Clays and Clay Minerals, Vol. 34, No. 2, 224–226, 1986.

NOTE

RECURSIVE METHOD FOR DETERMINING FREQUENCY FACTORS IN INTERSTRATIFIED CLAY DIFFRACTION CALCULATIONS

Key Words—Frequency factor, Interstratification, Recursive algorithm, Stacking sequence, X-ray powder diffraction.

The Fourier series technique of predicting X-ray diffraction patterns for interstratified clay minerals requires the calculation of a matrix of frequency factors, σ (Reynolds, 1980). In considering a clay composed of stacking sequences of two types of layers, designated A and B, the elements of σ give the probabilities of finding AA, AB, and BB layer pairs at each possible separation distance in all possible sequences or arrays. When the matrix is determined by considering each possible sequence individually, the computation time required increases exponentially with the number of layers in the longest sequences, N_{max}, and quickly dominates the time required for the diffraction calculation. The exponential nature of this increase is due to the great number of possible arrangements of A and B layers in even moderately long sequences (2^N) and has effectively limited the modeling of diffraction from interstratified clays to crystallites of $\leq$20 layers in thickness.

We suggest a recursive algorithm for calculating frequency factors. The method is so-named because the factors for a given separation are taken to be functionally dependent on more easily calculated factors for smaller separations, leaving σ elements on both sides of the derived equalities. The computational advantage is gained by treating groups of permutations of the clay layers at once instead of treating each possible sequence individually. The effort required in the calculation only increases linearly with the number of layers in the sequences, N, allowing the consideration of larger crystallites and the use of smaller computers than was previously possible. Further, the results are less likely to be affected by the numerical error introduced during the serial addition of floating-point numbers of varying magnitudes, a particular problem on computers with small word lengths (see Sterbenz, 1974).

DISCUSSION

The frequency factor matrix is divided into elements based on the types of layers in the pair considered and the separation of these layers, given by the numbers of A and B layers in the sequences:

$$\sigma_{AA,m,n} = \sigma_{A\rightarrow A,m,n},$$

$$\sigma_{AB,m,n} = \sigma_{A\rightarrow B,m,n} + \sigma_{B\rightarrow A,m,n},$$

$$\sigma_{BB,m,n} = \sigma_{B\rightarrow B,m,n}.$$

Here, m is the number of A layers in the sequences; n is the number of B layers; and AA, AB, and BB represent all sequences with an A on each end, an A on one end and a B on the other, and a B on each end. The directional terms A $\rightarrow$ A, A $\rightarrow$ B, etc., represent the sequences which begin with the first layer and end with the second. Both terms, $\sigma_{A\rightarrow B}$ and $\sigma_{B\rightarrow A}$, which comprise σ_{AB} are equal by symmetry.

A straightforward method of calculating σ which is in common use today (informally circulated MOD-4 and later programs of Reynolds) is to generate each possible sequence of A and B layers for $N \leq N_{max}$, compute the probability of the sequence occurring, determine which element in σ corresponds to the sequence, and add the probability to the value of that element. The probability of occurrence of a sequence is assumed to be the product of the appropriate layer fractions P_A and P_B and junction probabilities $P_{A.A}$, $P_{A.B}$, etc. (Reynolds, 1980, 252–255). Unfortunately, the relative computational effort required for each value of N in this process is approximately $(N-1)2^N$, and a practical limit to N_{max} is quickly reached. Although various hybrid forms of this algorithm which we have developed significantly improve efficiency, each retains an exponential increase in calculation time with N and some require large amounts of additional memory.

The recursive algorithm calculates the matrix elements for (m + n) = N from the elements in which (m + n) = (N − 1). Inasmuch as all possible sequences with AA terminations, m A layers, and n B layers may be constructed by adding an A layer to the end of each A $\rightarrow$ A and A $\rightarrow$ B sequence of (m − 1) A and n B layers,

$$\sigma_{AA,m,n} = \sigma_{A\rightarrow A,m-1,n}P_{A.A} + \sigma_{A\rightarrow B,m-1,n}P_{B.A}$$
$$= \sigma_{AA,m-1,n}P_{A.A} + \tfrac{1}{2}\sigma_{AB,m-1,n}P_{B.A},$$

assuming Reichweite 1 (or nearest-neighbor) ordering. Similarly,

$$\sigma_{AB,m,n} = \sigma_{AA,m,n-1}P_{A.B} + \tfrac{1}{2}\sigma_{AB,m,n-1}P_{B.B}$$
$$+ \sigma_{BB,m-1,n}P_{B.A} + \tfrac{1}{2}\sigma_{AB,m-1,n}P_{A.A},$$

and

$$\sigma_{BB,m,n} = \sigma_{BB,m,n-1}P_{B.B} + \tfrac{1}{2}\sigma_{AB,m,n-1}P_{A.B}.$$

https://doi.org/10.1346/CCMN.1986.0340213 Published online by Cambridge University Press

By seeding the matrix elements in which N = 2 at

$$\sigma_{AA,2,0} = P_A P_{A.A},$$

$$\sigma_{AB,1,1} = P_A P_{A.B} + P_B P_{B.A},$$

and

$$\sigma_{BB,0,2} = P_B P_{B.B},$$

(and taking all elements with negative m or n subscripts to be zero), σ is easily filled by computing the elements for progressively greater N using the recursion formulae.

