Acta Cryst. (1977). A33, 750–757

On the Application of Phase Relationships to Complex Structures. XI. A Theory of Magic Integers

BY PETER MAIN

Department of Physics, University of York, York YO1 5DD, England

(*Received* 9 *February* 1977; *accepted* 8 *April* 1977)

Magic integers enable several unknown phases to be expressed in terms of a single variable at the expense of introducing some error in the phases represented. The economy of variables is particularly useful in a multisolution direct-methods program like *MULTAN* where a successful solution may depend upon the ability to use a large number of phases at the beginning of phase determination and the computing time is proportional to the number of variables. Formulae are presented which give the phase errors in the magic-integer representation. A recipe is given for the generation of efficient magic-integer sequences in which the r.m.s. error is spread evenly over all the phases represented. These sequences minimize the overall phase error for a given maximum integer in the sequence. It is found that the errors are minimized when the differences between adjacent members of the magic-integer sequence form the terms of a geometric progression and the smallest integer is greater than half the largest.

Introduction

The concept of magic integers was introduced by White & Woolfson (1975) in paper VII of this series and an extension of their use was described by Derclercq, Germain & Woolfson (1975) in paper VIII.

For a sequence of n integers, denoted by $m_1, m_2, \ldots, m_n$, n phases are represented by the equations

$$\varphi_i = m_i x \bmod (1), \quad i = 1 \text{ to } n, \tag{1}$$

where the phases, φ_i, are in cycles and the set of equations is approximately satisfied by some value of x in the range $0 < x < 1$. Trials by White & Woolfson showed that the phase error in the magic-integer representation was acceptably low for sequences in which $n < 8$ and they demonstrated the use of magic integers in the solution of two known structures. Declercq, Germain & Woolfson described a more powerful method of using these ideas, called the *P–S* set method, which allows a large number of phases to be used simultaneously at the beginning of phase determination. The power of this method has been amply demonstrated by the solution of several crystal structures which had previously been very difficult or even impossible for *MULTAN* to solve (private communication from Dr G. Germain, 1976; see also paper VIII).

This paper critically examines the errors involved in the magic-integer representation of phases. It is observed that some integer sequences give rise to larger phase errors than other, apparently similar, sequences and a method of generating the more efficient magic-integer sequences is given. An analytical method of calculating the r.m.s. phase error for any particular sequence is described and a formula is developed which gives a lower bound to the r.m.s. error.

The magic-integer representation

A convenient and fruitful approach to the magic-integer representation of phases is to regard the set of n equations (1) as the parametric equation of a straight line in n-dimensional phase space. The most trivial example of this is a magic integer sequence of length one, so that a single phase is represented by

$$\varphi_1 = x. \tag{2}$$

In this case, the straight line fills the whole of the one-dimensional phase space in which it exists and each possible phase value corresponds exactly to a point on the line. The equation (2) can therefore represent the phase φ_1 in terms of the variable x without any error at all.

The first non-trivial case occurs with an integer sequence of length two as in

$$\left.\begin{aligned} \varphi_1 &= m_1 x \\ \varphi_2 &= m_2 x \end{aligned}\right\} \tag{3}$$

which is the equation of a straight line in two-dimensional phase space. For the special case where $m_1 = 2$, $m_2 = 3$, the line represented by (3) is shown in Fig. 1(*a*). When the phases are measured in cycles, modulo (1), and x is restricted to $0 < x < 1$, the line becomes that shown in Fig. 1(*b*). That is, it becomes a family of parallel straight lines of equal spacing in a box bounded by $\varphi_1 = 0, 1$ and $\varphi_2 = 0, 1$. Now, any pair of phase values (φ_1, φ_2) can be represented by a point in the box and the nearest point to (φ_1, φ_2) on any line will give the best magic-integer representation of the phases in a least-squares sense. For example, the point A in Fig. 1(*b*) is (0·600, 0·250) and the nearest point to A on the straight line is B. A simple calculation shows that B is the point (0·531, 0·296) and the value of x at B is 0·765. The errors in the magic-integer representa-

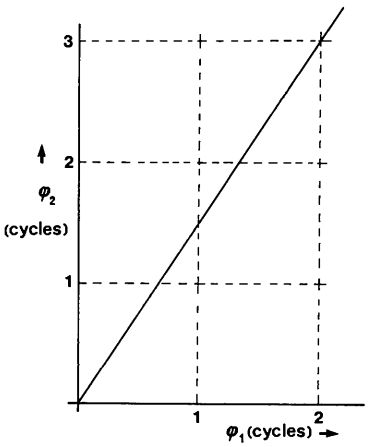
tion of these phases are therefore −0·069 and 0·046 cycles, *i.e.* −24·9° and 16·6° respectively. The length of the line AB is given by

$$AB=(\Delta\varphi_1^2+\Delta\varphi_2^2)^{1/2} \quad (4)$$

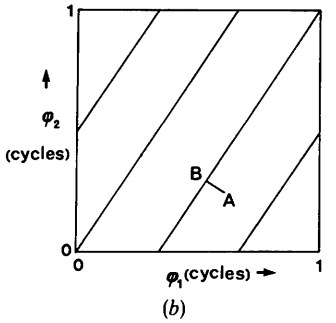
where $\Delta\varphi_1$ and $\Delta\varphi_2$ are the phase errors involved. Because B has been chosen to be the nearest point on the line to A, the phases represented by B are seen to be those which minimize the sum of the squares of the errors.

