}$$

Proof. The proof is elementary. From (4.2) we have

$$\frac{m_{i+1}}{n_{i+1}} \frac{n_i}{m_i} = \text{integer}.$$

Since n_{i+1} has no factor in common with m_{i+1}, it follows that n_{i+1} divides n_i. □

To proceed, we introduce s equations, each of which is an identity, and rewrite (4.1) in the form

$$\begin{pmatrix} V & 0 \\ 0 & U^{-1} \end{pmatrix} \begin{pmatrix} \tilde{A} \\ I \end{pmatrix} U = \begin{pmatrix} \delta \\ I \end{pmatrix}. \tag{4.5}$$

Here, as previously, I and U are $s \times s$ matrices and δ is a $t \times s$ diagonal matrix. The $\bar{t}$ nonzero diagonal elements of δ satisfy (4.3) and (4.4) above, with $(m_i, n_i) = 1$. We note that the $(t+s) \times (t+s)$ matrix on the left is unimodular. The thrust of the next lemma will be to provide a reduction in which the rational elements $\delta_{i,i}$ are replaced by integer inverse elements $1/n_i$.

Lemma 4.6. *For all m, n such that $(m, n) = 1$, there exists a 2×2 unimodular matrix V such that*

$$V \begin{pmatrix} m/n \\ 1 \end{pmatrix} = \begin{pmatrix} 1/n \\ 0 \end{pmatrix}. \tag{4.6}$$

Proof. Since $(m,n)=1$, there exist integers α,β such that $\alpha m+\beta n=1$. It is trivial to verify that

$$V=\begin{pmatrix}\alpha & \beta\\ -n & m\end{pmatrix} \tag{4.7}$$

satisfies (4.6) and has unit determinant. □

Corollary 4.7. *Let Δ be the $(t+s)\times s$ matrix on the right hand side of (4.3), its elements satisfying (4.4). Then there exists a $(t+s)\times(t+s)$ unimodular matrix $V^{(w)}$ such that $V^{(w)}\Delta$ differs from Δ only in the (w,w) element, which is replaced by $1/n_w$ and in the (w+t,w) element, which is replaced by zero.*

Proof. $V^{(w)}$ differs from the unit matrix I only in that the four elements required to carry out row operations on rows w and $t+w$ are replaced by the four in Lemma 4.6, with m_w and n_w replacing m and n. □

Theorem 4.8. *Given a $t\times s$ rational-valued matrix $\tilde{A}$, there exists an $s\times s$ unimodular matrix U and a $(t+s)\times(t+s)$ unimodular matrix $\tilde{V}$ having the property that*

$$\tilde{V}\begin{pmatrix}\tilde{A}\\ I\end{pmatrix}=\begin{pmatrix}\tilde{\delta}\\ J\end{pmatrix}U^{-1}, \tag{4.8}$$

where $\tilde{\delta}$ is a diagonal $t\times s$ matrix whose nonzero elements satisfy

$$\tilde{\delta}_{ii}=1/n_i \qquad i=1,2,\ldots,\bar{t}\le t$$

with integer n_i, where

$$n_{i+1}|n_i \qquad i=1,2,\ldots,\bar{t}-1, \tag{4.9}$$

and each row of J either is 0 or is $\mathbf{e}_u$ with $u>\bar{t}$.

Proof. As mentioned before, (4.1) is equivalent to (4.3). We may premultiply successively by $V^{(1)},V^{(2)},\ldots,V^{(\bar{t})}$, these being defined in Corollary 4.7. The effect on the left-hand side of (4.3) is to replace $\begin{pmatrix}V & 0\\ 0 & U^{-1}\end{pmatrix}$ by

$$\tilde{V}=V^{(\bar{t})}V^{(\bar{t}-1)}\ldots V^{(2)}V^{(1)}\begin{pmatrix}V & 0\\ 0 & U^{-1}\end{pmatrix}, \tag{4.10}$$

which is obviously a $(t+s)\times(t+s)$ unimodular matrix. The effect on the right-hand side is to successively replace the only nonzero element in the w-th row by $1/n_w$ and the $(w+t)$-th row by zero leaving a matrix of the form given by the left member of the right-hand side of (4.7). This establishes the theorem. □

Our major result follows simply from Theorem 4.8.