Frequency factors for Reichweite 2 (next-nearest neighbor) ordering are also easily calculated, although directional values must be stored for eight termination types for each possible m at the previous value of N (n is fixed by m at constant N). In Reichweite 2 sequences, the eight directional terms $\sigma_{A\to AA}$, $\sigma_{A\to AB}$, . . . , $\sigma_{B\to BB}$ are calculated from binary junction probabilities as

$$\sigma_{A\to AA,m,n} = \sigma_{A\to AA,m-1,n} P_{AA.A} + \sigma_{A\to BA,m-1,n} P_{BA.A},$$

$$\sigma_{A\to AB,m,n} = \sigma_{A\to AA,m,n-1} P_{AA.B} + \sigma_{A\to BA,m,n-1} P_{BA.B},$$

$$\vdots$$

$$\sigma_{B\to BB,m,n} = \sigma_{B\to AB,m,n-1} P_{AB.B} + \sigma_{B\to BB,m,n-1} P_{BB.B},$$

for (m + n) = N. Once the directional values for N are known, the frequency factors are formed by summing these terms according to their terminating layers:

$$\sigma_{AA,m,n} = \sigma_{A\to AA,m,n} + \sigma_{A\to BA,m,n},$$

$$\sigma_{AB,m,n} = \sigma_{A\to AB,m,n} + \sigma_{A\to BB,m,n} + \sigma_{B\to AA,m,n} + \sigma_{B\to BA,m,n},$$

and

$$\sigma_{BB,m,n} = \sigma_{B\to AB,m,n} + \sigma_{B\to BB,m,n}.$$

The terms for this N may now be used to calculate new directional terms for (m + n) = (N + 1).

Similarly, frequency factors for Reichweite 3 ordering are found by first calculating sixteen directional terms $\sigma_{A\to AAA}$, $\sigma_{A\to AAB}$, . . . , $\sigma_{B\to BBB}$ for each m as

$$\sigma_{A\to AAA,m,n} = \sigma_{A\to AAA,m-1,n} P_{AAA.A} + \sigma_{A\to BAA,m-1,n} P_{BAA.A},$$

$$\sigma_{A\to AAB,m,n} = \sigma_{A\to AAA,m,n-1} P_{AAA.B} + \sigma_{A\to BAA,m,n-1} P_{BAA.B},$$

$$\vdots$$

$$\sigma_{B\to BBB,m,n} = \sigma_{B\to ABB,m,n-1} P_{ABB.B} + \sigma_{B\to BBB,m,n-1} P_{BBB.B},$$

and then, as before, summing the intermediate terms according to their terminations to form the frequency factors σ_{AA}, σ_{AB}, and σ_{BB}.

Figure 1 shows the actual computer time required to determine σ for different values of N_{max} by the "straightforward" method in common use today, a "hybrid" of this algorithm, and the recursive algorithm. All tests were made on the University of Illinois CDC CYBER 175 computer using the FTN5 compiler at optimization level 2. A linear extrapolation of the data in semilog coordinates suggests that the first algorithm might take 300,000 years of computer time to determine σ for N_{max} = 50 layers, and the second might need 100 years; the recursive formulation finished in 0.01 seconds.

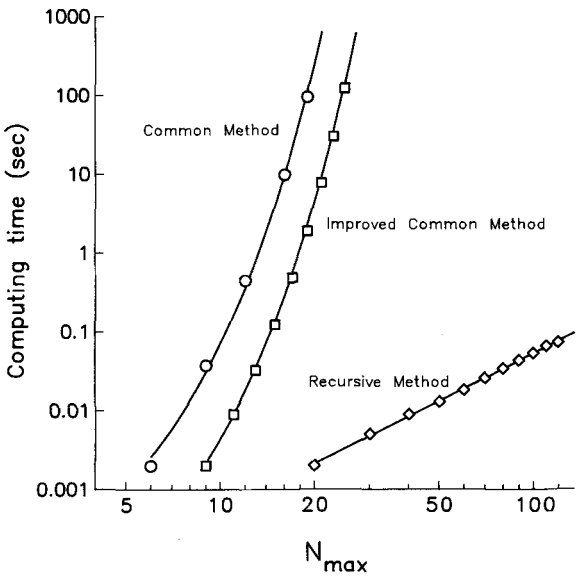

Figure 1. Log-log plot and semilog regressions of the computing time used by the common method of determining frequency factors (circles), an improvement on this method (squares), and the recursive formulation (diamonds) as a function of maximum sequence length.

Little work has been reported on three component systems, probably because of the very long run-times involved. The algorithm described above is easily modified so that it is appropriate for such calculations, and its efficiency compared to a 3^N approach should be even greater than the examples given here.

ACKNOWLEDGMENTS

We thank the University of Illinois for access to its computer, Karolyn Roberts for typing the manuscript, and D. R. Pevear for providing a FORTRAN version of the MOD-4 diffraction program. The manuscript benefited from thoughtful reviews by D. L. Bish, D. M. C. MacEwan, and D. R. Pevear.

Department of Geology CRAIG M. BETHKE
University of Illinois
Urbana, Illinois 61801

Department of Earth Sciences ROBERT C. REYNOLDS
Dartmouth College
Hanover, New Hampshire 03755

REFERENCES

Reynolds, R. C. (1980) Interstratified clay minerals: in *Crystal Structures of Clay Minerals and Their X-ray Identification,* G. W. Brindley and G. Brown, eds., Mineralogical Society, London, 249–303.

Sterbenz, P. H. (1974) *Floating-Point Computation:* Prentice Hall, Englewood, New Jersey, 316 pp.

(Received 18 April 1984; accepted 11 November 1984; Ms. 1354)