It should be clear from Fig. 1(*b*) that, provided the integers have no common factor, the higher the integers used, the closer together will be the lines in phase space. It follows from this that the larger the integers, the smaller will be the expected overall error in the phases. In addition, the more nearly equal the integers are, the more evenly divided will be the error between the individual phases. These conclusions are only strictly true for the sequences of length two just considered. They do, however, indicate tendencies for longer sequences and it is useful to bear them in mind.

The generalization of the magic-integer representation to n phases is now easily made. The set of equations



(*a*)



(*b*)

Fig. 1. (*a*) The line represented by $\varphi_1=2x$, $\varphi_2=3x$ in φ_1, φ_2 space. (*b*) Graph of $\varphi_1=2x$, $\varphi_2=3x$, mod(1), $0<x<1$. The pair of phase values (0·600, 0·250) is represented by the point A. The nearest point to A on any line is B where $x=0{\cdot}765$, giving $\varphi_1=0{\cdot}531$, $\varphi_2=0{\cdot}296$.

(1) represents a family of parallel straight lines in n-dimensional phase space. Any point $A(\varphi_1,\varphi_2,\ldots,\varphi_n)$, in this space represents a set of phase values and their magic-integer representation is given by the point on one of the lines which is nearest to A. The square of the shortest distance between A and the nearest line gives the sum of the squares of the phase errors.

Calculation of r.m.s. error

Since the magic-integer representation introduces errors in the phases, it is necessary to investigate the sizes of the errors involved. In particular, it is of interest to calculate the r.m.s. phase error of any given sequence, *i.e.* the r.m.s. error of all the phases taken together and the r.m.s. error of each individual phase.

Referring to Fig. 1(*b*), any point in phase space which is the same distance from the line as the point A will give rise to the same phase errors. When studying errors, therefore, it is convenient to consider the projection down the family of parallel lines onto a line parallel to AB. This consists of a one-dimensional lattice of points whose spacing is given by the spacing of the lines in Fig. 1(*b*). The r.m.s. phase error of the magic-integer sequence is now related to the r.m.s. distance of all points in the projection from the nearest lattice point. If the integers are m_1 and m_2, simple geometry shows that the spacing of the lines, and hence the lattice spacing in the projection, is given by:

$$\text{lattice spacing}=a=1/(m_1^2+m_2^2)^{1/2}. \quad (5)$$

The maximum distance of any point from the nearest lattice point must be half of this, *i.e.* $a/2$. The mean square distance of any point from the nearest lattice point is therefore given by:

$$\text{mean square distance}=\frac{2}{a}\int_0^{a/2}x^2\mathrm{d}x=\frac{a^2}{12}. \quad (6)$$

Substituting the value of a from (5) leads to the root-mean-square distance:

$$\text{r.m.s. distance}=\frac{1}{2\sqrt{3}(m_1^2+m_2^2)^{1/2}}. \quad (7)$$

This represents the root mean sum of the squares of the errors of two phases (measured in cycles). If the errors are assumed to be evenly distributed between the phases, the r.m.s. error of a typical phase represented by the sequence is half of the quantity in (7) above. Converting to radians, the r.m.s. error of a typical phase is therefore:

$$\Delta\varphi_{\text{rms}}=\frac{\pi}{2\sqrt{3}(m_1^2+m_2^2)^{1/2}}\ \text{radians}. \quad (8)$$

Alternatively, the phase error may be resolved into components in the φ_1 and φ_2 directions and the r.m.s. error of each phase individually can be shown to be:

$$\text{r.m.s. error in }\varphi_1=\frac{\pi}{\sqrt{3}}\frac{m_2}{m_1^2+m_2^2}\ \text{radians} \quad (9a)$$

$$\text{r.m.s. error in } \varphi_2 = \frac{\pi}{\sqrt{3}} \frac{m_1}{m_1^2+m_2^2} \text{ radians}. \quad (9b)$$