Theorem 4.11. *Let $Q(\Lambda)$ be given in a t-cycle D-Z form, and let $\tilde{A} = D^{-1}Z$. Let the Smith normal form of $\tilde{A}$ be $\delta = V\tilde{A}U$ and the nonzero elements of δ be $\delta_{i,i} = m_i/n_i$, $i = 1, 2, \ldots, \bar{t}$ in their lowest terms. Then an s-cycle canonical form of $Q(\Lambda)$ is given by*

$$D_c = \mathrm{diag}\{n_1, n_2, \ldots, n_{\bar{t}}, 1, \ldots, 1\}, \text{ and } Z_c = U^{-1}. \tag{4.11}$$

Proof. As discussed in Section 2, the lattice Λ is generated by the rows of $\begin{pmatrix} \tilde{A} \\ I \end{pmatrix}$. Since this is invariant under premultiplication by a unimodular matrix, this lattice is generated by the rows of the $(t+s)\times s$ matrix on the left-hand side of (4.8), which coincides with the $(t+s)\times s$ matrix on the right of (4.8). The $\bar{t}$ zero rows of J clearly play no part in this lattice generation and may be removed. The other $s-\bar{t}$ rows may be reordered in a natural way. We may identify D_c^{-1} and Z_c with the matrices remaining on the right-hand side of (4.8). □

In the expression for D_c in (4.11), there are $s-\bar{t}$ unit elements displayed. Besides these, some of the integers denoted by n_i may also be unity. The rank r is of course the number of non-unit diagonal elements in D_c and may be less than $\bar{t}$ which itself by definition cannot exceed min(s,t).

With this theorem at hand, we summarize the possible synthesis of an s-dimensional lattice rule $Q(\Lambda)$ given in a t-cycle D-Z form.

(a) Construct the $t\times s$ matrix $\tilde{A} = D^{-1}Z$.

(b) Construct the Smith Normal Form $\delta = V\tilde{A}U$ of $\tilde{A}$. This is a diagonal $t\times s$ matrix

(c) Put the elements of δ in its lowest terms, that is, $\delta_{i,i} = \frac{m_i}{n_i}$ with $(m_i, n_i) = 1$. Then a canonical form of $Q(\Lambda)$ is given by

$$D_c = \mathrm{diag}\{n_1, n_2, \ldots, n_{\bar{t}}, 1, \ldots, 1\}, \text{ and } Z_c = U^{-1}.$$

If one simply requires the invariants (perhaps to determine the rank), one does not need to calculate U. However, in the Smith normal form reduction, the calculation of U (or of U^{-1} or both) can be effected in situ.

(d) If one requires a generator matrix of Λ or of $\Lambda^{\perp}$, one calculates $A = D_c^{-1}U^{-1}$, or $B = D_c U^T$.

5 Numerical Examples

Our first example is to determine a canonical D-Z form of the three-dimensional integration lattice defined by five lattice points:

$$\mathbf{z}_j = \left(\frac{1}{3j-1}, \frac{1}{3j}, \frac{1}{3j+1}\right) \qquad j = 1, 2, 3, 4, 5. \tag{5.1}$$

We shall determine in passing whether the lattice generated by these five points is an integration lattice; that is, are the unit vectors $\mathbf{e}_i$ already included? Our matrix $\tilde{A}$ in (a) above is

$$\tilde{A} = D^{-1}Z = \begin{pmatrix} 1/2 & 1/3 & 1/4 \\ 1/5 & 1/6 & 1/7 \\ 1/8 & 1/9 & 1/10 \\ 1/11 & 1/12 & 1/13 \\ 1/14 & 1/15 & 1/16 \end{pmatrix} = \frac{1}{720720} \begin{pmatrix} 360360 & 240240 & 180180 \\ 144144 & 120120 & 102960 \\ 90090 & 80080 & 72072 \\ 65520 & 60060 & 55440 \\ 51480 & 48048 & 45045 \end{pmatrix}. \tag{5.2}$$

We now construct the Smith normal form

$$\delta = V\tilde{A}U = \mathrm{diag}\{1/720720, 1/280, 3/20\}, \tag{5.3}$$

and the inverse of the matrix U used in the reduction

$$Z_c = U^{-1} = \begin{pmatrix} 385164 & 148148 & 99 \\ -33120301 & -12739230 & -8513 \\ 11831180 & 4550687 & 3041 \end{pmatrix}. \tag{5.4}$$

The *invariants* are given by the denominators in δ. That is,

$$D_c = \mathrm{diag}\{720720, 280, 20\}. \tag{5.5}$$

D_c and Z_c give a canonical form in D, Z notation.

The elements $\mathbf{z}_i/n_i$ in this canonical form are given by the rows of $A = D_c^{-1} Z_c$. These elements are the rows of Z_c above divided by the corresponding elements of D_c above.

This matrix A contains elements greater in magnitude than unity. Since the lattice generated by A contains all elements of Λ_0, we may add or subtract any integer from any element of A. (In this case, this procedure corresponds to premultiplication by some unimodular matrix M.) We obtain in this way a matrix all of whose elements lie in $[0, 1)$, namely,

$$A_c' = \begin{pmatrix} 823/1540 & 37/180 & 1/7280 \\ 59/280 & 3/4 & 167/280 \\ 0 & 7/20 & 1/20 \end{pmatrix}. \tag{5.6}$$

Note that the lattice generated by the five rows of $\tilde{A}$ is not an integration lattice. We know this because the diagonal elements of δ, in their lowest terms, are not all inverse integers. Ignoring the numerator 3 in the element $\delta_{33} = 3/20$ has the effect of increasing the lattice density by this factor.

