5

Determination of the Mass-Ratio Distribution, II: Double-lined Spectroscopic Binary Stars

Submitted to: *Astrophys. Space Sci.*

Abstract

The application of the Richardson-Lucy iterative technique for the rectification of measurement errors, to the mass-ratio distribution of the double-lined spectroscopic binary stars (SBII) in the *Sixth Catalogue of the Orbital Elements of Spectroscopic Binary Stars*, is evaluated for the SBII systems in the *Eighth* version of the Catalogue.

The estimates of the real distribution, produced by the Richardson-Lucy method after successive iterations, do not converge in a trivial way. The stop criterion, suggested in the original paper on the method, is used in the application of the method to both real and simulated distributions. It is shown that, in the application to the mass-ratio distribution of the SBII systems, the iterations should stop considerably earlier than was previously assumed.

The enhanced peak at mass-ratios near unity, that was found to exist in the real mass-ratio distribution, based on the earlier application of the iterative technique to the SBII systems, is likely to be due to an overshoot effect that occurs when the iterations are carried too far. When the stop criterion is used in the application of the method to the SBII systems in the *Eighth* version of the Catalogue, the resulting estimate of the real mass-ratio distribution does not show evidence for an enhanced number of binary stars with mass ratios near unity.

Key words: spectroscopic binary stars, mass-ratio distribution, restoration, Richardson-Lucy method

5.1 Introduction

As part of a study of the mass-ratio distribution of spectroscopic binary stars (Hogeveen 1991, this thesis Chapter 3), the methods applied by Lucy and Ricco (1979, hereafter LR79) to determine the real mass-ratio distribution of double-lined spectroscopic binary stars (SBII) were investigated. Lucy and Ricco apply an iterative technique, as developed by Lucy (1974, hereafter L74), to correct the observed mass-ratio distribution of 173 SBII systems in the *Sixth Catalogue of the Orbital Elements of Spectroscopic Binary Stars* (DAO6) (Batten 1967) for measurement errors. The method is also used for image restoration (Adorf 1990), and is presently called the Richardson-Lucy method (hereafter R-L method), after the authors who independently devised it in 1972 and 1974.

In section 5.2, the R-L method is described in brief. In section 5.3 the criterion to stop the iterations, suggested in L74, is outlined. The stop criterion is tested in sections 5.4 and 5.6, in the application of the method to a sample of 237 SBII systems from the *Eighth Catalogue of the Orbital Elements of Spectroscopic Binary Stars* (DAO8) (Batten *et al.* 1989), and to a synthetic sample consisting of the same number of elements with about the same distribution. The estimates of the real distributions are, for both samples, evaluated for different (artificial) distributions of errors. In section 5.7 some aspects of the R-L method, and the results obtained with it, are discussed. In section 5.8 the consequences of the results for our understanding of the mass-ratios distribution of SBII systems are addressed.

5.2 The Richardson-Lucy method

An observed distribution $\phi(x)$ may be characterized by:

$$\phi(x) = \int \Psi(\xi) P(x|\xi) d\xi, \tag{5.1}$$

where $\Psi(\xi)$ is the real distribution, ξ the exact value of the observed parameter, x the actually observed value, and $P(x|\xi)d\xi$ the probability that the observed value x' falls within $(x, x+dx)$ when the exact value is ξ'. $P(x|\xi)$ will often represent the function by which the data are smeared when they are measured with an imperfect instrument.

In L74 it is assumed that the probability for the nth member of a set of observations to have a value $x \in (x, x+dx)$ is represented by:

$$P_n(x|\xi) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(-\frac{(x-\xi)^2}{2\sigma_n^2}\right) dx, \tag{5.2}$$

where ξ is the (unknown) *exact* value of the observed parameter. If the variances σ_n^2 are known, a sequence of estimates $\Psi^r(\xi)$ of the error-free distribution $\Psi(\xi)$ may be obtained from:

