BUSING, W. R. & LEVY, H. A. (1957). Acta Cryst. 10, 180–182.

BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). ORFLS. Report ORNL-TM-305. Oak Ridge National Laboratory, Tennessee.

CARPER, E. (1971). Wash. Post, March 31, p. A3.

CARROLL, R. E. (1966). J. Amer. Med. Assoc. 198, 267–269.

CROMER, D. T. (1965). Acta Cryst. 18, 17–23.

CROMER, D. T. & WABER, J. T. (1965). Acta Cryst. 18, 104–109.

DIETRICH, H. & CRAMER, F. D. (1954). Chem. Ber. 6, 806–817.

ENGEL, G. (1968). Z. anorg. allgem. Chem. 362, 273–280.

ENGEL, G. (1971). Private communication.

FUKAMACHI, T. (1971). Mean X-ray Scattering Factors Calculated from Analytical Roothaan–Hartree–Fock Wave Functions by Clementi, Tech. Rep. ISSP Ser. B, No. 12. Univ. of Tokyo, Japan.

IVANOV, JU. A., SIMONOV, M. A. & BELOV, N. V. (1976). Zh. Strukt. Khim. 17(2), 375–378.

LARSON, A. C., ROOF, R. B. JR & CROMER, D. T. (1964). Acta Cryst. 17, 1382–1386.

MACKIE, P. E. (1972). Precision Single-Crystal X-ray Diffractometry, PhD Thesis, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

SCHROEDER, H. A. (1965). J. Chronic Diseases, 18, 647–656.

SCHWEIZER, J., STRNAT, K. J. & TSUI, J. B. Y. (1971). Program and Abstracts, Amer. Cryst. Assoc. Summer Meeting, Ames, Iowa, p. 92 (M2).

SHANNON, R. D. (1976). Acta Cryst. A32, 751–767.

SHANNON, R. D. & PREWITT, C. T. (1969). Acta Cryst. B25, 925–946.

SUDARSANAN, K. & YOUNG, R. A. (1974). Acta Cryst. B30, 1381–1386.

SUDARSANAN, K., YOUNG, R. A. & DONNAY, J. D. H. (1973). Acta Cryst. B29, 808–814.

TOKONAMI, N. (1965). Acta Cryst. 19, 486.

TSUCHIYA, K. (1969). Keio J. Med. 18, 181–195.

WILSON, A. J. C., SUDARSANAN, K. & YOUNG, R. A. (1977). B33, 3142–3154.

The Structures of Some Cadmium 'Apatites' $Cd_5(MO_4)_3X$.
II. The Distributions of the Halogen Atoms in $Cd_5(VO_4)_3I$, $Cd_5(PO_4)_3Br$, $Cd_5(AsO_4)_3Br$, $Cd_5(VO_4)_3Br$ and $Cd_5(PO_4)_3Cl$

BY A. J. C. WILSON

The University of Birmingham, Birmingham, B15 2TT, England

AND K. SUDARSANAN AND R. A. YOUNG

Georgia Institute of Technology, Atlanta, Georgia 30332, USA

Many cadmium 'apatites' are 10 to 20% halogen deficient, with the halogen atoms distributed over a range of positions on the hexad axis instead of occupying the ideal positions $\pm(0,0,\frac{1}{4})$. Direct determination of the probability distribution of their positions has been attempted in three ways: (i) Fourier atomic-center density syntheses, (ii) least-squares refinement of the occupancy factors of a multiplicity of sites uniformly spaced along the c axis, and (iii) fitting of postulated distribution functions to the 'observed' structure factors corresponding to halogen centers only. All three methods are limited similarly by the resolving power afforded by the reflections in the accessible volume of reciprocal space. The results from all three methods are consistent with a 'chain' model in which the center of each chain occupies the modal position $(0,0,\frac{1}{4})$ or $(0,0,\frac{3}{4})$. Being 'oversize' (diameter $>\frac{1}{2}c$), halogen ions in succeeding half-cells are forced to deviate further and further from this modal position, until finally the steric strain is relieved by the occurrence of a vacancy, after which a new chain begins. Qualitative agreement is achieved with a uniform distribution of atomic centers over a range, δ, on each side of the modal positions, with δ varying from 0·04 in the chloride to 0·2 in the iodide. No other simple distribution tried, such as Cauchy or Gaussian, was applicable to all compounds.

1. Introduction

The structures of five cadmium apatites have been described in earlier papers (Sudarsanan, Young & Donnay, 1973; Sudarsanan, Young & Wilson, 1977).

Normalized to full occupancy of the oxygen sites in a structure of the fluorapatite type, they are halogen and cadmium deficient, with formulae $Cd_{1-x}(MO_4)_3X_{1-2x}$, and no definite positions on the hexad axis can be ascribed to the halogen atoms. These appear to be

distributed over a range of sites centered on the ideal positions $\pm(0,0,\frac{1}{4})$. Sudarsanan, Wilson & Young (1972) have proposed the following model: The halogen atoms are too big to pack stoichiometrically along the sixfold screw axis, the ionic diameter exceeding $\frac{1}{2}c$ by a fraction α of c. If one atom is at the ideal position $(0,0,\frac{1}{4})$, the next cannot be at $(0,0,\frac{3}{4})$, but must be at $(0,0,\frac{3}{4} + \alpha)$; the next cannot be at $(0,0,\frac{1}{4})$, but at $(0,0,\frac{1}{4} + 2\alpha)$; the next at $(0,0,\frac{3}{4} + 3\alpha)$; and so on. Eventually, when $n\alpha$ reaches some limit $\delta < \frac{1}{4}$, the steric strain is relieved by the omission of an atom, and a new sequence begins. The atoms thus occur in chains of length $2(\frac{1}{2} + \alpha)n_{max}c = \delta(1 + 2\alpha)c/\alpha$, with one atom missing for every $2n_{max} + 1$ present, and the probability of finding an atomic center at $(0,0,z)$ is more or less uniform over the ranges $\frac{1}{4} - \delta < z < \frac{1}{4} + \delta$ and $\frac{3}{4} - \delta < z < \frac{3}{4} + \delta$. The fractional deficiency, $2x$, from the stoichiometric halogen content is thus $1/(2\delta/\alpha + 2)$. Refinement of the model leads to values of $2x$ and δ, so that α and X–X distances can be determined.

If the arrangement described were perfectly regular, it would result in the formation of a superstructure with a long c axis, but a continuous scan along $d^*(00l)$ of the iodide showed no additional reflections in the range from $l < 1$ to $l \sim 8$. However, the above model is more precise than is necessary; instead of symmetrical chains containing an odd number of atoms centered on $(0,0,\frac{1}{4})$ or $(0,0,\frac{3}{4})$ one could have symmetrical chains with an even number of atoms centered on $(0,0,\frac{1}{4} \pm \frac{1}{2}\alpha)$, $(0,0,\frac{3}{4} \pm \frac{1}{2}\alpha)$. Equally, the chains could be unsymmetrical about the atoms nearest the ideal position. These further possibilities reduce the precision of the halogen siting, make a continuous distribution of sites over a range of z more plausible, and explain the absence of an observable superstructure. Possible approaches to the determination of this distribution are discussed in §2, and are applied to the experimental results for the various cadmium 'apatites' in §4.