As a second example, we give a rule in Sloan-Lyness form, which does not obviously appear repetitive, but which can easily be shown to be so, by using the above construction.

Let

$$Qf = \frac{1}{81} \sum_{j_1=1}^{9} \sum_{j_2=1}^{9} \bar{f}\left(\frac{j_1(0,8,4)}{9} + \frac{j_2(6,5,7)}{9} \right). \tag{5.7}$$

We wish to find out whether this is repetitive and, if so, to put it in nonrepetitive form.

We set

$$\tilde{A} = D^{-1} Z = \begin{pmatrix} 1/9 & 0 \\ 0 & 1/9 \end{pmatrix} \begin{pmatrix} 0 & 8 & 4 \\ 6 & 5 & 7 \end{pmatrix}$$

and find

$$\delta = \mathrm{snf}(\tilde{A}) = V \tilde{A} U = \begin{pmatrix} 1/9 & 0 & 0 \\ 0 & 4/3 & 0 \end{pmatrix}$$

with

$$U^{-1} = \begin{pmatrix} -6 & 11 & 1 \\ -2 & 3 & 0 \\ 1 & -1 & 0 \end{pmatrix}.$$

The fact that the product of the denominators in δ differs from the corresponding product in D^{-1} indicates that (5.7) is repetitive. To obtain a nonrepetitive form is straightforward. We proceed as before to construct $\tilde{\delta}$ from δ by removing the nonunit numerator. The rows of U^{-1} are then taken as the elements $\mathbf{z}_r$; we find after some trivial reduction that

$$Qf = \frac{1}{27} \sum_{j_1=1}^{9} \sum_{j_2=1}^{3} \bar{f}\left(\frac{j_1(3,2,1)}{9} + \frac{j_2(1,0,0)}{3} \right),$$

which is a nonrepetitive form of Q.

6 Concluding Remarks

From a technical point of view, the results in this paper merely show how to carry out various standard tasks relating to the manipulation of lattice rules. The tool is a standard technique to obtain the Smith Normal Form of an integer matrix. Using this normal form, we can readily find a Sloan-Lyness canonical form of $Q(\Lambda)$ from a generator matrix of Λ. And we can determine whether a given form of $Q(\Lambda)$ is repetitive by reducing it to a canonical form.

However, we believe that this paper has wider implications. The Smith Normal Form of an integer matrix is in fact the link between two apparently almost independent approaches to the theory of lattice rules. This is because the Smith normal form is a standard tool in the proof of the Kronecker decomposition theorem. The referee has pointed out that there is a sense in which this paper is in effect traversing a part of the proof of the decomposition theorem. In our opinion the principal virtue of the theorems in this paper is that they unite these two parts of the same theory to their mutual benefit.

References

G. H. Bradley, 1971. Algorithms for Hermite and Smith normal matrices and linear diophantine equations, Math. Comput. 25 (1971) 897–907.

B. Hartley and T.O. Hawkes, 1970. Rings, modules and linear algebra, Chapman and Hall (1970).

R. Kannan and A. Bachem, 1979. Polynomial algorithms for computing the Smith and Hermite Normal Forms of an integer matrix, SIAM J. Comput. 8 (1979) 499–507.

N. M. Korobov, 1959. The approximate computation of multiple integrals (Russian), Dokl. Akad. Nauk. SSSR 124 (1959) 1207–1210.

T.N.Langtry, 1992. The determination of canonical forms for lattice quadrature rules, submitted to J. Comp. Appl. Math.

J.N. Lyness, 1989. An introduction to lattice rules and their generating matrices, IMA J. Numer. Anal. 9 (1989) 405-419.

J. N. Lyness, T. Sorevik, and P. Keast, 1991. Notes on integration and integer sublattices, Math. Comput. 56 (1991) 243–255.

Maple V, 1991. Maple V language reference manual, Springer Verlag (1991).

M. Newman, 1972. Integral matrices, Academic Press Inc., New York (1972).

H. Niederreiter, 1988. Quasi-Monte Carlo methods for multidimensional numerical integration, in International Series of Numerical Mathematics, Vol. 85, Numerical Integration III, G. Hämmerlin and H. Brass, eds., Birkhauser Verlag, Basel, 1988, pp. 157–171.

A. Schrijver, 1986. Theory of linear and integer programming, Wiley & Sons (1986).

I.H. Sloan, 1985. Lattice methods for multiple integration, J. Comput. Appl. Math. 12 (1985) 131–143.

I. H. Sloan and J. N. Lyness, 1989. The representation of lattice quadrature rules as multiple sums, Math. Comput. 52 (1989) 81–94.

H. J. S. Smith, 1861. On systems of linear indeterminate equations and congruences, Phil. Trans. Roy. Soc. London (A) 151, (1861) 293–326.