This calculation may be generalized in principle to sequences of length n as follows. The projection down the family of parallel lines represented by (1) is a lattice of points in $(n-1)$-dimensional space. The mean square distance, $\overline{d^2}$, of any point in the $(n-1)$-dimensional space from the nearest lattice point is then the mean sum of the squares of the errors in the n phases. Assuming the error to be evenly distributed among the phases, the r.m.s. error of a typical phase is given by:

$$\Delta\varphi_{\text{rms}} = 2\pi\sqrt{\frac{\overline{d^2}}{n}} \text{ radians}. \quad (10)$$

The field of integration for the calculation of $\overline{d^2}$ must now be defined. This is the $(n-1)$-dimensional volume around each lattice point such that all points within the volume are closer to the lattice point at its centre than any other. It is the polyhedron, known in n-dimensional geometry as the Voronoi polyhedron (Voronoi, 1908; see also Rogers, 1964, pp. 74–80), whose plane faces are perpendicular bisectors of the lines between the lattice point and each of its near neighbours. The mean square distance is then given by:

$$\overline{d^2} = \frac{1}{V}\int_V r^2 \mathrm{d}V, \quad (11)$$

where V is the volume (or content) of the $(n-1)$-dimensional Voronoi polyhedron and r is the distance of a general point within V from its geometric centre. The equations (10) and (11) now define the r.m.s. error of any magic integer sequence of length n.

The error distributions for individual phases can also be obtained from the Voronoi polyhedron. Taking the origin as the geometric centre of the polyhedron, the probability of an error $\Delta\varphi_i$ in φ_i is proportional to the content of the $(n-2)$-dimensional section at $\varphi_i = \Delta\varphi_i$. The r.m.s. error of each individual phase can then be obtained from the corresponding probability density. This can most easily be expressed as:

$$\text{r.m.s. error in } \varphi_i = 2\pi\left(\frac{1}{V}\int_V r_i^2 \mathrm{d}V\right)^{1/2} \text{ radians}, \quad (12)$$

where r_i is the distance in the φ_i direction of a general point in the Voronoi polyhedron from its geometric centre. The field of integration is the same as in (11).

Unfortunately, the field of integration is not easily expressed in terms of the magic integers, except in the case of $n=2$ which has already been dealt with. It follows that no general formula has been found which simply relates the integers to the corresponding r.m.s. phase errors. The complete calculations must be performed for each sequence separately.

As an illustration of such a calculation, let us take the sequence 4 6 7. The projection down the lines these integers produce in three-dimensional phase space is the two-dimensional lattice shown in Fig. 2(*a*). The figure also gives the dimensions of the lattice. An enlarged portion of this lattice is shown in Fig. 2(*b*) in which has been constructed the two-dimensional Voronoi polyhedron – in this case a hexagon. The edges of the hexagon are perpendicular bisectors of the lines joining the central lattice point to its near neighbours and, clearly, every point in the hexagon is closer to the central lattice point than to any other. The vertices of the hexagon are at a maximum distance from the lattice points and must therefore represent the maximum phase errors which can occur with the sequence 4 6 7. Taking the central lattice point, O, as the origin of phase space, the coordinates of the vertices are given in Table 1. From this, the maximum possible phase errors in $\varphi_1, \varphi_2, \varphi_3$ are seen to be 65·9, 55·2, 46·3° respectively. Note that the largest error is in the phase represented by the smallest integer and conversely for the smallest error. The mean square distance, $\overline{d^2}$, of points within this hexagon from the origin, O, (calculated as in equation 11) is found to be 0·0163 cycles^2 and, from (10), this corresponds to an r.m.s. error for the sequence of 26·6°.

Table 1. *Coordinates* (°) *of the vertices of the hexagon in Fig.* 2(*b*)

The central lattice point is at (0,0,0)

	φ_1	φ_2	φ_3
A	65·9	−23·3	−17·8
B	51·7	19·6	−46·3
C	−1·8	55·2	−46·3
D	−65·9	23·2	17·8
E	−51·7	−19·6	46·3
F	1·8	−55·2	46·3

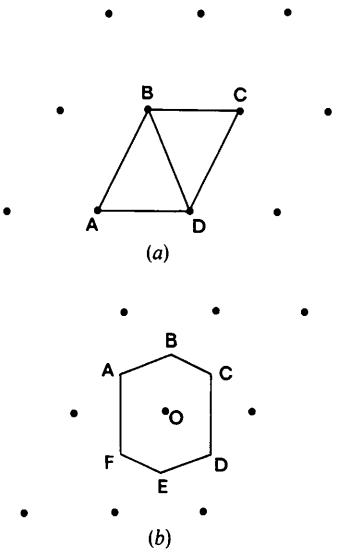


Fig. 2. (*a*) The two-dimensional lattice produced by the sequence 4 6 7. Dimensions of the lattice in cycles and degrees are: AB = 0·359; BD = 0·372; DA = 0·299; $\angle DAB$ = 68·3°; $\angle ABD$ = 48·2°; $\angle BDA$ = 63·5°. (*b*) The field of integration (Voronoi polyhedron) to calculate the r.m.s. error for the sequence 4 6 7.