2. Determination of the distribution of halogen sites: theoretical

Suppose that the fraction of halogen atoms having their centers between $(0,0,z)$ and $(0,0,z + dz)$ is $f(z)dz$, and that the parameters of the nonhalogen atoms are independent of the presence or absence of a halogen atom at z. The structure factor of the hkl reflection should then be given by

$$F(hkl) = F_1(hkl) + X(hkl)\int_0^1 f(z) \exp(2\pi ilz)\,dz, \quad (1)$$

where $X(hkl)$ is the scattering factor for the X ion at the Bragg angle in question (thermal-motion effects included) and $F_1(hkl)$ is the structure factor of the rest

of the unit cell. Because of the halogen deficiency,

$$\int_0^1 f(z)\,dz = 2 - 4x \quad (2)$$

instead of the ideal value 2. The quantity

$$\Delta(hkl) = \frac{F(hkl) - F_1(hkl)}{X(hkl)} \quad (3)$$

should thus be independent of h and k for a given l, since, by equation (1), it is given by

$$\Delta(hkl) = \int_0^1 f(z) \exp(2\pi ilz)\,dz. \quad (4)$$

This consequence of the model is testable, since

$$\Delta_{obs}(hkl) = \frac{F_{obs}(hkl) - F_{1\,calc}(hkl)}{X(hkl)} \quad (5)$$

is obtainable for each observed reflection, and should be constant (for fixed l) within the estimated error of observation and calculation. The observable information about the distribution of X atoms thus reduces to a small number of values $\Delta(**l)$ obtained by averaging the quantities given by equation (5) over h and k for fixed l. For Mo $K\alpha$ radiation the highest observed value of l is about 15 for these substances, but because of the space group $\Delta(**l) = \Delta(**l)$ for l even and $\Delta(**l) = 0$ for l odd, so that there are only six or seven distinguishable numerical data from which to determine $f(z)$. Three approaches suggest themselves: (i) A Fourier synthesis with the $\Delta(**l)$ as coefficients. (ii) Least-squares refinement of the occupancy factors of a sequence of X sites uniformly spaced along the z axis. (iii) Postulation of a plausible form for $f(z)$ and refinement of its adjustable parameters. There are two variants of the first approach: a one-dimensional Fourier synthesis with the $\Delta(**l)$'s only, or a line section through a three-dimensional synthesis based on all the $\Delta(hkl)$'s. In the second approach there is no fundamental necessity for the X sites to be uniformly spaced, but it is convenient both theoretically and practically, and uniform spacing was in fact used in part I (Sudarsanan, Young & Wilson, 1977). From the mathematical point of view it is simply a particular case of the third approach, fitting a postulated model, but it is not a model based on any physical picture.

2.1. Fourier synthesis; line section vs line synthesis

From equation (4),

$$\Delta(l) = \int_0^1 f(\zeta) \exp(2\pi il\zeta)\,d\zeta, \quad (6)$$

where the redundant indices h and k have been dropped

and z has been replaced by ζ. Multiplying each side by $\exp(-2\pi ilz)$ and adding gives

$$A(z) \equiv \sum_{l=-L}^{L} A(l)\exp(-2\pi ilz). \qquad (7)$$

On substituting for $A(l)$ from equation (6) and summing the geometric series, this becomes

$$A(z) = \int_0^1 f(\zeta)\frac{\sin[\pi(2L+1)(\zeta-z)]}{\sin[\pi(\zeta-z)]}d\zeta, \qquad (8)$$

where L is the largest value of l for which data are available. The left-hand side is simply a (truncated) one-dimensional Fourier synthesis with $A(l)$ as coefficients. The right-hand side is a convolution of the desired function $f(z)$ with an undesired smearing function; even if $f(z)$ were actually a delta function the syntheses would give peaks with a half-width of $\sim 1/2L$, or in the present case ~ 0.035. Any detail of $f(z)$ on a smaller scale will not be resolved, and the smearing implicit in the convolution (7) cannot be deconvoluted by the usual, or indeed any, methods (Jones & Misell, 1970). The non-observed Fourier coefficients are not available, and, as far as observation goes, could be assigned arbitrarily, giving arbitrarily different functions $f(z)$.

Rather than attempting to estimate the best values of $A(**l)$, it might be thought preferable to perform a three-dimensional synthesis

$$\sum_{h,k,l} A(hkl)\exp[-2\pi i(hx+ky+lz)]. \qquad (9)$$

This should give appreciable densities only near the line $0,0,z$; the distribution is $f(z)$ smeared with the transform of the limiting sphere, which acts as a three-dimensional window in reciprocal space. This approach has the advantage of giving an experimental determination of the smearing function; the variation perpendicular to $0,0,z$ should result only from lack of resolution, whereas the variation along $0,0,z$ is the convolution of the desired distribution function and the smearing function. The line section obtained by putting $x=y=0$ in equation (9),

$$B(z) \equiv \sum_{h,k,l} A(hkl)\exp(-2\pi ilz), \qquad (10)$$

is not identical with the line synthesis $A(z)$ given by equation (7). It may be written

$$B(z) = \sum_l D(l)\exp(-2\pi ilz), \qquad (11)$$

where

$$D(l) = \sum_{h,k} A(hkl). \qquad (12)$$

It seems fairly obvious that $D(l)/D(0)$ will decrease more rapidly with increasing l than $A(l)/A(0)$ does. In the ideal case postulated above, $D(l)$ will be $A(l)$ multiplied by the number of reflections in the lth layer line. This is approximately the area of the small circle

defined by the intersection of the limiting sphere and the lth layer of the reciprocal lattice divided by C^*, the area of the appropriate face of the reciprocal unit cell, and is thus approximately

$$K[(L+\varepsilon)^2 - l^2], \qquad (13)$$

where K and ε are constants depending on the dimensions of the reciprocal unit cell and the radius of the limiting sphere. Equation (12) becomes

$$D(l) = K[(L+\varepsilon)^2 - l^2]A(l). \qquad (14)$$

With the aid of equations (6) and (14) we can obtain the equivalent of equation (8) for the line section. Substitution in equation (11) gives

$$B(z) = K\int_0^1 f(\zeta)\left\{\sum_{l=-L}^{L}[(L+\varepsilon)^2-l^2]\right.$$
$$\left. \times \exp[2\pi il(\zeta-z)]\right\}d\zeta. \qquad (15)$$

The summation is rather tedious; it leads to

$$B(z) = K\int_0^1 f(\zeta)H(\zeta-z)\,dz, \qquad (16)$$

where

$$H(x) = (\{4L(2\varepsilon-1)+4\varepsilon^2+[4L(1-2\varepsilon)+2$$
$$-4\varepsilon^2]\cos^2\pi x\}\sin\pi(2L+1)x - 2(2L+1)$$
$$\times \sin\pi x\cos\pi x\cos\pi(2L+1)x)/4\sin^3\pi x. \qquad (17)$$

This is the Fourier-series analogue of the spherical Bessel function of Fourier-transform theory. The coefficient of $\sin\pi(2L+1)x$ depends somewhat on the value of ε, being simplest for $\varepsilon = \frac{1}{2}$, but effectively the true function $f(z)$ is convoluted with a $(\sin x - x\cos x)/x^3$ function in forming the line section, instead of with $(\sin x)/x$ as in forming the one-dimensional synthesis. The smearing is about 30% greater for the line section; this is to be expected, since $B(z)$ is effectively $A(z)$ calculated with an additional 'artificial temperature factor' of roughly $1 - l^2/L^2$. The one-dimensional synthesis [equation (7)] is thus to be preferred, on two grounds. (i) Its theoretical resolving power is better. (ii) Because of the techniques of data collection (especially the omission of reflections that might have been affected by simultaneous diffraction at the particular crystal setting used) the number of reflections actually available (N in Table 1) does not decrease smoothly with increasing l in accordance with the expression (13). The resulting distortion of $B(z)$ is difficult either to predict or to allow for.

2.2. Least-squares refinement of occupancy factors

The second approach is to consider the halogen atoms as occupying equally spaced sites along the

hexad axis, and to refine the occupancy factors of these sites by some least-squares process. If there are $2N + 1$ sites with spacing γ, and the site-occupancy factor of the jth site is f_j, the calculated structure factor for the halogens, for comparison with the 'observed' corrected and inflated factor $\Delta(l)$, will be

$$G(l) = \sum_{j=-N}^{N} f_j \exp(2\pi i j l \gamma). \qquad (18)$$

The problem is now to determine the best values of f_j. There cannot be a unique solution for N greater than L, for there would then be more unknowns than determining equations. If N is equal to L no least-squares technique is needed, since there are $2L + 1$ equations to be solved for $2L + 1$ unknowns:

$$\sum_{j=-L}^{L} f_j \exp(2\pi i l j \gamma) - \Delta(l) = 0, \quad -L \le l \le L. \qquad (19)$$

In general this will have a single solution, though in particular cases there may be convergence problems. One particular choice of γ,

$$\gamma = (2L + 1)^{-1}, \qquad (20)$$

leads to an easy and illuminating solution; this choice just covers the length of the axis with equally spaced atomic sites. To find a particular f_j, say f_m, for this choice of sites, divide each of the $(2L + 1)$ equations (19) by $\exp(2\pi i l m \gamma)$ and add:

$$\sum_{j=-L}^{L} \sum_{l=-L}^{L} f_j \exp[2\pi i (j-m) l \gamma]$$

$$- \sum_{l=-L}^{L} \Delta(l) \exp(-2\pi i l m \gamma) = 0, \qquad (21)$$

or, after performing the summation over l in the first term and making use of equation (7),

$$\sum_{j=-L}^{L} f_j \frac{\sin[(2L + 1)\pi(j-m)\gamma]}{\sin[\pi(j-m)\gamma]} = A(m\gamma). \qquad (22)$$

Equation (22) is in fact valid, whatever the choice of γ, but for the choice (20) all the sine terms on the left-hand side vanish except that with $j = m$, giving

$$f_m = (2L + 1)^{-1} A[m/(2L + 1)]. \qquad (23)$$

In other words, within a constant factor the site-occupancy factors are just the values given by the Fourier synthesis at the points $m/(2L + 1)$. One may expect this result to hold *approximately* for other choices of γ, since the coefficient of f_m in equation (22) remains $(2L + 1)$, and the coefficients of the other terms are small (though not very small) in comparison.