$$\Psi^{(r+1)}(\xi) = \frac{1}{N} \sum_n Q_n^r(\xi|x_n), \tag{5.3}$$

where

$$Q_n^r(\xi|x_n) = \frac{\Psi^r(\xi) P_n(x_n|\xi)}{\phi^r(x_n)} \tag{5.4}$$

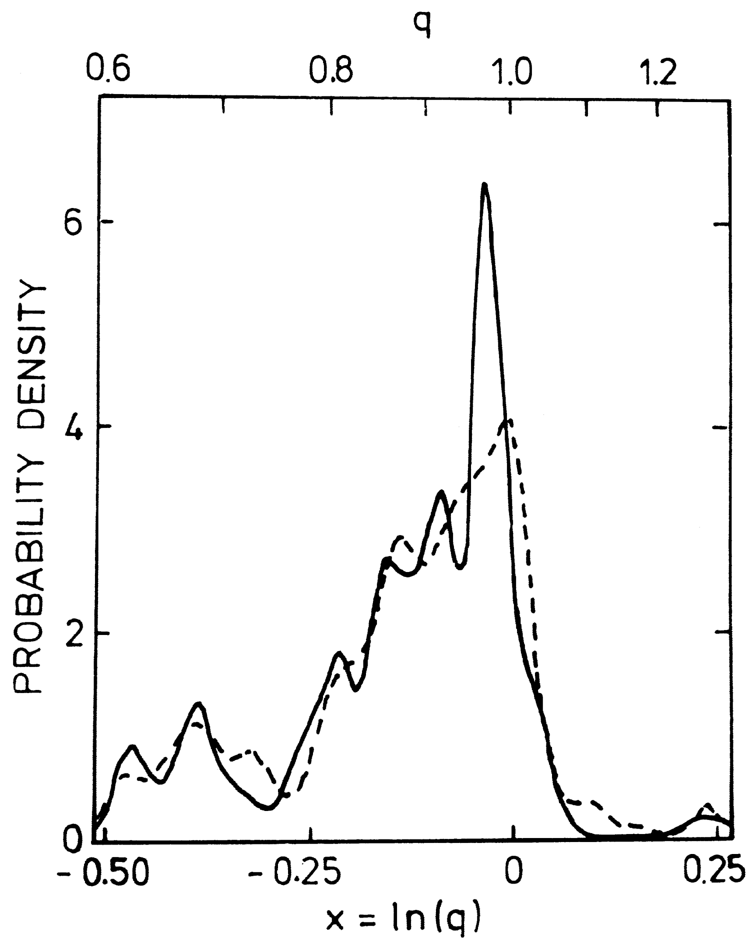

Figure 5.1: Dashed line: smoothed histogram of $\phi(x)$ ($x = \ln q$) for the SBII systems in DAO6, based on raw data, as presented by Lucy and Ricco (1979). Solid line: distribution $\Psi^5(\xi)$ obtained by Lucy and Ricco after applying the iterative technique.

and

$$\phi^r(x_n) = \int \Psi^r(\xi) P_n(x_n|\xi) d\xi. \tag{5.5}$$

After a few iterations the technique is expected to give a better estimate of the undistorted $\Psi(\xi)$.

In LR79 the iterative technique is applied to the observed distribution $\phi(q)$ of mass ratios $q = M_{\rm sec}/M_{\rm prim}$ of 173 SBII systems that were selected from DAO6. In LR79 the parameter $x = \ln q$ is used in the above equations. The technique requires a first estimate of the distribution $\Psi^0(\xi)$. In LR79 $\Psi^0(\xi) = 1/(\xi_2 - \xi_1)$ is assumed (i.e., a flat distribution in $\ln q$) for $\xi \in (\xi_1, \xi_2)$, where $\xi_1 = \ln q_1$ and $\xi_2 = \ln q_2$. Here q_1 and q_2 are the boundaries of the q interval considered. The variances of the measurements were either taken from the original publications of the SBII orbits in DAO6, or they were re-calculated by the authors of LR79. Figure 5.1 shows the smoothed histogram (dashed line) of $\phi(x)$ ($x = \ln q$) for the SBII systems in DAO6, based on raw data, as presented in LR79. The same figure shows the distribution the authors obtained after five iterations with the R-L method (solid line).

In LR79 it is argued that the latter distribution is a better estimate of the real q–distribution of SBII systems. The enhanced peak in the distribution for $q = 1$ leads the authors to the conclusion that the number of binary stars with nearly identical components is indeed significantly higher than suggested by the original, not corrected histogram. They attribute the large number of systems with $q \simeq 1$ to the formation of binaries through fragmentation.

5.3 The stop criterion

In the Richardson-Lucy method, the estimates $\Psi^r(\xi)$ of the real distribution do not converge in the sense that $\Psi^{r+1}(\xi) - \Psi^r(\xi) \to 0$ as $r \to \infty$. In L74 it was already indicated that, after obtaining a best estimate from iteration r_o, $\Psi^r(\xi)$ tends to fit (and amplify) the statistical fluctuations in the data for every $r > r_o$. In the absence of knowledge of the real distribution, an independent criterion is needed to decide when to stop the iterations.

In L74 it is suggested to use

$$\chi^2\{\phi^r\} = N \sum_{i=1}^{I} \frac{(\phi_i - \phi_i^r)^2}{\phi_i + \phi_i^r} \tag{5.6}$$

for this purpose. Here ϕ is the observed distribution according to Eq. (5.1), and ϕ^r is the estimate of this distribution at iteration r, as it is occurring in Eq. (5.5). N is the number of elements in the sample. The indices i indicate that we are referring to the histogrammed versions of the distributions: $\phi_i = \phi(x_i)\Delta x$, and $\phi_i^r = \phi^r(x_i)\Delta x$, where i is the number of the bin, and Δx the bin width. I is the total number of bins. $\chi^2\{\phi^r\}$, according to Eq. (5.6), is then a measure of the agreement of the estimate of the observed distribution ϕ^r with the actual observed distribution ϕ.