The probability density distributions of the errors in the three phases can be obtained as already described. In this case, the probability of an error $\Delta\varphi_i$ in φ_i is proportional to the length of the line of intersection of the hexagon with the plane normal to the φ_i axis at $\varphi_i=\Delta\varphi_i$. The probability density functions obtained in this way are shown in Fig. 3. From these, the r.m.s. errors in $\varphi_1,\varphi_2,\varphi_3$ can be obtained as 30·4, 24·3, 24·6° respectively.

Lower bound of r.m.s. error

As the Voronoi polyhedron cannot be simply described in terms of the magic integers, making calculations of r.m.s. error extremely clumsy, it becomes necessary to make some approximations in order to produce a general formula for the r.m.s. error. A useful simplification is to take the field of integration in (11) as an $(n-1)$-dimensional hypersphere of the same volume as the polyhedron. This will always give a value of $\overline{d^2}$ which is a minimum for the volume considered and therefore will lead to a lower bound of the r.m.s. error as calculated from (10).

A formula for the lower bound of the r.m.s. error of a sequence, $\Delta\varphi_{1b}$, can be developed as follows. The family of lines in n-dimensional phase space intersect the $(n-1)$-dimensional section perpendicular to the φ_j axis in m_j points arranged on a lattice. The $(n-1)$-dimensional volume associated with each point is therefore $1/m_j$. If this section makes an angle θ_j with the direction of the lines, the volume associated with each lattice point in the section perpendicular to the lines must be $\cos\theta_j/m_j$. That is, the volume, V, of the Voronoi polyhedron is:

$$V=\frac{\cos\theta_j}{m_j}. \tag{13}$$

The value of $\cos\theta_j$ may be obtained from the inner product of the unit vectors representing the directions of the two sections,

$$i.e.\ \cos\theta_j=\frac{(m_1,m_2,\ldots,m_j,\ldots,m_n).(0,0,\ldots,1,\ldots,0)}{(\sum_i m_i^2)^{1/2}}$$

$$=\frac{m_j}{(\sum_i m_i^2)^{1/2}}. \tag{14}$$

From (13) and (14), the volume of the Voronoi polyhedron is:

$$V=\frac{1}{\left(\sum_{i=1}^{n} m_i^2\right)^{1/2}}. \tag{15}$$

Let us assume that this $(n-1)$-dimensional volume is a hypersphere centred on a lattice point. The volume of an $(n-1)$-dimensional hypersphere of radius a is:

$$V=\frac{2a^{n-1}}{n-1}.\frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)}. \tag{16}$$

So from (15) and (16) we obtain the radius of the hypersphere as:

$$a=\frac{1}{\pi^{1/2}}\left[\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(\sum_{i=1}^{n} m_i^2\right)^{1/2}}\right]^{1/(n-1)}. \tag{17}$$

The mean-square distance, $\overline{d^2}$, of all points in the hypersphere from its centre is obtained by using an $(n-1)$-dimensional spherical polar coordinate system. Equation (11) then becomes:

$$\overline{d^2}=\frac{(n-1)\Gamma\left(\frac{n-1}{2}\right)}{2a^{n-1}\pi^{(n-1)/2}}\int_0^a r^2\mathrm{d}r\int_0^\pi \mathrm{d}\theta_1\int_0^\pi \mathrm{d}\theta_2,\ldots,\int_0^{2\pi}\mathrm{d}\theta_{n-2}$$
$$\times r^{n-2}\cos^{n-3}\theta_1\cos^{n-4}\theta_2,\ldots,\cos\theta_{n-3}. \tag{18}$$

The variables are separable and the integrations can be carried out as in Kendall (1961, p. 35). The result is:

$$\overline{d^2}=\frac{(n-1)\Gamma\left(\frac{n-1}{2}\right)}{2a^{n-1}\pi^{(n-1)/2}}.\frac{a^{n+1}}{n+1}.\frac{\pi^{(n-3)/2}}{\Gamma\left(\frac{n-1}{2}\right)}.2\pi=\frac{(n-1)a^2}{n+1}. \tag{19}$$

Using (10), (17) and (19), we obtain an expression for the lower bound of the r.m.s. error, $\Delta\varphi_{1b}$, as

$$\Delta\varphi_{1b}=2\sqrt{\frac{\pi(n-1)}{n(n+1)}}\left[\frac{\Gamma\left(\frac{n+1}{2}\right)}{(\sum_i m_i^2)^{1/2}}\right]^{1/(n-1)}\ \text{radians}. \tag{20}$$

The actual r.m.s. error of a sequence for which $n>2$ must always be greater than $\Delta\varphi_{1b}$, but the lower bound