When N is less than L the occupancy factors can be

chosen to give the best least-squares fit. The simplest residual to manipulate is

$$R_1' = \sum_{l=-L}^{L} [G(l) - \Delta(l)]^2, \qquad (24)$$

and it is reasonable to use it in the present problem, since the signs of the Δ's are known. Differentiating R_1' with respect to f_m and equating to zero gives

$$\frac{\partial R_1'}{\partial f_m} = 0 = 2 \sum_{l=-L}^{L} [G(l) - \Delta(l)] \exp(2\pi i m l \gamma), \qquad (25)$$

$$\sum_{l=-L}^{L} G(l) \exp(2\pi i m l \gamma) = \sum_{l=-L}^{L} \Delta(l) \exp(2\pi i m l \gamma), \qquad (26)$$

$$\sum_{j=-N}^{N} f_j \frac{\sin[(2L + 1)\pi(j + m)\gamma]}{\sin[\pi(j + m)\gamma]} = A(-m\gamma). \qquad (27)$$

With a change in the sign of m this takes the same form as equation (22), though the summation has $(2N + 1)$ terms instead of $(2L + 1)$, and one would expect that

$$f_m \sim (2L + 1)^{-1} A(m\gamma), \qquad (28)$$

as before. For

$$\gamma = n/(2L + 1), \qquad (29)$$

where n is any integer, equation (28) becomes exact.

A similar argument can be applied to the three-dimensional synthesis [equation (9)]. The analogue of equation (24) is

$$R_1' = \sum_{hkl} [G(l) - \Delta(hkl)]^2, \qquad (30)$$

$$\frac{\partial R_1'}{\partial f_m} = 0 = 2 \sum_{hkl} [G(l) - \Delta(hkl)] \exp(2\pi i m l \gamma), \qquad (31)$$

$$\sum_{hkl;j} f_j \exp[2\pi i(m + j)l\gamma]$$

$$= \sum_{hkl} \Delta(hkl) \exp(2\pi i m l \gamma), \qquad (32)$$

$$= B(-m\gamma) \qquad (33)$$

from equation (10). The left-hand side is easily seen to be

$$\sum_{j} f_j H[(m + j)\gamma], \qquad (34)$$

where

$$H(x) = \sum_{hkl} \exp(2\pi i l x), \qquad (35)$$

and is approximately given by equation (17). There is no choice of γ that will make $H[(m + j)\gamma]$ vanish for all but one value of j, but it is equal to the total number of reflections for $j = -m$, and has appreciably smaller values for all other j. Then, approximately,

$$f_m = N^{-1} B(m\gamma), \qquad (36)$$

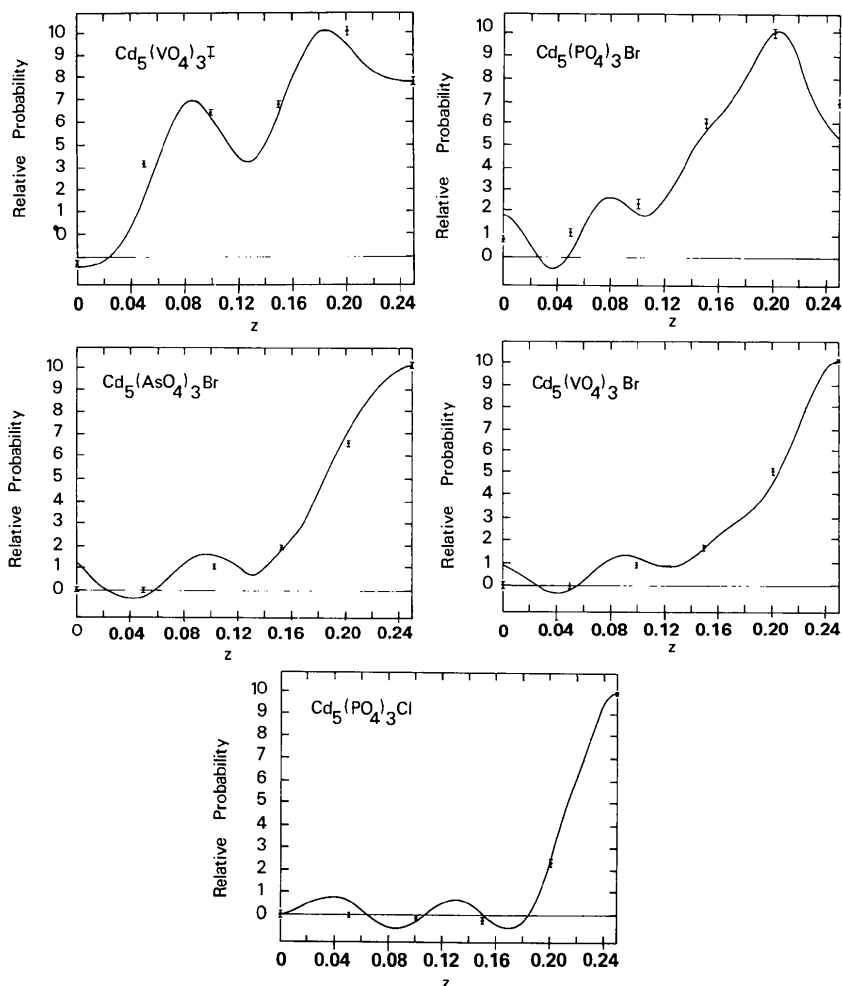where N is the total number of reflections.

Fig. 1. Comparison of halogen distributions as indicated by the $\Delta(**l)$ line syntheses (solid curve) and by least-squares refinements of site-occupancies at uniformly spaced sites (ⵣ). The error bars are e.s.d.'s.

It is interesting, though disappointing, to find that the least-squares refinement of occupancy factors suffers from the same defects as the Fourier-synthesis approach. The occupancy factors do not refine towards the 'true' distribution $f(z)$, but towards the distribution represented by the Fourier synthesis, reproducing its ripples exactly for sites chosen in accordance with equation (20) or (29) and approximately for other choices. A more general argument leading to similar conclusions about the relation of Fourier synthesis and least-squares refinement has been given by Wilson (1976a). The Fourier line-synthesis results and the least-squares site-occupancy refinement results for the five 'apatites' are graphically compared in Fig. 1.

2.3. Testing and optimization of a postulated distribution

If some particular form is expected for $f(z)$, depending on one or more adjustable parameters, the values of the parameters can be adjusted to give the best fit, in some sense, between the postulated distribution, say $g(z)$, and the 'true' distribution $f(z)$. It can be shown (Wilson, 1976a) that the mean-square difference between $f(z)$ and $g(z)$,

$$\int_0^1 [f(z) - g(z)]^2 \, dz, \tag{37}$$

has its minimum value for the same values of the parameters as those that minimize the residual R'_1 based on the Fourier coefficients of $f(z)$ and $g(z)$. Minimizing this residual is thus exactly equivalent to obtaining the least-squares fit between the functions. Weighting the terms in the residual corresponds to obtaining a least-squares fit between functions distorted so that their Fourier coefficients are multiplied by square-roots of the weights.

