Efficient Computation of Polymer Conformation Energy

In Monte Carlo studies of polymer properties [1–3], the nonbonded contribution to the total conformational energy of a polymer is ordinarily taken as

$$U_{\mathrm{nb}} = \sum_{i=1}^{N-2} \sum_{j=i+2}^{N} f(r_{ij}^2), \tag{1}$$

where U_{nb} is the conformational energy due to nonbonded interactions, r_{ij}^2 is the squared distance between atoms i and j, N is the number of atoms in the chain, and the summation extends over all nonbonded atomic pairs in the polymer; $f(r_{ij}^2)$ is the potential energy contribution of atoms i and j as a function of r_{ij}^2.

No matter what functional form is assumed for $f(r_{ij}^2)$ [4–6] [except for the trivial case $f \equiv 0$, ignoring excluded volume effects entirely], the sum in Eq. (1) requires the function f to be evaluated approximately $\frac{1}{2}N^2$ times for each sample polymer examined. Since the conformational energy of each polymer generated must be evaluated regardless of the sampling technique used, the total computational time increases roughly as the square of the number of backbone atoms. This situation would easily put the most interesting lengths out of range of present-day computers, if not for the following stratagem [7].

Replace the function f by an approximating function $\hat{f}$, and define a cutoff distance ρ such that

$$\hat{f}(r_{ij}^2) = 0, \qquad r_{ij}^2 \geq \rho^2,$$
$$\hat{f}(r_{ij}^2) = f(r_{ij}^2), \qquad r_{ij}^2 < \rho^2. \tag{2}$$

ρ is a natural parameter for hard-sphere potential energy models. For a continuous function, such as a Lennard–Jones potential, ρ may be defined as the separation beyond which the energy of interaction between two atoms is negligible. A "negligible" energy in this case should be chosen so that the sum of all such neglected contributions in the molecule does not exceed about 5–10% of the average observed U_{nb}.

Let a be the fixed bond length for the backbone chain, as in Fig. 1. Then in the case $r_{ij}^2 \geq \rho^2$, the increment k to be added to j (for fixed i) to find the index j' of the next atom along the chain where any interaction with atom i is possible is

$$k = \left[\frac{r_{ij} - \rho}{a} \right], \tag{3}$$

where $[x]$ denotes the greatest integer not exceeding x. In other words, if one has just determined that no interaction exists for the current values of i and j, then use of

$$j' = j + k, \tag{4}$$

with k given by Eq. (3), determines the second index of the next pair, ij', which must be examined. All intervening pairs may be skipped as geometrically excluded. This is illustrated in Fig. 1.
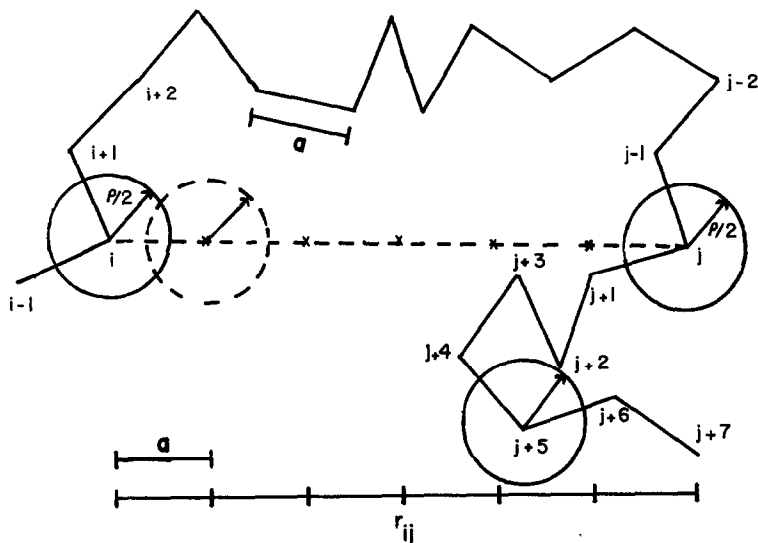

FIG. 1. Example of a random-walk polymer in two dimensions, with step length a and interaction parameter ρ, chosen here to be $\frac{3}{2}a$. Atoms i and j are separated by $6a > \rho$, and hence do not interact. If by unlikely chance the next six steps after j lie along the line connecting it with i, the next possible interaction would not occur until $j' = j + k$, where, by Eq. (6), $k \leqslant 6$. Hence, by taking $k = 5$ all intervening pairs which are geometrically excluded are skipped. The unlikely worst case of direct approach is shown with crosses representing atomic positions and the nearest possible overlap with atom i shown as a broken circle. Other apparent overlaps have not yet been detected, since they await incrementing of i.