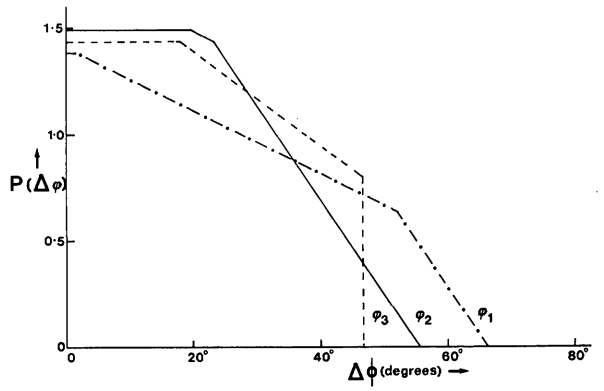


Fig. 3. Probability density distributions of the errors in individual phases represented by $\varphi_1=4x$, $\varphi_2=6x$, $\varphi_3=7x$, mod(1).

forms a very useful check on its quality. The closer the r.m.s. error is to the lower bound, the more efficient will the sequence be. For $n=1$, equation (20) gives $\Delta\varphi_{\rm lb}=0$, which has already been deduced from (2). Also, for $n=2$, equation (20) reduces to the exact formula (8) for this case. Taking the sequence 4 6 7, equation (20) gives the lower bound as 26·2°. Since the exact r.m.s. error of 26·6° is only slightly greater than this, 4 6 7 is seen to be a particularly efficient sequence.

Note how $\Delta\varphi_{\rm lb}$ depends upon the length of the sequence and the integers themselves. Using the main term in Stirling's formula for the gamma function, *i.e.*

$$\Gamma(x)\doteqdot\left(\frac{x}{e}\right)^x\sqrt{\frac{2\pi}{x}}, \tag{21}$$

we obtain from (20) for large n:

$$\Delta\varphi_{\rm lb}\doteqdot\sqrt{\frac{2\pi}{e}}\left[\frac{1}{\sum_i m_i^2}\right]^{1/2(n-1)}. \tag{22}$$

Clearly, the larger the integers the smaller will be the lower bound, but $\Delta\varphi_{\rm lb}$ becomes more insensitive to the size of the integers as n increases. In addition, $\Delta\varphi_{\rm lb}$ increases with n for the same size of integers.

Efficient magic-integer sequences

We now come to the important question of how to predict efficient magic-integer sequences. An efficient sequence must fulfil the following criteria as well as possible:

(*a*) the integers should be small;
(*b*) the overall r.m.s. error should be small;
(*c*) the r.m.s. error should be evenly divided among the phases represented.

The criteria (*a*) and (*b*) are mutually opposed and a compromise must be made according to the particular application of the sequence. In the extreme case of criterion (*c*) where all phases have identical errors, the Voronoi polyhedron must be a hypersphere and the r.m.s. error is given by the lower-bound formula (20). It is clearly impossible to produce a lattice in which the unit cell is a hypersphere, but a good approximation is produced if the $n(n-1)$ nearest neighbours of each lattice point are equidistant. Even this is impossible in $(n-1)$-dimensional space, but it is an ideal which must be approached as closely as possible. Obviously, (*c*) may be used as a single criterion in place of both (*a*) and (*b*).

In order to reduce the number of variables in the problem, let us seek the most efficient sequence of integers, *i.e.* that which minimizes $\Delta\varphi_{\rm rms}$, for which the largest integer, m_n, is a fixed value. This also turns out to be a very convenient way of thinking about magic-integer sequences in practical applications.

It is already clear that the larger the integers, the smaller will be the r.m.s. phase error. The first rule for creating efficient sequences is therefore quite obvious – that, with the largest integer kept fixed, the remaining integers in the sequence should be as large as possible.

It has also been observed that the more nearly equal the integers in the sequence are, the more evenly will the error be distributed among the phases. The second rule for creating efficient sequences must therefore be that the integers should be as close to each other in value as possible.

A third rule for creating efficient sequences is also required. If any integer in the sequence is complemented, *i.e.* if m_i is changed to m_n-m_i, the effect in n-dimensional phase space is merely to reverse the direction of the ith axis. No other changes take place which will affect the quality of the sequence. The result of this change is to decrease the lower-bound r.m.s. error, $\Delta\varphi_{\rm lb}$, if $m_n-m_i>m_i$, *i.e.* if the complemented integer is larger than the original. Conversely, $\Delta\varphi_{\rm lb}$ rises if $m_n-m_i<m_i$. The effect on the actual r.m.s. error is the same and this is consistent with the earlier observation that the larger the integers the smaller the phase errors. This makes it possible to minimize the r.m.s. error, without increasing the size of the largest integer, by complementing all integers that are less than half the maximum. The resulting sequence will have $m_1>m_n/2$.