The postulated function of immediate interest is the uniform distribution discussed in §1. This leads to

$$f(z) = f, \quad \left. \begin{array}{l} \tfrac{1}{4} - \delta \le z \le \tfrac{1}{4} + \delta \\ \text{or } \tfrac{3}{4} - \delta \le z \le \tfrac{3}{4} + \delta \end{array} \right\} \tag{38}$$
$$= 0, \text{ otherwise,}$$

where f and δ are adjustable parameters. One has then

$$G(l) = f \int_{1/4-\delta}^{1/4+\delta} \exp(2\pi ilz)dz$$

$$+ f \int_{3/4-\delta}^{3/4+\delta} \exp(2\pi ilz)dz \tag{39}$$

$$= 2f(-)^{l/2} \frac{\sin 2\pi l\delta}{\pi l}, \quad l \text{ even} \left. \vphantom{\begin{array}{c}1\\1\\1\end{array}} \right\}$$
$$\tag{40}$$
$$= 0, \quad\quad\quad\quad\quad\quad l \text{ odd.} \left. \vphantom{\begin{array}{c}1\\1\end{array}} \right\}$$

Equation (24) becomes

$$R'_1 = \sum_{l(\text{even})} \left[2f(-)^{l/2} \frac{\sin 2\pi l\delta}{\pi l} - \Delta(l) \right]^2$$

$$+ \sum_{l(\text{odd})} \Delta^2(l), \tag{41}$$

and differentiation with respect to f and δ gives

$$\frac{\partial R'_1}{\partial f} = 0 = \sum_{l(\text{even})} \left[2f(-)^{l/2} \frac{\sin 2\pi l\delta}{\pi l} - \Delta(l) \right]$$

$$\times 4(-)^{l/2} \frac{\sin 2\pi l\delta}{\pi l}, \tag{42}$$

$$f = \frac{\left[\displaystyle\sum_{l(\text{even})} (-)^{l/2}\Delta(l) \frac{\sin 2\pi l\delta}{\pi l} \right]}{\left[2 \displaystyle\sum_{l(\text{even})} \frac{\sin^2 2\pi l\delta}{(\pi l)^2} \right]}; \tag{43}$$

$$\frac{\partial R'_1}{\partial \delta} = 0 = \sum_{l(\text{even})} \left[2f(-)^{l/2} \frac{\sin 2\pi l\delta}{\pi l} - \Delta(l) \right]$$

$$\times 8f(-)^{l/2} \cos 2\pi l\delta, \tag{44}$$

$$f \sum_{l(\text{even})} \frac{\sin 4\pi l\delta}{\pi l} = \sum_{l(\text{even})} (-)^{l/2}\Delta(l) \cos 2\pi l\delta. \tag{45}$$

Equations (43) and (45) must be solved simultaneously for f and δ; there is no obvious way of doing this except trial-and-error, and it was found just as easy to fit $G(l)$, as given by equation (40), directly to the observed $\delta(l)$ by trial-and-error.

It should be noted that a Fourier series with effectively $L/2$ terms, where L is at most 14, is not capable of producing a very good representation of the truncated uniform distribution postulated in equation (38). For small δ the representation is bell-shaped,

vaguely like a Cauchy distribution with some ripples. For larger values of δ the representation (not the distribution) becomes bimodal (approximately when the width of the distribution exceeds one-tenth of the length of the axis), and the ripples become more prominent. In Fig. 1, the $Cd_5(PO_4)_3Br$ and $Cd_5(VO_4)_3I$ cases are particular examples of this effect of producing false minima in the representation.

The factor $(-)^{l/2}$ in equation (40) arises only from the choice of origin; it would disappear if the origin were shifted to $(0,0,\tfrac{1}{4})$, but does appear in the expressions for $G(l)$ corresponding to any model distributions centered on $(0,0,\tfrac{1}{4})$ and $(0,0,\tfrac{3}{4})$. For values of δ greater than $1/2L$ (about 0·04 in the present case) the factor $(\sin 2\pi l\delta)/\pi l$ in equation (40) produces non-trivial changes in sign and non-monotonic decreases in magnitude within the observable range of l. Other model distributions that might be considered (Gaussian or normal, Cauchy or Lorentzian, Laplacian ...) lead to Fourier components $G(l)$ decreasing smoothly with l, and thus cannot be made to match values of $\Delta(l)$ that show anything other than alternation of sign and monotonic decrease in magnitude. The observed signs of $\Delta(l)$ thus give some immediate information about the type of distribution, but it is difficult to utilize it positively; it can rule out some model distributions, but may not he'p in finding satisfactory ones.

The remark following equation (35) applies also to the optimization of postulated distributions. The fit obtained is between the model and the observed distribution as represented imperfectly by its observed Fourier components, and not between the model and the real distribution.

3. Experimental results for $\Delta_o(**l)$

According to the theory developed in equations (1)–(5), the quantities

$$\Delta_o(hkl) = \frac{F_o(hkl) - F_1(hkl)}{X(hkl)}, \tag{46}$$

where $F_o(hkl)$ is the observed structure factor, $F_1(hkl)$ is the structure factor calculated for the non-halogen atoms only, and $X(hkl)$ is the temperature-corrected atomic scattering factor for the halogen atoms, should have the value zero for l odd, and should have the same value for all h and k for any given even l. Since $\Delta_o(hkl)$ depends on a small difference between two large quantities, one expects, and finds, a considerable scatter in its 'observed' values. The amount of the scatter gives some indication of the reliability of the measurements. The average values (fixed l, variable h and k) for $\Delta_o(**l)$ are collected in Table 1 for the available cadmium 'apatites' (Sudarsanan, Young & Donnay, 1973; Sudarsanan, Young & Wilson, 1977). The value tabulated in each case is in fact that corresponding to

Table 1. *Values of* $\Delta_o(**l)$ ($\times 10^4$) *and standard deviations for five cadmium 'apatites'*

	Cd$_5$(VO$_4$)$_3$I			Cd$_5$(PO$_4$)$_3$Br			Cd$_5$(AsO$_4$)$_3$Br			Cd$_5$(VO$_4$)$_3$Br			Cd$_5$(PO$_4$)$_3$Cl		
l	$\Delta_o(**l)$	σ	N	$\Delta_o(**l)$	σ	N	$\Delta_o(**l)$	σ	N	$\Delta_o(**l)$	σ	N	$\Delta_o(**l)$	σ	N
0	6233	264	141	7605	291	108	8476	196	178	8059	320	100	8916	390	149
1	−21	45	155	−212	108	81	47	59	125	21	71	145	339	290	159
2	−2103	71	155	−3750	192	83	−5902	208	138	−5507	184	148	−7465	424	161
3	12	50	157	141	89	79	−52	75	149	−53	97	171	−137	280	159
4	−577	52	151	113	122	92	2720	79	142	2784	103	151	7147	338	144
5	7	60	128	−108	118	77	−60	70	153	66	84	148	−74	274	156
6	−382	56	155	797	161	69	−1408	79	127	−1773	109	145	−4775	393	152
7	80	70	148	353	153	71	−120	83	87	103	130	140	140	343	134
8	−942	69	138	−937	167	88	255	120	85	940	154	101	3704	463	125
9	−22	82	117	−120	167	71	93	134	70	110	192	93	74	432	111
10	267	83	107	824	258	68	346	175	67	−158	252	94	−1952	806	86
11	127	132	81	217	199	64	−65	222	55	275	388	74	282	1166	64
12	435	157	47	−19	478	41	316	276	44	902	321	55	−809	1866	52
13	−558	478	24	−421	594	58	36	519	31	−379	850	21	—	—	
14	−28	284	40	814	735	50	—	—		—	—		—	—	
15	—	—		−76	1112	18	—	—		—	—		—	—	

half the unit cell; if the compound were stoichiometric the tabulated $\Delta_o(**l)$ for $l = 0$ should be unity. The table also gives the number of observations, N, on which each average is based, and a standard deviation, σ, calculated as

$$\sigma^2 = \frac{\sum_{h, \, \cdot} [\Delta_o(hkl) - \Delta_o(**l)]^2}{N(N - 1)}. \tag{47}$$

It should be noted that σ is merely a measure of internal consistency; it is not explicitly based on counting statistics and other random errors, and makes no explicit allowance for effects arising from atomic scattering factors, temperature factor, scaling factors, *etc.*

The line syntheses produced with these $\Delta(**l)$ values are shown in Fig. 1 along with the site-occupancy refinement results for the same physical feature, the distribution of halogen ions along the 6$_3$ axis.

