

APPENDIX A

R-FACTOR FOR UNCORRELATED STRUCTURE FACTORS

We wish to calculate the R-factor for the comparison of two sets of structure factors that are unrelated to each other. We will assume that they are correctly scaled together, i.e. that they are taken from identical probability distributions.

To obtain the ideal distribution of structure factor amplitudes, consider a Shannon sampled random structure in which the density at every sample point is uncorrelated with that at every other. The calculation of a non-centric structure factor from this structure is essentially a random walk in the complex plane. Thus the real and imaginary parts of the structure factor are Gaussian distributed. The distribution function for amplitudes is therefore of the form,

$$P(F) dF \propto F e^{-F^2/2\sigma^2} dF$$

We can simplify the calculation by normalising the structure factors to $f = F/\sigma$, whence the normalised distribution is

$$P(f) df = f e^{-f^2/2} df$$

The joint probability distribution for two sets of structure

factors on the same scale is the product of two such distributions,

$$P'(f_1, f_2) = P(f_1)P(f_2) = f_1 f_2 e^{-f_1^2/2} e^{-f_2^2/2}$$

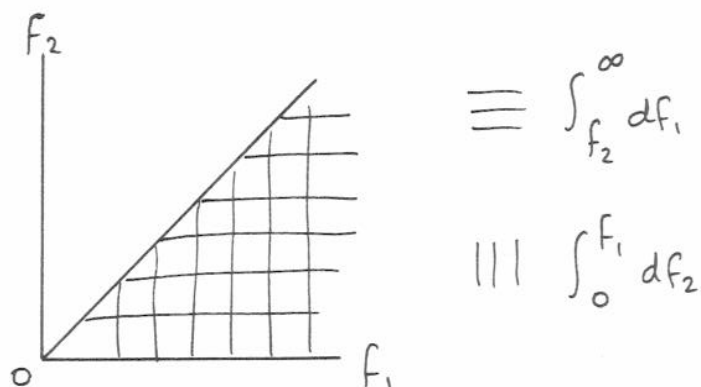
The expression for the R-factor is

$$R = \frac{R_{\text{num}}}{R_{\text{denom}}} = \frac{\int_0^{\infty} df_1 \int_0^{\infty} df_2 |f_1 - f_2| f_1 e^{-f_1^2/2} f_2 e^{-f_2^2/2}}{\int_0^{\infty} f_1^2 e^{-f_1^2/2} df_1}$$

The expression for the numerator can be expanded as

$$R_{\text{num}} = \int_0^{\infty} df_1 \left\{ \int_0^{f_1} (f_1 - f_2) df_2 + \int_{f_1}^{\infty} (f_2 - f_1) df_2 \right\} f_1 f_2 e^{-(f_1^2 + f_2^2)/2}$$

The integrals over $\int_0^{\infty} df_1 \int_0^{f_1} df_2$ and over $\int_{f_2}^{\infty} df_1 \int_0^{\infty} df_2$ correspond to the same region of $f_1 f_2$ space with a different order of integration.



Thus the four integrals in the numerator consist of two identical pairs:

$$R_{\text{num}} = 2 \int_0^{\infty} df_1 f_1^2 e^{-f_1^2/2} \left\{ \int_0^{f_1} df_2 - \int_{f_1}^{\infty} df_2 \right\} f_2 e^{-f_2^2/2}$$

The integral over f_2 is an analytic form:

$$R_{num} = 2 \int_0^{\infty} df_1 f_1^2 e^{-f_1^2/2} \left\{ -e^{-f_2^2/2} \Big|_0^{f_1} + e^{-f_2^2/2} \Big|_{f_1}^{\infty} \right\}$$

$$= 2 \int_0^{\infty} df_1 f_1^2 e^{-f_1^2/2} \left[1 - e^{-f_1^2/2} - e^{-f_1^2/2} + 0 \right]$$

Using $\int_0^{\infty} x^2 e^{-2x^2} dx = \int_0^{\infty} \frac{y^2}{2} e^{-y^2} dy / \sqrt{2}$

$$R_{num} = 2 \int_0^{\infty} df_1 f_1^2 e^{-f_1^2/2} \left[1 - 2 \left(\frac{1}{2\sqrt{2}} \right) \right]$$

$$= (2 - \sqrt{2}) R_{denom}$$

$$\Rightarrow R = 2 - \sqrt{2} \quad \text{Q.E.D.}$$

BIBLIOGRAPHY

- "International Tables of Crystallography", Kynoch Press, Birmingham (1965)
- Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, A. G. W., Rayment, I., Rossmann, M. G., Suck, D., Tsukihara, T., *Nature* 286 33 (1980)
- Adolph, K. W., Butler, P. J. G., *J. Mol. Biol.* 88 327 (1974)
- Agarwal, R. C., *Acta C.* A34 791 (1978)
- Arndt, U. W., Wonacott, A. J., "The Rotation Method in Crystallography", North Holland (1977)
- Baldwin, R. L., *Ann. Rev. Biochem.* 453 (1975)
- Bancroft, J. B., Hills, G. J., Markham, R., *Virology* 31 354 (1967)
- Bancroft, J. B., *Adv. in Virus Res.* 16 99 (1970)
- Bernal, J. D., Fankuchen, I., Riley, D. P., *Nature* 142 1075 (1938)
- Blundell, T., Johnson, L., "Protein Crystallography", Academic Press (1976)
- Bricogne, G., *Acta C.* A30 395 (1974)
- Bricogne, G., "The Joy of Skewing", unpublished (1975)
- Bricogne, G., *Acta C.* A32 832 (1976)
- Bricogne, G., In "Report of CECAM Workshop on Virus Crystallography", Centre Europeen de Calcul Atomique et Moleculaire, Orsay, Paris (1977)
- Buerger, M. J., "The Precession Method", J. Wiley (1964)
- Casjens, S., King, J., *Ann. Rev. Biochem.* 44 555 (1975)
- Caspar, D. L. D., *Nature* 177 475 (1956)
- Caspar, D. L. D., Klug, A., *Cold Spring Harbor Symp. Quant. Biol.* 27 1 (1962)
- Caspar, D. L. D., *Adv. in Protein Chem.* 18 37 (1963)