The fourth, and very important, rule which efficient sequences must obey is best illustrated by an actual example. We already know 4 6 7 to be a good sequence and its two-dimensional lattice is shown in Fig. 2(*a*). It can be seen that the lattice is close to hexagonal, which approaches most closely the impossible ideal in two-dimensional space of criterion (*c*). It has already been shown that the actual r.m.s. error of this sequence is 26·6° and the lower bound is 26·2°. Let us compare this with the sequence 3 4 7, which is deceptively similar to the first example. However, the lattice it generates is shown in Fig. 4 which also gives the lattice dimensions. It is immediately apparent that it corresponds to very poor packing of spheres in two dimensions. Consequently, the r.m.s. error is expected to be much higher than the lower bound. Actual calculation shows $\Delta\varphi_{\rm rms}$ to be 36·9° while $\Delta\varphi_{\rm lb}$ is 28·3°. The reason why the actual error is so much greater than the lower bound is the uneven distribution of lattice points as seen in Fig. 4. Examination of the process which sets up the lattice shows immediately that this effect is caused by two of the integers (3 and 4) summing to a

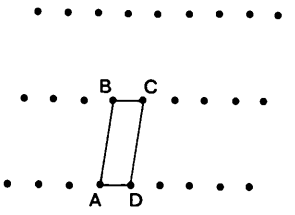


Fig. 4. The two-dimensional lattice produced by the sequence 3 4 7. Dimensions of the lattice in cycles and degrees are: $AB=0{\cdot}581$; $BD=0{\cdot}593$; $DA=0{\cdot}201$; $\angle DAB=83{\cdot}4°$; $\angle ABD=19{\cdot}8°$; $\angle BDA=76{\cdot}9°$.

Table 2. *Magic-integer sequences based on the geometric progression* 1 2 4 8 16 ... *of common ratio* $r=2$

n	Sequence										$\Delta\varphi_{lb}$	$\Delta\varphi_{rms}$	R.m.s. errors of individual phases									
2	2	3									20·4	20·4	24·0	16·0								
3	4	6	7								26·2	26·6	30·4	24·3	24·6							
4	8	12	14	15							29·5	29·9	34·0	28·3	28·7	28·5						
5	16	24	28	30	31						31·8	32·5	36·2	31·9	31·3	31·4	31·6					
6	32	48	56	60	62			63			33·5	34·2	37·7	34·1	33·3	33·2	33·6	32·8				
7	64	96	112	120	124	126		127			34·7	35·5	38·9	35·7	35·0	34·4	34·8	34·9	34·4			
8	128	192	224	240	248	252	255	254			35·7	36·5	40·2	36·7	36·5	36·2	35·6	35·8	36·0	35·1		
9	256	384	448	480	496	504	510	508	511		36·5	37·4	41·0	37·7	37·8	37·1	36·7	36·9	36·8	36·8	35·9	
10	512	768	896	960	992	1008	1020	1016	1022	1023	37·1	38·3	42·2	39·3	37·9	38·2	37·8	37·3	37·7	37·6	38·2	36·4

third integer (7). This simple linear relationship among the integers will always produce such a result. Therefore, it may be stated as a general rule that if the sum or difference of two integers is also a member of the sequence, the r.m.s. error of the sequence will be higher than it could otherwise be. In view of rule 3, the sums and differences of integers in the sequence being unique must also apply when the integers are complemented. This rule leads to a sequence in which the differences between adjacent integers are the terms of a geometric progression. Such a sequence ensures that the sums and differences of pairs of integers are all unique and are not of the sequence.

A combination of these four rules leads to the most efficient magic-integer sequences found so far which are of practical use. A typical sequence has $2m_1=m_n+1$ and the differences $m_n-m_{n-1},m_{n-1}-m_{n-2},\ldots,m_3-m_2,m_2-m_1$ form a geometric progression with a common ratio, $r>1$. Examples of such sequences are shown in Table 2 where they are based on the geometric progression 1 2 4 8 16,...,2^p,.... Also shown in Table 2 are the actual r.m.s. errors and the lower bounds of the sequences. In each case, the actual error is only slightly more than the lower bound, showing that the sequences are very good. The r.m.s. errors were nearly all calculated using a Monte Carlo technique rather than the cumbersome analytical method described in this paper. This means the values of $\Delta\varphi_{rms}$ shown are subject to an error of about $\pm 0{\cdot}3°$. Table 2 also contains the r.m.s. errors of the individual phases represented by each sequence. Because of the Monte Carlo technique, these are not as accurately calculated as the overall r.m.s. errors, but it should be clear from the table that the errors are distributed fairly evenly among the phases. This is the property that makes these sequences particularly efficient [in fulfilment of the criterion (*c*)] and is important for all practical applications.