3.1. *Allowance for dispersion*

Equation (46) is formally correct, whether or not the atomic scattering factors have appreciable in-quadrature components. If the numerator is in fact complex, of the form $A + iB$, and the denominator of the form $a + ib$, then

$$\Delta_o(hkl) = \frac{A + iB}{a + ib} \tag{48}$$

$$= \frac{(A + iB)(a - ib)}{(a + ib)(a - ib)} \tag{49}$$

$$= \frac{Aa + Bb}{a^2 + b^2} + i \frac{Ba - Ab}{a^2 + b^2}. \tag{50}$$

Ideally the numerator and denominator in equation (48) would have the same phase, and the imaginary part of equation (50) would be zero. In fact, the

imaginary term was always small, and subject to a large statistical error. The quantity given in Table 1 is the first term, $(Aa + Bb)/(a^2 + b^2)$.

3.2. *Effects of systematic error*

The signs of the $\Delta(hkl)$'s depend on the relative magnitudes of the observed structure factors and of the structure factors calculated for the non-halogen atoms only. The temperature-corrected atomic scattering factor for the halogen atom, $X(hkl)$ in equation (46), enters into the determination of the *magnitude* of $\Delta(hkl)$, but cannot affect its *sign*. Since the thermal parameter β_{33} of the halogen atoms is not reliably determinable from the structure refinement (being too strongly correlated with other parameters of the model), only a value estimated by analogy with similar compounds can be used in calculating $\Delta(hkl)$; one must, therefore, conclude that the signs of the $\Delta(hkl)$'s, and particularly of the $\Delta(**l)$'s, are more reliable than their magnitudes. This statement is especially true for the higher values of l; $\Delta(**0)$ does not depend on β_{33}, but only on β_{11} and β_{22}, and $\Delta(**2)$ is little affected.

The signs as well as the magnitudes of the $\Delta(hkl)$'s are, however, strongly dependent on the temperature factors of the nonhalogen atoms, and on the overall scaling factor. It is becoming recognized (Wilson, 1974; Lomer & Wilson, 1975) that the overall scaling factors obtained by different least-squares refinement procedures may differ somewhat from each other and from that required to produce calculated intensities that are equal, on the average, to the observed intensities, and the same is true for the overall temperature factor (Wilson, 1976b). What is wanted in the present case is, presumably, a third, again slightly different, set of parameters that makes the average observed and calculated structure factors equal. The values of $\Delta(**l)$ for l odd give some indication of whether uncertainties

Table 2. *Estimates of halogen content of half unit cell* (*stoichiometric composition* = 1·000)

Figures in parentheses are standard deviations. For explanation see text.

| Compound | Single atom at $(0,0,\tfrac{1}{4})$[(i)] | Sum of occupancy factors[(ii)] | $\Delta_o(**0)$ for all reflexions and for stated ranges of $\sin^2\theta$ | | | | | | From length of c axis with | |
			All[(iii)]	≤ 0.11	>0.11 ≤ 0.22	>0.22 ≤ 0.33	>0.33 ≤ 0.44	>0.44	ionic radii	crystal radii
$Cd_5(VO_4)_3I$	0·800 (11)	0·734 (9)	0·623 (26)	0·670 (50)	0·684 (58)	0·641 (59)	0·525 (76)	0·561 (54)	0·74	0·79
$Cd_5(PO_4)_3Br$	0·838 (11)	0·830 (19)	0·761 (29)	0·846 (17)	0·705 (68)	0·825 (43)	0·685 (72)[(iv)]		0·83	0·89
$Cd_5(AsO_4)_3Br$	0·911 (6)	0·925 (16)	0·848 (20)	0·849 (15)	0·872 (29)	0·763 (66)	0·847 (43)	0·903 (45)	0·83	0·90
$Cd_5(VO_4)_3Br$	0·880 (8)	0·874 (12)	0·806 (32)	0·835 (59)	0·776 (88)	0·740 (70)	0·868 (27)[(iv)]		0·83	0·90
$Cd_5(PO_4)_3Cl$	0·836 (10)	0·845 (19)	0·892 (39)	0·774 (109)	0·894 (52)	0·886 (41)	0·877 (90)	0·975 (146)	0·90	0·97

(i) Taken from Table 5 of Sudarsanan, Young & Wilson (1977). (ii) Taken from Table 6 of Sudarsanan, Young & Wilson (1977). (iii) Taken from Table 1 of this paper. (iv) All reflections with $\sin^2\theta > 0.33$.

in the overall scaling factors and in the temperature parameters of the nonhalogen atoms are large enough to vitiate any inferences drawn from the signs and magnitudes of the $\Delta(**l)$'s for l even. An examination of Table 1 permits the following conclusions.

3.2.1. The magnitudes of $\Delta(**l)$ for l odd, theoretically zero, are in fact satisfactorily close to zero for all odd values of l and for all five substances. For about three-quarters of the entries in Table 1 the magnitude of $\Delta(**l)$ for l odd is less than its standard deviation, as given by equation (47); it exceeds twice its standard deviation only for (**7) of $Cd_5(PO_4)_3Br$.

3.2.2. There is no tendency for $\Delta(**l)$ for l odd to be predominantly positive or predominantly negative for any substance.

3.2.3. There is no detectable tendency for a change in predominant sign with increasing l, either for any substance, or for all substances taken together.

3.2.4. Conclusion 3.2.2 is sensitive to errors in overall scaling factors, and conclusion 3.2.3 to errors in the temperature parameters of the nonhalogen atoms. It thus appears that, within the accuracy implied by the standard deviation of $\Delta(**l)$ for l odd, uncertainties in the scaling factors and in the nonhalogen temperature parameters are not a source of substantial systematic

errors. Errors in the positional parameters of the nonhalogen atoms will add to the scatter of the $\Delta(hkl)$'s for l odd, and will thus increase the standard deviation of $\Delta(**l)$ for l odd.

3.2.5. It is reasonable to regard the standard deviation of $\Delta(**l)$ for l odd, therefore, as an estimate for the magnitude of the effect of errors and uncertainties in the nonhalogen part of the structure on the magnitude of $\Delta(**l)$, even when l is even. It is thus prudent to regard the magnitude of $\Delta(**l)$ for even l as being possibly subject to a 'zero error' of the order of the standard deviations of $\Delta[**(l-1)]$ and $\Delta[**(l+1)]$, in addition to a random error given by its own standard deviation.

3.2.6. For l even the values of $\Delta_o(**l)$ depend critically on the assumed values of the atomic scattering factor and of the thermal parameters for the halogen atom. In principle, the correctness of these can be checked by examining the behavior of $\Delta_o(**l)$ calculated for annuli in reciprocal space corresponding to increasing mean values of $\sin^2\theta$. The results for the $hk0$ reflections are given in Table 2, columns 5 to 9, the ranges of $\sin^2\theta$ being 0–0·11, 0·11–0·22, 0·22–0·33, 0·33–0·44, >0·44. These ranges were chosen as containing approximately equal numbers of reflections

Table 3. *Values of $\Delta_o(**4)$ for various ranges of $\sin^2\theta$*

Figures in parentheses are standard deviations.

| Compound | Range of $\sin^2\theta$ | | | | | |
	All	≤ 0.11	>0.11 ≤ 0.22	>0.22 ≤ 0.33	>0.33 ≤ 0.44	>0.44
$Cd_5(VO_4)_3I$	−0·058 (5)	−0·071 (18)	−0·050 (13)	−0·056 (9)	−0·055 (11)	−0·064 (12)
$Cd_5(PO_4)_3$ Br	0·011 (12)	0·020 (15)	0·023 (12)	0·017 (26)	−0·062 (58)	—
$Cd_5(AsO_4)_3Br$	0·272 (8)	0·278 (12)	0·268 (11)	0·273 (10)	0·275 (20)	0·269 (20)
$Cd_5(VO_4)_3Br$	0·278 (10)	0·274 (19)	0·290 (24)	0·309 (11)	0·266 (25)	0·238 (28)
$Cd_5(PO_4)_3Cl$	0·715 (34)	0·751 (86)	0·648 (37)	0·711 (41)	0·688 (89)	0·827 (124)

($\sim$25); for the two compounds marked (iv) in the table there were comparatively few higher-angle reflections and the last two zones were combined. For most of the substances there is no indication of a trend with $\sin^2\theta$. For one, $Cd_5(VO_4)_3I$, $\Delta_o(**0)$ appears to decrease with $\sin^2\theta$, but the change is within the possible statistical scatter; this matter is discussed further in §4.2.4. It seems, therefore, that the assumed values of the atomic scattering factor and of the thermal parameters are confirmed within the (rather large) statistical scatter.

