

## Derivation of a New Tangent Formula from Patterson-Function Arguments

BY JORDI RIUS

*Institut de Ciència de Materials de Barcelona, CSIC, Campus Universitari de Bellaterra,  
08193 Cerdanyola, Barcelona, Spain*

(Received 15 June 1992; accepted 2 October 1992)

### Abstract

Most Patterson-search methods are based on functions measuring the coincidence of the observed Patterson function with the model Patterson function calculated as a function of some variables, e.g. the rotation angles in the case of the rotation function. In parallel with such methods, a new phase-refinement function based on the maximization of  $Z_R(\Phi) = \sum_{\mathbf{H}} (E_{\mathbf{H}} - \langle E_{\mathbf{H}} \rangle) G_{\mathbf{H}}(\Phi)$  as a function of the collectivity  $\Phi$  of phases of the reflections with large  $E$  values is presented, where  $E_{\mathbf{H}} - \langle E_{\mathbf{H}} \rangle$  and  $G_{\mathbf{H}}(\Phi)$  are the Fourier coefficients of two Patterson-type functions, the first with the origin peak removed and the second taking into account the atomicity condition. It will be shown how  $Z_R(\Phi)$  can be maximized by a new tangent formula actively using the small  $E$  values and not possessing any weighting factor. The utility of this formula is illustrated on the basis of two representative test examples.

### 1. Introduction

Direct interpretation of the Patterson (1935) function constitutes a simple way of solving crystal structures containing heavy atoms. However, the fact that many substances, such as organic compounds, do not contain heavy atoms seriously restricts its application.

In the case where a related molecular fragment is available, the so-called Patterson-search methods can be applied. These methods use the known geometry of the fragment to compensate for the absence of strong individual Patterson peaks. One such method is the rotation function of Rossman & Blow (1962). Its most widespread use is for determining the relative orientation of a given search fragment in the unit cell of the unknown structure and can be expressed, for this particular case, in the form

$$R(\Omega) = V \int_{\nu} P_0(\mathbf{u}) P_{\text{model}}(\mathbf{u}, \Omega) d\mathbf{u}. \quad (1)$$

$R(\Omega)$  measures the coincidence of the observed Patterson function ( $P_0$ ) with the Patterson function of the fragment or model ( $P_{\text{model}}$ ) computed as a function of the rotational operator  $\Omega$ , i.e. the three rotation angles. For the correct rotations, the values of  $R(\Omega)$  will be, in general, maximal.

Should no suitable molecular fragment be available, the Patterson function can be interpreted using direct methods, provided that the intensity data extend to atomic resolution. This implies that the atomic peaks in the true and squared structures are discrete, so that the respective structure factors  $E_{\mathbf{H}} = E_{\mathbf{H}} \exp i\varphi_{\mathbf{H}}$  and  $G_{\mathbf{H}} = G_{\mathbf{H}} \exp i\phi_{\mathbf{H}}$  will have similar associated phases  $\varphi_{\mathbf{H}} \approx \phi_{\mathbf{H}}$ . This allows the modulus  $G_{\mathbf{H}}$  to be expressed in terms of the collectivity  $\Phi$  of phases of the set  $\{K\}$  of reflections with large  $E$  magnitudes,

$$\begin{aligned} G_{\mathbf{H}}(\Phi) &= \exp i\phi_{-\mathbf{H}} G_{\mathbf{H}} \\ &= \text{Re} \left( \exp i\phi_{-\mathbf{H}} \sum_{\mathbf{h}'} E_{\mathbf{h}'} E_{\mathbf{H}-\mathbf{h}'} \right) \\ &= \sum_{\mathbf{h}'} E_{\mathbf{h}'} E_{\mathbf{H}-\mathbf{h}'} \cos(\varphi_{-\mathbf{H}} + \varphi_{\mathbf{h}'} + \varphi_{\mathbf{H}-\mathbf{h}'}) \end{aligned} \quad (2)$$

with  $\mathbf{h}'$  and  $\mathbf{H}-\mathbf{h}' \in K$ . Now let  $P_0(\mathbf{u})$  and  $\mathcal{P}(\mathbf{u}, \Phi)$  be the Patterson-type Fourier syntheses (Ramachandran & Raman, 1959)

$$P_0(\mathbf{u}) = (1/V) \sum_{\mathbf{H}} E_{\mathbf{H}} \exp(-2\pi i \mathbf{H} \cdot \mathbf{u}), \quad (3)$$

$$\mathcal{P}(\mathbf{u}, \Phi) = (1/V) \sum_{\mathbf{H}} G_{\mathbf{H}}(\Phi) \exp(-2\pi i \mathbf{H} \cdot \mathbf{u}). \quad (4)$$