A sequence of any length can easily be generated in the way shown by the examples in Table 2 and its r.m.s. errors may be estimated quite accurately as slightly more than the lower bound calculated from equation (20). However, long sequences will give rise to very large integers, making them less practical, and it becomes necessary to find other sequences with smaller integers. Reducing the sizes of the integers will result in higher phase errors, but these can be tolerated. It is only possible to produce sequences using smaller integers if the common ratio of the generating geometric progression is made smaller than two. A true geometric progression is now impossible, but a good approximation can be obtained if each term is made a simple linear function of other terms.

The ratio of adjacent members of the series $F_1,F_2,\ldots,F_i$ will converge to a definite limit when the series is produced by:

$$F_n=F_{n-i}+F_{n-j}+\ldots, \tag{23}$$

where $i,j\ldots$ are any integers and the limit will be the

largest positive root of the polynomial

$$r^n = r^{n-i} + r^{n-j} + \ldots . \qquad (24)$$

For example, the Fibonacci series 1 1 2 3 5 8 13 ... is generated when each term is the sum of the previous two. The ratio of adjacent terms rapidly converges to the limit $(1+\sqrt{5})/2$, known in this case as the 'golden ratio', and the series is a very good approximation to a geometric progression. Using this series as the differences between adjacent members of magic-integer sequences formed such that $2m_1 = m_n + 1$ produces the sequences shown in Table 3. A further example is shown in Table 4 where the magic-integer sequences are based on the series 1 1 1 1 2 3 4 5 7 10 14 .... This series is generated by

$$F_n = F_{n-1} + F_{n-4} \qquad (25)$$

and the ratio of adjacent terms converges to a limit of 1·380. The final example in Table 5 is of sequences based on the series 1 1 1 1 ... which may be regarded as a geometric progression of common ratio 1.

In all of these examples, the r.m.s. error is only slightly more than the lower bound. This means that equation (20) can be used to estimate $\Delta\varphi_{\mathrm{rms}}$ for all sequences generated in this way instead of using either the analytical or Monte Carlo techniques. Note that the difference between $\Delta\varphi_{\mathrm{rms}}$ and $\Delta\varphi_{\mathrm{lb}}$ increases as the common ratio of the generating integer-geometric series decreases. This is because the rules for setting up efficient sequences, especially the important fourth rule, are less well obeyed as r decreases from a value of 2. In addition, these sequences all have the important property that the errors in phase representation are distributed evenly among the phases concerned. This is not true of sequences generated in any other way tried so far.

Very long sequences

It is interesting to see how $\Delta\varphi_{\mathrm{lb}}$ for efficient magic-integer sequences varies as a function of the length of the sequence, n, and the common ratio, r, of the generating integer-geometric progression. Equation (20)

Table 3. *Magic-integer sequences based on the Fibonacci series* 1 1 2 3 5 8 13 ..., *generated by* $F_n = F_{n-1} + F_{n-2}$ *and limiting ratio* $r = 1{\cdot}618$

n	Sequence										$\Delta\varphi_{\mathrm{lb}}$	$\Delta\varphi_{\mathrm{rms}}$
2	2	3									20·4	20·4
3	3	4	5								31·2	32·0
4	5	7	8	9							35·2	36·1
5	8	11	13	14	15						38·4	39·2
6	13	18	21	23	24	25					40·5	41·4
7	21	29	34	37	39	40	41				42·1	43·2
8	34	47	55	60	63	65	66	67			43·4	44·4
9	55	76	89	97	102	105	107	108	109		44·4	45·7
10	89	123	144	157	165	170	173	175	176	177	45·2	46·6

Table 4. *Magic-integer sequences based on the series* 1 1 1 1 2 3 4 5 7 10 ..., *generated by* $F_n = F_{n-1} + F_{n-4}$ *and limiting ratio* $r = 1{\cdot}380$

n	Sequence										$\Delta\varphi_{\mathrm{lb}}$	$\Delta\varphi_{\mathrm{rms}}$
2	2	3									20·4	20·4
3	3	4	5								31·2	32·0
4	4	5	6	7							38·6	39·7
5	5	6	7	8	9						44·1	45·6
6	7	9	10	11	12	13					46·5	48·2
7	10	13	15	16	17	18	19				48·2	49·9
8	14	18	21	23	24	25	26	27			49·6	51·6
9	19	24	28	31	33	34	35	36	37		51·1	52·9
10	25	33	38	42	45	47	48	49	50	51	52·2	54·1