As the $hk0$ reflections are a special case, a similar correlation was made for the $hk4$ reflections. The results are given in Table 3. Except for $Cd_5(PO_4)_3Br$ there are about 10 reflections in the first range of $\sin^2\theta$ and about 30 in the others; $Cd_5(PO_4)_3Br$ has 20 in the first range, but only 10 in the fourth, for which the sign is anomalous. No compound shows a consistent trend of $\Delta_o(**4)$ with increasing $\sin^2\theta$, again confirming the general correctness of the β's and atomic scattering factors. The values for $Cd_5(VO_4)_3I$ are too small, in comparison with their standard deviations, to confirm or contradict the trend suggested by the $hk0$'s for this substance.

4. Interpretation

4.1. *Deductions about the halogen distribution function from the signs, only, of the $\Delta_o(**l)$'s*

We have seen that (i) the signs of $\Delta_o(**l)$ are better determined than their magnitudes, (ii) a uniform-distribution model gives values of $\Delta_o(**l)$ that do not necessarily decrease monotonically with l and that do not necessarily exhibit the simple alternation of sign in accordance with $(-)^{l/2}$, and (iii) Gaussian distributions and similar smooth models necessarily exhibit the properties denied for the uniform-distribution model. Examination of the signs in Table 1 leads to the following conclusions.

(i) $Cd_5(VO_4)_3I$. Gaussian and similar distributions are excluded by the signs for $l = 4$, 8 and 10. A uniform distribution with δ about 0·20 would give the correct signs up to $l = 12$; the negative sign for $l = 14$ is not significant since $\Delta_o(**14)$ is only one-tenth of its standard deviation.

(ii) $Cd_5(PO_4)_3Br$. Gaussian and similar distributions are excluded by the signs for $l = 6$ and 8, and are contra-indicated by the signs for $l = 10$, 12 and 14, though these are not strongly determined. A uniform distribution with $\delta \sim 0·11$ is contra-indicated only by $l = 10$.

(iii) $Cd_5(AsO_4)_3Br$. Gaussian and similar distributions are contra-indicated by the sign for $l = 10$, though this is not strongly determined. A uniform distribution with $\delta \sim 0·06$ gives the signs up to $l = 10$ correctly, but is contra-indicated by the sign for $l = 12$;

this is not significant, since $\Delta_o(**12)$ is only 1·1 times its standard deviation.

(iv) $Cd_5(VO_4)_3Br$ and $Cd_5(PO_4)_3Cl$. The signs are consistent with any narrow distribution (except for a very slight contra-indication of the Gaussian and similar distributions by $l = 12$ for the last-named). If the distribution is uniform, $\delta \sim 0·04$ for both compounds.

4.2. *Non-stoichiometry: halogen deficiency*

Direct chemical analysis of bulk material from the same batch cannot be expected to give a reliable indication of the composition of the single crystal actually used for structure determination. In addition to the usual uncertainties in chemical analysis for minor constituents – the halogen is of the order of 5% of the total – the proportion of halogen may differ in different crystals from the same batch, and analysis of the bulk material may be further vitiated by surface contaminants. It is felt, therefore, that refinement of the structure gives a better indication of the halogen content of the actual crystal. This and part I (Sudarsanan, Young & Wilson, 1977) give four estimates of the halogen content of each substance. These estimates, collected in Table 2, differ in their statistical reliability and the assumptions made in obtaining them.

4.2.1. *From length of the hexad axis.* The simplest estimate is the ratio of half the length of the hexad axis to the accepted ionic diameter of the halogen in question. Since the halogen ions do pack together along c, the effective diameter of a halogen ion along c may be more than that of the ion coordinated only by cations but less than that in the crystalline form of the halogen. Both types of diameters were used for the estimates shown in the last two columns of Table 2. These estimates of halogen content are based on the assumption that the hexad axis is uniformly filled with halogen ions, which is contradicted by the observation that $\Delta_o(00l)$ is non-zero for some even $l \neq 0$. Therefore, the estimates of halogen content obtained with the crystal diameters, especially, should be regarded as an upper limit.

4.2.2. *From refinements with atoms at $\pm(0,0,\frac{1}{4})$ only.* Part I gives site-occupancy factors obtained by refinement of a model structure with halogen atoms placed at $\pm(0,0,\frac{1}{4})$ only; the values obtained are reproduced in Table 2, column 2. The corresponding temperature parameters β_{33} are very large, and it is easy to see that this refinement is equivalent to the use of a model with a Gaussian distribution of halogens in the sense of §2.3 of the present paper. It was shown in §4.1 that a Gaussian distribution is excluded for the first two compounds, and possibly for the third. The halogen content obtained by this method is thus subject to systematic error for the first two compounds, and suspect for the third.

4.2.3. *From refinement with uniformly spaced atoms.* Part I gives occupancy factors obtained by refining a model structure with halogen atoms placed at uniform intervals of $0.05c$ along the hexad axis. The sum of these occupancy factors is given in Table 2, column 3, and should be equal to the halogen content of the half-cell. The method is subject to some systematic error, in that it approximates what is probably a continuous distribution by a discrete set of positions. The error could, in principle, be investigated by refining models with different spacings, but in practice the interval is limited on one side by the need to use enough points to simulate the distribution, and on the other side by indeterminateness if the number of positions is greater than about $\frac{1}{2}L$. This estimate of halogen content may perhaps be expected to be more reliable for wider distributions (*i.e.* for those compounds at the top of the table) than for the narrower ones.

4.2.4. *From mean structure factor for $l = 0$.* The values of $\Delta_o(**0)$ given in Table 1 and reproduced in column 4 of Table 2 should be equal to the halogen content of a half-cell. In principle these are the best estimates of the halogen content, as they are independent of any model of this halogen distribution. In practice, they depend critically on the scaling factors, the positional and thermal parameters of the nonhalogen atoms, and on the values of β used in determining $X(hk0)$ in equation (3). The discussion in §§3.2.1 to 3.2.4 indicates that the scaling factors and the nonhalogen part of the structure are correct within the statistical scatter. The measures of halogen content discussed in §§4.2.2 and 4.2.3 also depend on these parameters, but the dependence is not so obvious. The discussion in §3.2.6 shows that, within the considerable statistical scatter, $Cd_5(VO_4)_3I$ gives some reason to suspect an underestimate of β_{11}. If the trend of $\Delta_o(**0)$ with $\sin^2 \theta$ exhibited by this substance is real, the halogen content by this method, extrapolating back to $\theta = 0$, should be about 0.715 instead of 0.623. There is a similar but much weaker indication for $Cd_5(PO_4)_3Br$; if accepted as real, $\Delta_o(**0)$ would be raised to about 0.785.

The preceding discussion leads to the estimates of the content of the half-cell shown in Table 4. The estimates are in reasonable agreement with each other.

4.3. Attempts at quantitative agreement for halogen-distribution models

4.3.1. $Cd_5(VO_4)_3I$. The signs of $\Delta_o(**l)$ are well determined $(\Delta/\sigma > 3)$ up to $l = 10$. To obtain the correct sign for $l = 10$, $2\pi\delta \times 10$ must be less than 4π $(\delta < 0.2)$, and for $l = 8$, $2\pi\delta \times 8$ must be greater than 3π $(\delta > 0.1875)$. Any choice of δ in this range will also give the observed sign for $l = 12$ $(\Delta/\sigma = 2.8)$, and no choice will give the observed sign for $l = 14$ $(\Delta/\sigma = 0.1)$. The best (but very poor) quantitative agreement overall is obtained for δ near the upper limit (0.199) and $f = 0.864$ (Table 5). The agreement cannot be substantially improved by adjustment of β_{33}; a change that improves $l = 4$ and 6 worsens $l = 2$ and 8.

4.3.2. $Cd_5(PO_4)_3Br$. The signs of $\Delta_o(**l)$ are well determined up to $l = 8$, that for $l = 10$ is borderline $(\Delta/\sigma = 3.1)$. The signs for $l = 6$ and 8 require that $\pi/6 < 2\pi\delta < 2\pi/8$ $(0.083 < \delta < 0.125)$. The small magnitude of $\Delta_o(**4)$ suggests a value near the upper limit; the sign for $l = 10$, if accepted, requires $2\pi\delta < 2\pi/10$ $(\delta < 0.10)$. It is not possible, however, to obtain quantitative agreement (Table 6), and the sign for $l = 10$ is incorrect. The sign for $l = 12$ is not significant $(\Delta/\sigma = 0.04)$. The value of δ implied is 0.113.