By analogy with the rotation function, the 'coincidence' function can be defined as

$$\begin{aligned} Z(\Phi) &= V \int_{\nu} P_0(\mathbf{u}) \mathcal{P}(\mathbf{u}, \Phi) d\mathbf{u} \\ &= \sum_{\mathbf{H}} E_{\mathbf{H}} G_{\mathbf{H}}(\Phi). \end{aligned} \quad (5)$$

Consideration in (5) of the reflections with large  $E$  values only, followed by substitution of (2) in (5), leads to the well known expression

$$Z(\Phi) = \sum_{\mathbf{h}} \sum_{\mathbf{h}'} E_{-\mathbf{h}} E_{\mathbf{h}'} E_{\mathbf{H}-\mathbf{h}'} \cos(\varphi_{-\mathbf{h}} + \varphi_{\mathbf{h}'} + \varphi_{\mathbf{H}-\mathbf{h}'}). \quad (6)$$

Fig. 1(a) shows the observed  $P_0(\mathbf{u})$  function and the  $\mathcal{P}(\mathbf{u}, \Phi)$  function, calculated with true phases, of a one-dimensional equal-atom model structure. Fig. 1(b) shows the corresponding product function  $P_0(\mathbf{u}) \times \mathcal{P}(\mathbf{u}, \Phi)$ . It can be seen that the largest contribution to  $Z(\Phi)$  is determined by the origin peak 0 of  $P_0$  and that the contribution due to the remaining

part,  $R$ , of  $P_0$  is comparatively insignificant. If the  $E_{\mathbf{H}} - \langle E_{\mathbf{H}} \rangle$  are the Fourier coefficients of the  $P_0$  function with the origin peak removed, then these two contributions can be written separately in the form

$$\begin{aligned} Z(\Phi) &= Z_0(\Phi) + Z_R(\Phi) \\ &= \langle E_{\mathbf{H}} \rangle \sum_{\mathbf{H}} G_{\mathbf{H}}(\Phi) + \sum_{\mathbf{H}} (E_{\mathbf{H}} - \langle E_{\mathbf{H}} \rangle) G_{\mathbf{H}}(\Phi) \\ &\approx \langle E_{\mathbf{H}} \rangle \sum_{\mathbf{H}} G_{\mathbf{H}}(\Phi). \end{aligned} \quad (7)$$

Consequently, any solution of  $\Phi$  giving rise to a large origin peak in  $\mathcal{P}$  will maximize  $Z(\Phi)$ .

In comparison with the three variables of the rotation function, the large number of phases increases the difficulty of finding the correct maximum of  $Z(\Phi)$ . As shown by Debaerdemaeker, Tate & Woolfson (1985),  $Z(\Phi)$  can be maximized by means of the tangent formula of Karle & Hauptman (1956). In practice, however, refinement of initially random phases with this tangent formula tends to refine them to overconsistency, *i.e.* the  $\mathcal{P}(\mathbf{u}, \Phi)$  function com-

puted with these phases shows an origin peak larger than expected while the remaining part of  $\mathcal{P}(\mathbf{u}, \Phi)$  is, in general, incorrect. This problem is especially important in structures with symmorphic space groups. As demonstrated below, a simple solution to this problem is the maximization of  $Z_R(\Phi)$  instead of  $Z(\Phi)$ . Since the origin peak of  $P_0$  has been removed (Fig. 2a), the structure is now solved by combination of the information of the nonorigin peaks of  $P_0(\mathbf{u})$  with the structural restriction implicit in (2), *i.e.* the atomicity condition. Fig. 2(b) shows the corresponding product function.