- Chauvin, C., Witz, J., Jacrot, B., J. Mol. Biol. 124 641 (1978)
- Chauvin, C. Pfeiffer, P., Witz, J., Jacrot, B., Virology 88 138 (1978A)
- Cohn, E. J., Edsall, J. T., "Proteins, Amino Acids and Peptides", Hafner Press (1943)
- Cooley, J. W., Tukey, J. W., Math. of Computation 19 297 (1965)
- Crawford, J. L., Thesis, Harvard U. (1977)
- Crick, F. H. C., Watson, J. D., Nature 177 473 (1956)
- Crowther, R. A., Acta C. 22 758 (1967)
- Crowther, R. A., Acta C. B25 2571 (1969)
- Debye, P., Annal. Physik 43 49 (1914)
- Debye, P., Z. physik. Chem. 130 56 (1927)
- Dickerson, R. E., Kendrew, J. C., Strandberg, B. E., Acta C. 14 1188 (1961)
- Dickerson, R. E., Kopka, M. L., Varnum, J. C., Weinzierl, J. E., Acta C. 23 511 (1967)
- Dickerson, R. E., Weinzierl, J. E., Palmer, R. A., Acta C. B24 997 (1968)
- Durham, A. C. H., Hendry, D. A., von Wechmar, M. B., Virology 77 524 (1977)
- Falicov, L. M., "Group Theory and its Physical Applications", U. Chicago Press (1967)
- Finch, J. T., Klug, A., Leberman, R., J. Mol. Biol. 50 215 (1970)
- Fox, G. C., Holmes, K. C., Acta C. 20 886 (1966)
- Golden, J. S., Harrison, S. C., to be submitted (1981)
- Guinier, A., "X-ray diffraction in Crystals, Imperfect Crystals and Amorphous Bodies", W. H. Freeman (1963)
- Harrison, B. D., Finch, J. T., Gibbs, A. J., Hollings, M., Shepherd, R. J., Valenta, V., Wetter, C., Virology 45 356 (1971)
- Harrison, S. C., J. Appl. Cryst. 1 84 (1968)

- Harrison, S. C., J. Mol. Biol. 42 457 (1969)
- Harrison, S. C., Jack, A., J. Mol. Biol. 97 173 (1975)
- Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K., Bricogne, G., Nature 276 368 (1978)
- Harrison, S. C., Biophys J. 32 139 (1980)
- Hodgkin, D. C., Cold Spring Harbor Symp. Quant. Biol. 14 65 (1950)
- Hogle, J. M., Harrison, S. C., to be submitted (1981)
- Horne, P., Wildy, R. W., Virology 15 348 (1961)
- Incardonna, N. L., Kaesberg, P., Biophys J. 4 11 (1964)
- Jack, A., Harrison, S. C., Crowther, R. A., J. Mol. Biol. 97 163 (1975)
- Jacrot, B., J. Mol. Biol. 95 433 (1975)
- Jacrot, B., Chauvin, C., Witz, J., Nature 266 417 (1977)
- Kam, Z., Shore, H. B., Feher, G., J. Mol. Biol. 123 539 (1978)
- Leberman, R., Finch, J. T., J. Mol. Biol. 50 209 (1970)
- Lipscomb, W. N., Coppola, J. C., Hartsuck, J. A., Ludwig, M. L., Muirhead, H., Searl, J., Steitz, T. A., J. Mol. Biol. 19 423 (1966)
- Luftig, R., J. Ultrastructure Res. 20 91 (1967)
- McPherson, A., Meth. Biochem. Analysis 23 249 (1975)
- McPherson, A., J. Biol. Chem. 251 6300 (1976)
- Monod, J., Wyman, J., Changeux, J-P., J. Mol. Biol. 12 88 (1965)
- Munowitz, M. G., Dobson, C. M., Griffin, R. G., Harrison, S. C., J. Mol. Biol. 141 327 (1980)
- Perutz, M. F., Nature 228 726 (1970)
- Pfeiffer, P., Hirth, L., Virology 61 160 (1974)
- Pfeiffer, P., Hirth, L., FEBS Lett. 56 144 (1975)
- Richards, F. M., J. Mol. Biol. 37 225 (1968)

- Richardson, J. S., Nature 268 495 (1977)
- Schutt, C. E., Thesis, Harvard U. (1976)
- Shannon, C. E., Proc. Inst. Radio Engineers, New York, 37 10
(1949)
- Stryer, L., Kendrew, J. C., Watson, H. C., J. Mol. Biol. 8
96 (1964)
- Ten Eyck, L. F., Acta C. A29 183 (1973)
- Unge, T., Lijias, L., Strandberg, B. E., Vaara, I., Kannan,
K. K., Fridborg, K., Nordman, C. E., Lentz, P. J.,
Nature 285 373 (1980)
- Williams, R. C., Backus, R. C., J. Amer. Chem. Soc. 71 4052
(1949)
- Wilson, A. J. C., Acta C. 2 328 (1949)
- Winkler, F. K., Schutt, C. E., Harrison, S. C., Bricogne,
G., Nature 265 509 (1977)
- Winkler, F. K., Schutt, C. E., Harrison, S. C., Acta C. A35
901 (1979)
- Ziegler, A., Harrison, S. C., Leberman, R., Virology 59 509
(1974)