Table 5. *Observed and calculated values of $\Delta(**l)$ for* $Cd_5(VO_4)_3I$

l	$\Delta_o(**l)(\sigma)$	$(-)^{l/2}0.55(\sin 1.25l)/l$	Difference $\frac{}{\sigma}$
0	0.623 (26) [0.715]	0.687	2.5 |1.1|
2	−0.210 (7)	−0.165	6
4	−0.058 (5)	−0.132	15
6	−0.038 (6)	−0.086	8
8	−0.094 (7)	−0.037	8
10	0.027 (8)	0.004	2.9
12	0.044 (16)	0.030	0.9
14	−0.003 (28)	0.038	1.5

Table 4. *Summary of halogen-deficiency estimates*

	From single-halogen occupancy factor	From multiple-halogen occupancy factors	From $\Delta_o(**0)$	From c-axis length
$Cd_5(VO_4)_3I$	0.800 (11)	0.734 (9)	>0.623 (26) [0.715?]†	0.74–0.79
$Cd_5(PO_4)_3Br$	0.838 (11)	0.830 (19)	>0.761 (29) [0.785?]†	0.83–0.89
$Cd_5(AsO_4)_3Br$	0.911 (6)	0.925 (16)	0.848 (20)	0.83–0.90
$Cd_5(VO_4)_3Br$	0.880 (8)	0.874 (12)	0.806 (32)	0.83–0.90
$Cd_5(PO_4)_3Cl$	0.836 (10)	0.845 (19)	0.892 (39)	0.90–0.97

† See text (§4.2.4).

Table 6. *Comparison of observed and calculated* $\Delta(\!*\!*l)$ *for* $Cd_5(PO_4)_3Br$

l	$\Delta_o(\!*\!*l)(\sigma)$	$(-)^{l/2}0.99(\sin 0.71l)/l$	Difference $\overline{\sigma}$
0	0.7605 (291)	0.703	2.0
2	−0.3750 (192)	−0.489	5.9
4	0.0113 (122)	0.0735	5.1
6	0.0796 (161)	0.148	4.3
8	−0.0937 (167)	−0.070	1.4
10	0.0824 (258)	−0.0722	6.0
12	−0.0019 (478)	0.065	1.4
14	0.0814 (735)	0.035	0.6

4.3.3. $Cd_5(AsO_4)_3Br$. Since the signs for $l = 8$ and 10 are positive, $\pi/10 < 2\pi\delta < \pi/8$, so that $0.05 < \delta < 0.0625$. Reasonable qualitative agreement is obtained for $2\pi\delta = 0.36$ ($\delta = 0.057$) (Table 7), but the calculated magnitudes are systematically large for large l. An additional temperature factor would improve the agreement considerably, but empirically a factor of the form $\exp(\alpha|l|)$ is better than one of the form $\exp(\beta l^2)$. With $\alpha = 0.136$ the only serious discrepancy is for $l = 4$ (columns 6 and 7 of Table 7). This discrepancy can be reduced, at the expense of larger discrepancies elsewhere, by altering the distribution parameters slightly, as in the last three columns of Table 7 ($\delta = 0.060$).

4.3.4. $Cd_5(VO_4)_3Br$. The signs of $\Delta_o(\!*\!*l)$ alternate simply as $(-)^{l/2}$, and Gaussian and similar distributions cannot be excluded on the basis of signs alone. The magnitude of $\Delta_o(\!*\!*l)$ for $l = 12$ is nearly three times its standard deviation; its sign suggests $2\pi\delta \times 12 < \pi$ ($\delta < 0.04$), but no choice of δ gives numerical agreement for

lower values of l. Reasonable agreement can be obtained by inflating the observed values by a factor of $\exp(\alpha|l|)$, as for $Cd_5(AsO_4)_3Br$, with $2\pi\delta = 0.22, \alpha = 0.195$.

It must be remembered that the assumption of a uniform distribution is unnecessary for this compound. If the $\Delta_o(\!*\!*l)$ were inflated by $\exp(0.267l)$ the result would be nearly constant, corresponding to 0.81 of an atom at $\pm(0,0,\frac{1}{4})$. The equivalent reverse comparison is made in Table 8, in which the seventh column contains

$$\Delta(\!*\!*l) = (-)^{l/2}0.806 \exp(-0.23l). \qquad (51)$$

The agreement is about as good as that between columns 3 and 4. [The ratio of difference to standard deviation is the same, whether the observed Δ's are inflated or the calculated ones are deflated.] Calculated $\Delta(\!*\!*l)$'s of the form of equation (51) correspond to a Cauchy distribution of atomic centers.

4.3.5. $Cd_5(PO_4)_3Cl$. The small size of $\Delta_o(\!*\!*10)$ and the weakly indicated sign of $\Delta_o(\!*\!*12)$ suggest $\pi/12 <$

Table 9. *Observed and calculated values of* $\Delta(\!*\!*l)$ *for* $Cd_5(PO_4)_3Cl$

l	$\Delta_o(\!*\!*l)(\sigma)$	$(-)^{l/2}3.25 (\sin 0.264l)/l$	Difference $\overline{\sigma}$
0	0.8916 (390)	0.858	0.9
2	−0.7465 (424)	−0.818	1.7
4	0.7147 (338)	0.707	0.2
6	−0.4775 (393)	−0.542	1.6
8	0.3704 (463)	0.348	0.5
10	−0.1952 (806)	−0.157	0.5
12	−0.0809 (1866)	−0.007	0.4

Table 7. *Observed and calculated values of* $\Delta(\!*\!*l)$ *for* $Cd_5(AsO_4)_3Br$

l	$\Delta_o(\!*\!*l)\,(\sigma)$	$(-)^{l/2}2.35 \times$ $(\sin 0.36l)/l$	$\Delta_o(\!*\!*l) \times$ $\exp(0.0232l^2)$	Difference σ	$\Delta_o(\!*\!*l) \times$ $\exp(0.136l)$	Difference σ	$(-)^{l/2}2.23 \times$ $(\sin 0.38l)/l$	$\Delta_o(\!*\!*l) \times$ $\exp(0.147l)$	Difference σ
0	0.8476 (196)	0.846	0.848 (20)	0.1	0.848 (20)	0.1	0.847	0.848 (20)	0.1
2	−0.5902 (208)	−0.775	−0.648 (26)	4.9	−0.775 (27)	0	−0.768	−0.792 (28)	0.9
4	0.2720 (79)	0.582	0.394 (11)	17	0.469 (14)	8	0.557	0.490 (14)	4.8
6	−0.1408 (79)	−0.326	−0.325 (18)	0.1	−0.318 (18)	0.4	−0.291	−0.340 (19)	2.6
8	0.0255 (120)	0.076	0.113 (53)	0.7	0.076 (36)	0	0.028	0.083 (39)	1.4
10	0.0346 (175)	0.104	0.352 (178)	1.4	0.135 (68)	0.5	0.136	0.150 (76)	0.1
12	0.0316 (276)	−0.181	0.892 (779)	1.4	0.162 (141)	2.4	−0.184	0.184 (161)	2.3

Table 8. *Observed and calculated* $\Delta(\!*\!*l)$ *for* $Cd_5(VO_4)_3Br$

l	$\Delta_o(\!*\!*l)(\sigma)$	$\Delta_o(\!*\!*l) \exp(0.195l)$	$(-)^{l/2}3.66 (\sin 0.22l)/l$	Difference σ	$(-)^{l/2}0.806 \exp(-0.23l)$	Difference σ
0	0.8059 (320)	0.806 (32)	0.805	0.0	0.806	0.0
2	−0.5507 (184)	−0.813 (27)	−0.779	1.3	−0.499	2.9
4	0.2784 (103)	0.607 (22)	0.705	4.5	0.309	3.0
6	−0.1773 (109)	−0.571 (35)	−0.591	0.6	−0.191	1.3
8	0.0940 (154)	0.447 (73)	0.449	0.0	0.118	1.6
10	−0.0158 (252)	−0.111 (177)	−0.296	1.0	−0.073	2.3
12	0.0902 (321)	0.936 (333)	0.147	2.4	0.045	1.4

$2\pi\delta < \pi/10$, and good agreement is obtained for $\delta = 0\cdot0420$ (Table 9). Good agreement cannot be obtained for calculated $\Delta(**l)$'s of the form of equation (51).