## 2. The 'coincidence' function $Z_R(\Phi)$

The function  $Z_R(\Phi)$  measuring the coincidence of the  $P_0$  function without the origin peak with the  $\mathcal{P}(\mathbf{u}, \Phi)$  function is given by

$$Z_R(\Phi) = \sum_{\mathbf{H}} (E_{\mathbf{H}} - \langle E_{\mathbf{H}} \rangle) G_{\mathbf{H}}(\Phi). \quad (8)$$

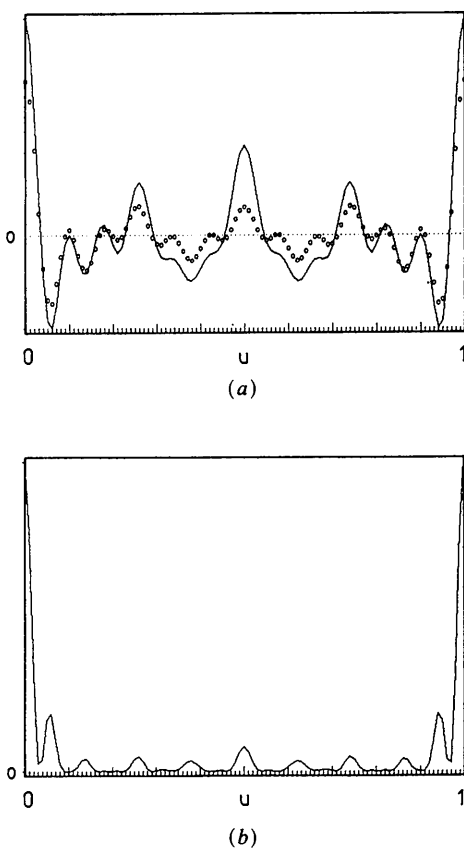


Fig. 1. (a) Patterson-type function  $P_0$  (small circles) of a one-dimensional model structure with atoms at 0.00, 0.20, 0.50 and 0.75, and Patterson-type function  $\mathcal{P}(\Phi)$  (continuous line) calculated with true phases ( $E_{\min} = 0.71$ ). Both functions without zero Fourier term,  $[(\sin \theta) / \lambda]_{\max} = 0.6 \text{ \AA}^{-1}$ . (b) Product function  $P_0 \times \mathcal{P}(\Phi)$ . Notice the dominance of the origin peak over the rest.

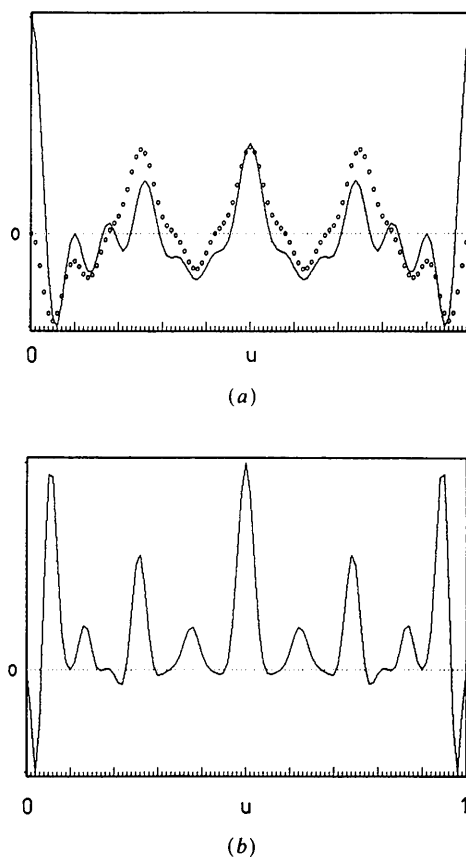


Fig. 2. (a)  $P_0$  function with removed origin peak (small circles) and  $\mathcal{P}(\Phi)$  function calculated with true phases (continuous line) for the same example as in Fig. 1. (b) Product function  $P_0$  (without origin peak)  $\times \mathcal{P}(\Phi)$ . The disappearance of the origin peak gives greater significance to the remaining peaks compared with Fig. 1(b).