Table 5. *Magic-integer sequences based on the series* 1 1 1 1 1 1 ..., *generated by* $F_n = 1$ *and common ratio* $r = 1$

n	Sequence										$\Delta\varphi_{\mathrm{lb}}$	$\Delta\varphi_{\mathrm{rms}}$
2	2	3									20·4	20·4
3	3	4	5								31·2	32·0
4	4	5	6	7							38·6	39·7
5	5	6	7	8	9						44·1	45·6
6	6	7	8	9	10	11					48·4	50·0
7	7	8	9	10	11	12	13				51·7	53·1
8	8	9	10	11	12	13	14	15			54·5	56·2
9	9	10	11	12	13	14	15	16	17		56·9	58·3
10	10	11	12	13	14	15	16	17	18	19	58·8	60·1

can be used to calculate $\Delta\varphi_{lb}$ as a function of n and the results for the four different series in Tables 2 to 5 are shown in Fig. 5. Clearly, $\Delta\varphi_{lb}$ tends to a limit in each case which is dependent upon the value of r. The limits can be obtained by expressing Σm^2 in (22) in terms of the parameters of the geometric progression. To a good approximation, $(\Sigma m^2)^{1/2}$ is linearly dependent upon r^{n-2}, which is the highest power of r appearing in the complete expression. Putting

$$\left(\sum_{i=1}^{n} m_i^2\right)^{1/2} = Ar^{n-2}, \tag{26}$$

where A is a constant, from (22) and (26) we obtain:

$$\lim_{n\to\infty} \Delta\varphi_{lb} = \frac{1}{r}\left(\frac{2\pi}{e}\right)^{1/2} \text{ radians}. \tag{27}$$

It is clear from this and from Fig. 5 that the larger the value of r of the generating g.p. the smaller will be the r.m.s. error of the corresponding magic-integer sequence. This has practical significance and, indeed, leads to a further rule for generating efficient sequences. If more than one sequence can be generated, using the rules outlined so far, of the same length n and using the same integers m_1 and m_n at each end, the sequence corresponding to the larger value of r will have the smaller r.m.s. error.

Note that for $r=1$, equation (27) gives a lower bound of 87·2°. This is a high error, but it certainly does not correspond to random phases (the r.m.s. error of random phases is 104·9°). If the actual error remains close to the lower bound for very long sequences, this means n phases can be represented by a single variable using the consecutive integers from n to $2n-1$ with an error significantly less than that corresponding to random phases. In this case n can be large enough to represent most of the strong reflexions in a complete data set.

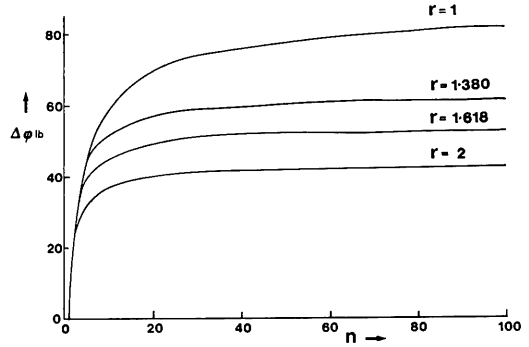


Fig. 5. Plot of lower bound of r.m.s. error in degrees, $\Delta\varphi_{lb}$, against the length of sequence, n, for four different values of r, the limiting ratio of the generating integer-geometric progression.

In conclusion

To summarize the recipe for generating efficient magic integer sequences:

The sequence $m_1, m_2, \ldots, m_n$ is such that $2m_1 = m_n + 1$ and the differences $m_n - m_{n-1}, m_{n-1} - m_{n-2}, \ldots, m_3 - m_2, m_2 - m_1$ form an integer-geometric progression with a common ratio, r, in the range $1 \leq r \leq 2$. The larger the value of r, the more efficient the sequence will be. The integer-geometric progression is generated by relationships of the type shown in equation (23).

No proof can be offered that this recipe always produces the most efficient sequences. Such a proof would relate either to the most efficient covering or the most efficient packing of n-dimensional space with hyperspheres and these are as yet unsolved problems in mathematics. However, the lower-bound formula (20) is quite rigorous and, as the sequences produced have errors which are only slightly greater than the lower bound, there remains very little room for improvement.

The author has not investigated sequences based on geometric progressions for which $r > 2$, since these have not yet found any practical use.

The author wishes to thank Mr D. Taylor for many stimulating discussions. Some of this work would have taken much longer without his keen interest in the problem.

References

DECLERCQ, J. P., GERMAIN, G. & WOOLFSON, M. M. (1975). *Acta Cryst.* A**31**, 367–372.

KENDALL, M. G. (1961). *A Course in the Geometry of n-Dimensions. Griffin's Statistical Monographs and Courses*, Vol. 8. New York: Hafner.

ROGERS, C. A. (1964). *Packing and Covering*. Cambridge Univ. Press.

VORONOI, G. (1908). *J. reine angew. Math.* **134**, 198–287.

WHITE, P. S. & WOOLFSON, M. M. (1975). *Acta Cryst.* A**31**, 53–56.