5. Summary

5.1. Fourier synthesis

The probability distribution of atomic centers obtained by Fourier synthesis based on the observable reflections is broadened and rippled through termination effects. These effects are more prominent for a line section through the three-dimensional synthesis than for a one-dimensional synthesis. Deconvolution to remove the termination effects is not possible; observation of reflections of higher order would make it possible to reduce them.

The ripples in the syntheses (Fig. 1) are perhaps more prominent than would be expected from termination effects; in particular, the dip at $z \sim 0\cdot12$ may be significant (§5.3).

5.2. Site-occupancy refinement

The theory in §2.2 indicates that refinement of site occupancies does not avoid termination effects; site occupancies may be expected to refine toward approximately the same broadened and rippled distribution as is obtained by Fourier synthesis. This conclusion is supported by the comparisons in Fig. 1.

5.3. Fitting of postulated distributions

The uniform distribution is the only simple model, among those that have been tried, that gives qualitative agreement with the 'observed' values of $\Delta(**l)$ for all five compounds. Gaussian and Cauchy distributions are inconsistent with the signs for the first two (§§4.3.1 and 4.3.2). A Cauchy component improves the agreement for the third and fourth (§§4.3.3 and 4.3.4); in fact for the fourth, $Cd_5(VO_4)_3Br$, a Cauchy is at least as satisfactory as a uniform distribution. Only for the last, least imperfect, compound, $Cd_5(PO_4)_3Cl$ (§4.3.5),

is agreement with a uniform-distribution model entirely satisfactory.

The discrepancies between the 'observed' $\Delta(**l)$'s and those calculated for a uniform distribution are different for the different compounds, but there appears to be one possibly significant feature: $\Delta(**4)$ is smaller than expected for all compounds but $Cd_5(PO_4)_3Cl$. This may be a further manifestation of whatever feature of the distribution is associated with the dip at $z \sim 0\cdot12$; this is repeated by the symmetry at $z \sim 0\cdot38, 0\cdot62, 0\cdot88$, i.e. at intervals of approximately one-quarter of the hexad axis.

5.4. Non-stoichiometry

The estimates of halogen content by four independent methods, collected in Table 4, are in reasonable agreement with each other. That obtained by close-packing of the halogens along the hexad axis cannot be considered quantitative, since (i) literal close-packing would imply that $\Delta_o(00l) = 0$ for $l \neq 0$, contrary to observation, and (ii) the effective ionic diameter is unknown, being dependent on environment. The other three estimates of halogen content agree, in general, within two standard deviations. Their average value leads to the estimates of halogen deficiency $2x$ given in the first column of Table 10. The estimate of scatter given in parentheses is based on their mean deviation from the mean. In all five cases the deficiency is several times larger than the scatter, so it must be concluded that it corresponds with some physical reality, but the possibility is not excluded that there is some substitution of a lighter ion, perhaps F or O, for an occasional heavy halogen, rather than an actual vacancy.

5.5. Deductions from the model

The number of halogen atoms in a chain, as pictured in the model in §1 of this paper, is

$$2n_{max} + 1 = (2x)^{-1} - 1, \qquad (52)$$

and runs from about three for $Cd_5(VO_4)_3I$ to about

Table 10. *Halogen-ion deficiency, width of distribution, chain length, and effective X diameter in cadmium 'apatites'*

Compound	Deficiency $2x$ (Table 4)	δ (§§4.3.1– 4.3.5)	Number of X atoms in chain (equation 52)	Effective X diameter (equation 53) (Å)	Ionic diameter* (Å)	Crystal diameter* (Å)	Separation in X_2 molecule (Å)
$Cd_5(VO_4)_3I$	0·250 (34)	0·199 (4)	3·0 (0·6)	4·2–5·0	4·40	4·12	2·66
$Cd_5(PO_4)_3Br$	0·182 (23)	0·113 (3)	4·5 (0·8)	3·65 (10)			2·29
$Cd_5(AsO_4)_3Br$	0·105 (31)	0·058 (2)	8·5 (3·0)	3·36 (5)	3·92	3·64	2·29
$Cd_5(VO_4)_3Br$	0·147 (32)	0·035 (2)	5·8 (1·9)	3·36 (4)			
$Cd_5(PO_4)_3Cl$	0·142 (23)	0·042 (2)	6·0 (1·4)	3·35 (3)	3·62	3·34	1·99

* Shannon (1976).

eight for $Cd_5(AsO_4)_3Br$ (Table 10, column 3). The effective ionic diameter depends on the spread, δ, of atomic-center sites about $z = \frac{1}{4}$ as well as on the halogen deficiency. From the model in §1 it is easily seen that the interionic spacing is

$$X - X = \tfrac{1}{2}c + \delta \frac{4x}{1 - 4x} c, \qquad (53)$$

where c is the length of the hexad axis. Values of δ from §§4.3.1–4.3.5 are given in column 2 of Table 10; the estimates of reliability are based on the amount by which δ can be varied without seriously affecting the sign agreements for the $\Delta_o(**l)$'s. A check against blunders in δ is given by comparing $\frac{1}{4} - \delta$ with the value of z at which the relative probability falls to one-half in Fig. 1; agreement is within 0·01.

Effective ionic diameters (interionic spacings) determined from equation (53) and $2x$, δ, and c are given in column 4 of Table 10, with uncertainties corresponding to those in $2x$ and δ. That for the iodide has a very large uncertainty. The others are appreciably smaller than the usual ionic diameters (Shannon, 1976), the difference ranging from 7% to 16%. They are very much larger than the homopolar distances in X_2 molecules, and agree better with the 'crystal diameters' given by Shannon.

6. Conclusions

The four different analytical approaches have all led to the conclusion that these compounds are halogen-deficient and that the halogens are distributed over a variety of positions, along the c axis, centered on the ideal position $(0,0,\frac{1}{4})$.

The different estimates of the amount of halogen present generally agreed within 10–15%. It was not possible to state with certainty which method led to the better estimate, one interesting reason being that least-squares refinement of the distributed occupancy factors, it is shown, suffers from the same defects as does the Fourier-synthesis approach: the occupancy factors do not refine to the true values but toward the distribution represented by the Fourier series, reproducing its ripples exactly in a special case of uniformly distributed occupancy points and approximately in other cases.

For the actual distribution of the halogens a square-wave model showed no serious disagreements with observation whereas Gaussian and Cauchy distri-

butions did and, hence, could be ruled out as a class. On this model, average chain lengths could be calculated and were found to be physically reasonable, ranging from about three ions in the iodide case to about eight ions in the arsenate bromide case.

Wholly unequivocal determinations of the distribution of halogen require a resolution of adjacent scattering points (effectively in the Fourier transform of the reciprocal-space intensity data) better than that available even though Mo $K\alpha$ radiation was used. This resolution is noted to be different from the precision with which the position of an individual atom may be determined when it is at least $\lambda/2$ from its nearest neighbor, λ being the wavelength associated with the highest-order Fourier term (farthest out reciprocal-lattice point) used.

It was found that the two ordinarily used methods of determining the scattering density along a line using three-dimensional data, a line synthesis and a line section, did differ in resolution, the line synthesis exhibiting about 30% less smearing with the same data.

These five compounds constitute examples of sterically imposed non-stoichiometry (i) occurring and (ii) being a mechanism of accommodation of oversize ions.

References

JONES, A. F. & MISELL, D. L. (1970). *J. Phys. A: Math. Gen.* **3**, 462–472.

LOMER, T. R. & WILSON, A. J. C. (1975). *Acta Cryst.* **B31**, 646–648.

SHANNON, R. D. (1976). *Acta Cryst.* **A32**, 751–767.

SUDARSANAN, K., WILSON, A. J. C. & YOUNG, R. A. (1972). *Acta Cryst.* **A28**, S151.

SUDARSANAN, K., YOUNG, R. A. & DONNAY, J. D. H. (1973). *Acta Cryst.* **B29**, 808–814.

SUDARSANAN, K., YOUNG, R. A. & WILSON, A. J. C. (1977). *Acta Cryst.* **B33**, 3136–3142.

WILSON, A. J. C. (1974). Paper presented at the IUCr Conf. on Anomalous Scattering, Madrid, 25 April 1974; included in *Anomalous Scattering* (1975), edited by S. RAMASESHAN & S. C. ABRAHAMS, pp. 325–332. Copenhagen: Munksgaard.

WILSON, A. J. C. (1976a). *Acta Cryst.* **A32**, 781–783.

WILSON, A. J. C. (1976b). *Acta Cryst.* **A32**, 994–996.