After replacement of  $G_{\mathbf{H}}(\Phi)$  by (2),  $Z_R(\Phi)$  becomes

$$\begin{aligned} & \sum_{\mathbf{h}} (1 - \langle E_{\mathbf{H}} \rangle / E_{\mathbf{h}}) \sum_{\mathbf{h}'} E_{-\mathbf{h}} E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'} \\ & \times \cos(\varphi_{-\mathbf{h}} + \varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'}) + \sum_{\mathbf{l}} (E_{\mathbf{l}} - \langle E_{\mathbf{H}} \rangle) \\ & \times \sum_{\mathbf{h}} E_{\mathbf{h}} E_{\mathbf{l}-\mathbf{h}} \cos(\phi_{-\mathbf{l}} + \varphi_{\mathbf{h}} + \varphi_{\mathbf{l}-\mathbf{h}}) \end{aligned}$$

and, after some manipulation, it takes the final form

$$\begin{aligned} Z_R(\Phi) = & \sum_{\mathbf{h}} E_{-\mathbf{h}} \left[ \sum_{\mathbf{h}'} X_{\mathbf{h},\mathbf{h}'} E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'} \cos(\varphi_{-\mathbf{h}} + \varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'}) \right. \\ & \left. + \sum_{\mathbf{l}} (E_{\mathbf{l}} - \langle E_{\mathbf{H}} \rangle) E_{\mathbf{l}-\mathbf{h}} \cos(\varphi_{-\mathbf{h}} + \phi_{\mathbf{l}} + \varphi_{\mathbf{l}-\mathbf{h}}) \right] \end{aligned} \quad (9)$$

with

$$\begin{aligned} X_{\mathbf{h},\mathbf{h}'} = & [(1 - \langle E_{\mathbf{H}} \rangle / E_{-\mathbf{h}}) + (1 - \langle E_{\mathbf{H}} \rangle / E_{\mathbf{h}'}) \\ & + (1 - \langle E_{\mathbf{H}} \rangle / E_{\mathbf{h}-\mathbf{h}'})] / 3 \end{aligned} \quad (10)$$

and  $\mathbf{l}$  being the reflections not belonging to the basis set  $\{K\}$ . By use of the same procedure as Debaeremaeker *et al.* (1985), the new phase estimates maximizing  $Z_R(\Phi)$  can be found if the condition is solved for the limit of  $Z_R(\Phi)$ ,

$$(\partial / \partial \varphi_{\mathbf{h}}) Z_R(\Phi) = 0 \quad \forall \mathbf{h} \in K, \quad (11)$$

which leads to the tangent formula

$$\begin{aligned} \varphi_{\mathbf{h}} = & \text{phase of} \left\{ \sum_{\mathbf{h}'} X_{\mathbf{h},\mathbf{h}'} E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'} \right. \\ & \left. + \sum_{\mathbf{l}} [E_{\mathbf{l}} - \langle E_{\mathbf{H}} \rangle] E_{\mathbf{l}-\mathbf{h}} \exp i\phi_{\mathbf{l}} \right\}, \end{aligned} \quad (12)$$

where the phases  $\phi_{\mathbf{l}}$  are periodically recalculated from  $G_{\mathbf{l}}(\Phi)$ . Besides the largest  $E$  values, (12) also requires the small  $E$  values since the corresponding Fourier coefficients  $E_{\mathbf{l}} - \langle E_{\mathbf{H}} \rangle$  are large negative quantities.

Finally, it is interesting to see what happens when the same theory is applied to the Patterson function. Since the corresponding Fourier coefficients are the squared moduli of the structure factors, (8) becomes

$$Z'_R(\Phi) = \sum_{\mathbf{H}} (E_{\mathbf{H}}^2 - \langle E_{\mathbf{H}}^2 \rangle) G_{\mathbf{H}} G_{\mathbf{H}}(\Phi), \quad (13)$$

where  $\langle E_{\mathbf{H}}^2 \rangle$  is the mean value of  $E_{\mathbf{H}}^2$  averaged over all  $\mathbf{H}$ . For the reflections belonging to  $\{K\}$ ,  $G_{\mathbf{h}}$  can be estimated using the relationship

$$G_{\mathbf{h}} = s_{\mathbf{h}} E_{\mathbf{h}}, \quad (14)$$

where  $s_{\mathbf{h}}$  is a calculable scaling factor. For the remaining reflections, the  $G_{\mathbf{l}}$  are derived from the available  $\varphi_{\mathbf{h}}$  estimates,

$$G_{\mathbf{l}} = \left| \sum_{\mathbf{h}''} E_{\mathbf{h}''} E_{\mathbf{l}-\mathbf{h}''} \right| \quad (15)$$

with  $\mathbf{h}''$  and  $\mathbf{l}-\mathbf{h}'' \in K$ , so that (13) can be rewritten

Table 1. Number of correct solutions (%) of the test structure MBH2

40 sets of random phases refined with tangent formula (12) under different conditions specified by  $N_{\text{large}}$  (=number of large  $E$  values) and by  $N_{\text{small}}$  (=number of small  $E$  values). The best result is obtained for  $N_{\text{large}} = 220$  and  $N_{\text{small}} = 200$  with  $\langle E_{\mathbf{H}} \rangle = 1.18$ . A solution is considered to be correct if  $|\Delta\varphi_{\mathbf{h}}| < 35^\circ$ .