To calculate k, we begin with the easily proved inequality

$$\left[\frac{r_{ij} - \rho}{a}\right] \leqslant \left[\frac{r_{ij}}{a}\right] - \left[\frac{\rho}{a}\right] + 1, \tag{5}$$

thus obtaining

$$k \leqslant k' = [\sqrt{r_{ij}^2 t^2}] - C, \tag{6}$$

where $t^2 = a^{-2}$ and $C = [\rho/a] - 1$. C and t are constants for a particular Monte Carlo experiment. We use t^2 instead of a^{-2} because computer multiplication is always faster than division.

A function subprogram to evaluate the greatest integer in the square root of a number was written in order to circumvent the slower machine-supplied square root routine. For the sake of even faster code, we allow the integer part of the square root to be slightly underestimated. [Since, in Eq. (6), k' can never over-estimate k by more than unity, it is always safe to add $(k' - 1)$, rather than k' or k, to j; this allowance may slightly lengthen the search, but never invalidates it.] A table is constructed beforehand containing the square roots of the first 1000 integers, along with a subsidiary table containing the first five powers of 10. If the argument, $r_{ij}^2 t^2$, exceeds 1000, it is multiplied successively by 0.01 until the result is less than 1000, note being taken of the number, n, of multiplications needed. The result is then truncated to an integer, which is used as the index for the table look-up. The table entry is then multiplied by 10^n, this factor being brought from the second table. The entire subprogram, coded in COMPASS (CDC assembly language), occupies seven words of memory in the CDC 6600, exclusive of the tables. In a time test, the code proved to be 40 % faster than the machine-supplied square root routine. Usual underestimation was by one integer, and never exceeded two.
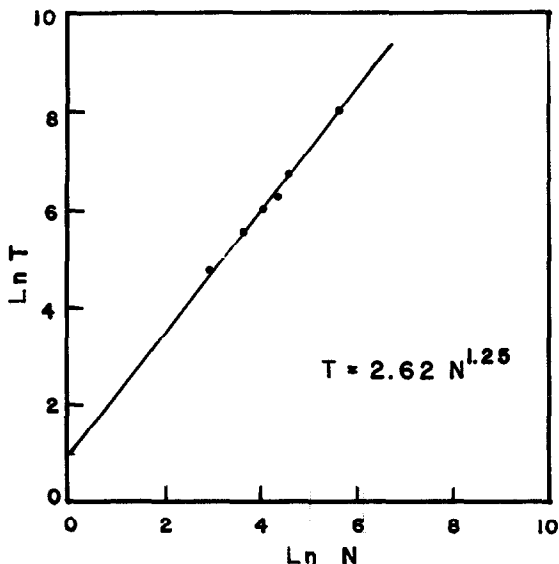

FIG. 2. Log–log plot of running time, T, for generation and analysis of 10,000 polymer samples of chain length N, vs N.

The efficiency imparted to the entire Monte Carlo program by this technique is difficult to estimate, but the overall efficiency may be appreciated by reference to Fig. 2, in which the natural logarithm of the running time, T, for generation and analysis of 10,000 samples by the method of Stellman and Gans [1] is plotted against ln N. The line is quite straight (standard deviation of the fitted slope is 0.03), and yields the relationship

$$T = \alpha N^{\beta}, \tag{7}$$

where $\alpha = 2.62$ and $\beta = 1.25$. Thus, in spite of many complicating factors, the time rises only as the 5/4 power of N, rather than with the square. This development offers much promise for future studies.

REFERENCES

1. S. D. STELLMAN AND P. J. GANS, to be published.
2. F. T. WALL, S. WINDWER, AND P. J. GANS, "Methods in Computational Physics," Vol. 1, Academic Press, New York, 1963.
3. J. MAZUR AND F. L. MCCRACKIN, J. Chem. Phys. 49 (1968), 648.
4. F. T. WALL AND J. MAZUR, Ann. N. Y. Acad. Sci. 89 (1961), 608.
5. R. A. SCOTT AND H. A. SCHERAGA, Biopolymers 4 (1966), 237; J. Chem. Phys. 44 (1966), 3054.
6. A. ABE, R. L. JERNIGAN, AND P. J. FLORY, J. Amer. Chem. Soc. 88 (1966), 631.
7. Other loop-checking devices previously described include a doubly-linked list method, P. J. GANS, J. Chem. Phys. 42 (1965), 4159, and a space-cubing method, C. LEVINTHAL, Sci. Amer. 214 (6), (1966), 42.

STEVEN D. STELLMAN, MARK FROIMOWITZ, AND PAUL J. GANS

Department of Chemistry,
New York University,
New York, New York 10003