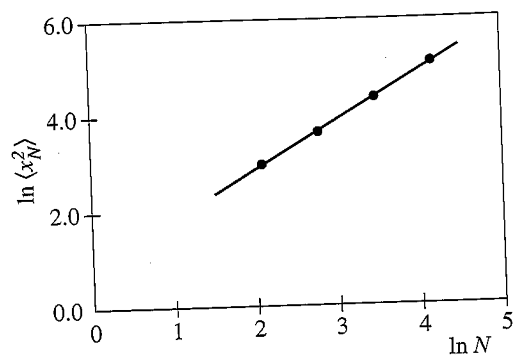

Figure 7.5 Plot of $\ln \Delta x^2$ versus $\ln N$ for the data listed in Table 7.2. The straight line $y = 1.02x + 0.83$ through the points is found by minimizing the sum (7.34).

nonlinear can be turned into a linear problem by a change of variables. In this case we convert the nonlinear relation (7.40) to a linear relation by taking the logarithm of both sides:

$$\ln(\Delta x^2) = \ln a + 2\nu \ln N. \tag{7.41}$$

The values of $y = \ln(\Delta x^2)$ and $x = \ln N$ in Table 7.2 and the least squares fit are shown in Figure 7.5. We use (7.39) and find that $m = 1.02$ and $b = 0.83$. Hence, we conclude from our limited data and the relation $2\nu = m$ that $\nu \approx 0.51$, which is consistent with the expected result $\nu = 1/2$.

The least squares fitting procedure also allows us to estimate the uncertainty or the most probable error in m and b by analyzing the measurements themselves. The result of this analysis is that the most probable error in m and b, σ_m and σ_b, respectively, is given by

$$\sigma_m = \frac{1}{\sqrt{n}}\frac{\Delta}{\sigma_x} \tag{7.42a}$$

$$\sigma_b = \frac{1}{\sqrt{n}}\frac{(\langle x^2\rangle)^{1/2}\Delta}{\sigma_x}, \tag{7.42b}$$

where

$$\Delta^2 = \frac{1}{n-2}\sum_{i=1}^{n} d_i^2, \tag{7.43}$$

and d_i is given by (7.33).

Because there are n data points, we might have guessed that n rather than $n-2$ would be present in the denominator of (7.43). The reason for the factor of $n-2$ is related to the fact that to determine Δ, we first need to calculate *two* quantities m and b, leaving only $n-2$ independent degrees of freedom. To see that the $n-2$ factor is reasonable, consider the special case of $n = 2$. In this case we can find a line that passes exactly through the two data points, but we cannot deduce anything about the reliability of the set of measurements because the fit is exact. If we use (7.43), we see that both the numerator and denominator would be zero, and hence Δ would be undetermined. If a factor of n rather than $n-2$ appeared in (7.43), we would conclude that $\Delta = 0/2 = 0$, an absurd conclusion. Usually $n \gg 1$, and the difference between n and $n-2$ is negligible.

For our example, $\Delta = 0.03$, $\sigma_b = 0.07$, and $\sigma_m = 0.02$. The uncertainties δm and $\delta\nu$ are related by $2\delta\nu = \delta m$. Because $\delta m = \sigma_m$, we conclude that our best estimate for ν is $\nu = 0.51 \pm 0.01$.

If the values of y_i have different uncertainties σ_i, then the data points are weighted by the quantity $w_i = 1/\sigma_i^2$. In this case it is reasonable to minimize the quantity

$$\chi^2 = \sum_{i=1}^{n} w_i(y_i - mx_i - b)^2. \tag{7.44}$$

The resulting expressions in (7.39) for m and b are unchanged if we generalize the definition of the averages to be

$$\langle f\rangle = \frac{1}{n\langle w\rangle}\sum_{i=1}^{n} w_i f_i, \tag{7.45}$$

where

$$\langle w\rangle = \frac{1}{n}\sum_{i=1}^{n} w_i. \tag{7.46}$$

Problem 7.27 Example of least squares fit

(a) Write a program to find the least squares fit for a set of data. As a check on your program, compute the most probable values of m and b for the data shown in Table 7.2.
(b) Modify the random walk program so that steps of length 1 and 2 are taken with equal probability. Use at least 10,000 trials and do a least squares fit to Δx^2 as done in the text. Is your most probable estimate for ν closer to $\nu = 1/2$? ■

For simple random walk problems the relation $\Delta x^2 = aN^\nu$ holds for all N. However, in many random walk problems, a power law relation between Δx^2 and N holds only asymptotically for large N, and hence we should use only the larger values of N to estimate the slope. Also, because we are finding the best fit for the logarithm of the independent variable N, we need to give equal weight to all intervals of $\ln N$. In the above example, we used $N = 8$, 16, 32, and 64, so that the values of $\ln N$ are equally spaced.

7.7 ■ APPLICATIONS TO POLYMERS

Random walk models play an important role in polymer physics (cf. de Gennes). A polymer consists of N repeat units (monomers) with $N \gg 1$ ($N \sim 10^3 - 10^5$). For example, polyethylene can be represented as $\cdots -CH_2-CH_2-CH_2-\cdots$. The detailed structure of the polymer is important for many practical applications. For example, if we wish to improve the fabrication of rubber, a good understanding of the local motions of the monomers