$N_{\text{large}}$	$N_{\text{weak}}$					
	140	160	180	200	220	240
240	0.0	12.5	12.5	12.5	10.0	—
220	7.5	25.0	20.0	32.5	30.0	15.0
200	0.0	5.0	17.5	7.5	7.5	—
180	0.0	0.0	0.0	0.0	—	—

in the form

$$\begin{aligned} Z'_R(\Phi) = & \sum_{\mathbf{h}} s_{\mathbf{h}} (E_{\mathbf{h}}^2 - \langle E_{\mathbf{H}}^2 \rangle) E_{\mathbf{h}} G_{\mathbf{h}}(\Phi) \\ & + \sum_{\mathbf{l}} (E_{\mathbf{l}}^2 - \langle E_{\mathbf{H}}^2 \rangle) \left| \sum_{\mathbf{h}''} E_{\mathbf{h}''} E_{\mathbf{l}-\mathbf{h}''} \right| G_{\mathbf{l}}(\Phi). \end{aligned} \quad (16)$$

It can be demonstrated\* that  $Z'_R(\Phi)$  can be maximized with a tangent formula similar to (12), *i.e.*

$$\begin{aligned} \varphi_{\mathbf{h}} = & \text{phase of} \left\{ \sum_{\mathbf{h}'} X_{\mathbf{h},\mathbf{h}'} E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'} + \sum_{\mathbf{l}} (E_{\mathbf{l}}^2 - \langle E_{\mathbf{H}}^2 \rangle) E_{\mathbf{l}-\mathbf{h}} \right. \\ & \left. \times \left| \sum_{\mathbf{h}''} E_{\mathbf{h}''} E_{\mathbf{l}-\mathbf{h}''} \right| \exp i\phi_{\mathbf{l}} \right\}, \end{aligned} \quad (17)$$

where the principal differences are the presence of the factor  $|\sum_{\mathbf{h}''} E_{\mathbf{h}''} E_{\mathbf{l}-\mathbf{h}''}|$  in the  $\mathbf{l}$  summation and the redefinition of  $X_{\mathbf{h},\mathbf{h}'}$  according to the expression

$$\begin{aligned} X_{\mathbf{h},\mathbf{h}'} = & [s_{-\mathbf{h}} (E_{-\mathbf{h}}^2 - \langle E_{\mathbf{H}}^2 \rangle) + s_{\mathbf{h}'} (E_{\mathbf{h}'}^2 - \langle E_{\mathbf{H}}^2 \rangle) \\ & + s_{\mathbf{h}-\mathbf{h}'} (E_{\mathbf{h}-\mathbf{h}'}^2 - \langle E_{\mathbf{H}}^2 \rangle)] / 3. \end{aligned} \quad (18)$$

Although the derivation of (17) is quite simple, it requires further experimental confirmation, which is not provided in the present paper.

### 3. Test calculations

To test the tangent formula (12), two structures from the Sheldrick data bank (Göttingen) have been selected: the structure with code name MBH2, of formula  $C_{15}H_{24}O_3$ , triclinic, space group  $P1$ ,  $Z=3$ , cell volume =  $1106 \text{ \AA}^3$  (Poyser, Edwards, Anderson, Hursthouse, Walker, Sheldrick & Walley, 1986); and the triclinic modification of valinomycin (code name = TVAL), of formula  $C_{54}H_{90}N_6O_{18}$ , triclinic, space group  $P1$ ,  $Z=2$  (Karle, 1975; Smith, Duax, Langs, DeTitta, Edmonds, Rohrer & Weeks, 1975).

The results of the application of (12) to MBH2 are summarized in Table 1. It can be seen that the best

\* The mathematical derivation has been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 55600 (3 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

result is obtained for  $N_{\text{large}} = 220$  and  $N_{\text{small}} = 200$  with 32.5% of the refined sets possessing mean phase errors less than  $35^\circ$ : 7.5, 10.0 and 15.0% in the ranges 0–25, 25–30 and 30–35°, respectively.

The TVAL structure was selected to show the capability of the tangent formula (12) to refine random phases in the case of a relatively large structure (156 nonhydrogen atoms in the unit cell). The best result was obtained for  $N_{\text{large}} = 300$  and  $N_{\text{small}} = 300$  ( $\langle E_{\text{H}} \rangle = 1.24$ ). From a total of 200 sets, 12 (6%) show a mean phase error between 30 and 35°.

#### 4. Concluding remarks

The viability of solving crystal structures from the direct interpretation of the nonorigin Patterson peaks as a function of the phases has been demonstrated. This result, however, should not be surprising since these peaks contain all the information regarding the atomic arrangement in the structure and, in addition,

the atomicity information contained in the removed Patterson origin peak has already been considered in the derivation of the Fourier coefficients  $G_{\text{H}}(\Phi)$ .

This work was supported by the DGICYT (Project PB89-0036).

#### References

- DEBAERDEMAEKER, T., TATE, C. & WOOLFSON, M. (1985). *Acta Cryst.* **A41**, 286–290.  
 KARLE, I. (1975). *J. Am. Chem. Soc.* **97**, 4379–4386.  
 KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.  
 PATTERSON, A. (1935). *Z. Kristallogr. Teil A*, **90**, 517–542.  
 POYSER, J., EDWARDS, R., ANDERSON, J., HURSTHOUSE, M., WALKER, N., SHELDRIK, G. & WALLEY, A. (1986). *J. Antibiot.* **39**, 167–169.  
 RAMACHANDRAN, G. N. & RAMAN, S. (1959). *Acta Cryst.* **12**, 957–964.  
 ROSSMANN, M. & BLOW, D. (1962). *Acta Cryst.* **15**, 24–31.  
 SMITH, G., DUAX, W., LANGS, D., DETITTA, G., EDMONDS, J., ROHRER, D. & WEEKS, C. (1975). *J. Am. Chem. Soc.* **97**, 7242–7247.

*Acta Cryst.* (1993). **A49**, 409–421

## Parametrization of Triply Periodic Minimal Surfaces. III. General Algorithm and Specific Examples for the Irregular Class

BY A. FOGDEN

*Department of Applied Mathematics, Research School of Physical Sciences and Engineering, Australian National University, Box 4, Canberra 2601, Australia*

(Received 11 July 1991; accepted 5 October 1992)

#### Abstract

The construction algorithm of the first two papers of this series [Fogden & Hyde (1992). *Acta Cryst.* **A48**, 442–451, 575–591] is extended to a general treatment of triply periodic minimal surfaces (containing as a special case the ‘regular’ class analysed previously). A detailed outline of the parametrization procedure for an arbitrary ‘irregular’ class surface is provided and verified *via* a systematic rederivation of the C(P) surface described by Neovius [*Bestimmung Zweier Speziellen Periodische Minimalflächen* (1883), Helsinki: Frenckel]. The method is further employed in parametrizing various empirically generated surfaces.

#### 1. Introduction

Paper I of this series (Fogden & Hyde, 1992*a*) outlines the mathematical foundation of this study, in extending the local representation of a general minimal surface, due to Weierstrass, to a rigorous connection between infinite (triply) periodic minimal surfaces

(IPMS) and the finite-sheeted Riemann surface of their algebraic complex Weierstrass functions. Topological considerations impose simple conditions relating the fundamental global characteristic of the IPMS – the genus – to the two principal global features of the Riemann surface – the total branch point order and number of sheets. Considerations of differential geometry local to the degenerate points of the IPMS (the ‘flat’ points, at which the Gaussian curvature is zero) constrain the Riemann surface structure to be local to the corresponding branch points. The local and global aspects are then coupled by the symmetries of the IPMS – the plane lines of curvature, linear asymptotes and rotational invariances – which reduce to Weierstrass functional relations. In summary, the specifying properties of an IPMS are readily translated into those of the Riemann surface, making the latter a natural and extremely useful means of describing the former.

The remainder of papers I and II of this series (Fogden & Hyde, 1992*a,b*) deals with a special subset