The probability $P(\chi^2\{\phi^r\}|\nu)$, that a χ^2 greater than the observed value is found by chance when the number of degrees of freedom involved in the distributions is ν, is the measure that is used to decide when to stop the iterations. According to L74 a reasonable criterion is $P(\chi^2|\nu) > 0.05$, and $1 - P > 0.05$.

For the comparison of two distributions by means of the χ^2-test, and the determination of probability $P(\chi^2|\nu)$, see Press *et al.* (1986).

5.4 A new sample of SBII systems

The *Eighth* version of the Catalogue, DAO8, contains 237 SBII systems of luminosity classes IV, IV–V, and V, in the mass-ratio interval $0.6 < q < 1.3$, which comprise a sample quite similar to that considered in LR79, but 37 % larger. Figure 5.2 shows the smoothed and the ordinary histogram of the sample. The smoothed histogram of the new sample is the equivalent of the histogram in Fig. 1 in LR79. From the high frequency features in Fig. 1 in LR79 it is estimated that the original histogram consisted of approximately 20 bins in the interval $0.6 < q < 1.3$.

The smoothing technique employed in the present paper is the following. The contribution $\phi_n(q)$ of each q_n measurement to the observed mass-ratio distribution

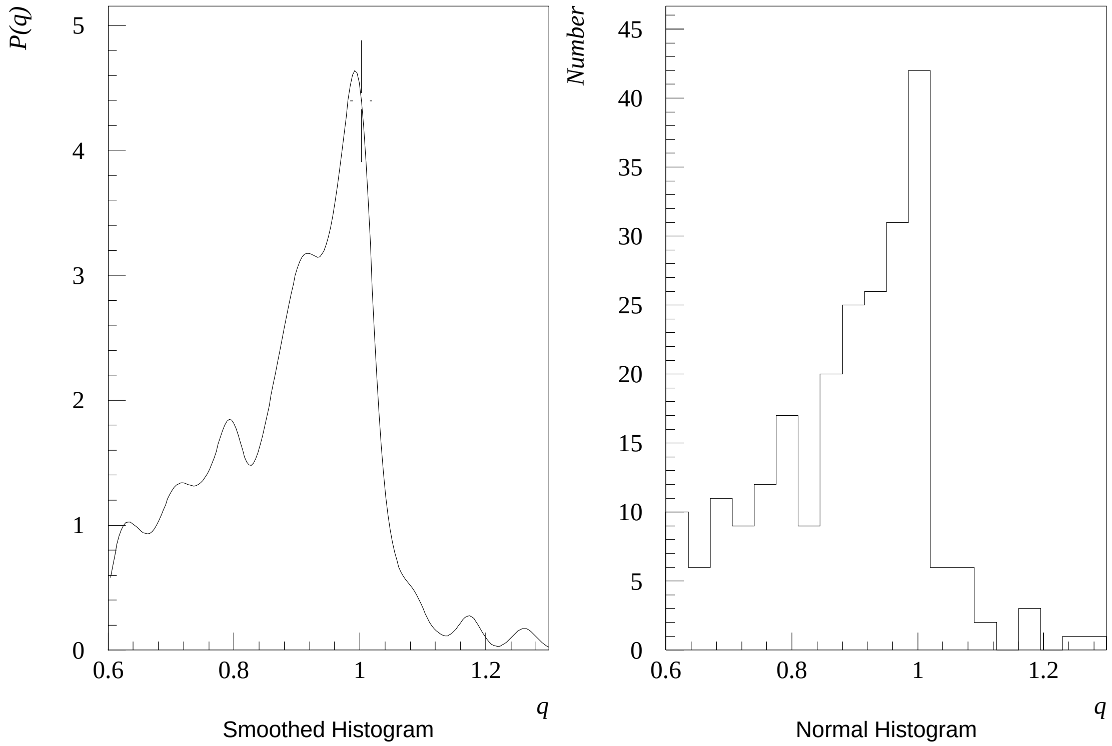

Figure 5.2: Smoothed (left hand panel) and ordinary histogram (right hand panel) of 237 SBII systems of luminosity classes IV, IV–V, and V from DAO8. Both histograms are based on 20 bins of width $\Delta q = 0.035$. The cross in the smoothed histogram represents the standard deviation of the bin concerned. The horizontal bar of the cross indicates the smoothing length or bin width. The smoothing technique that was employed to obtain the left hand histogram is described in the text.

$\phi(q)$ is represented by a normal distribution centered at q_n, with a standard deviation determined by the required binwidth of the histogram, Δq

$$\phi_n(q) = \frac{1}{\frac{1}{2}\Delta q\sqrt{2\pi}} \exp\left(-\frac{(q_n - q)^2}{\frac{1}{2}(\Delta q)^2}\right).$$

The standard deviation σ associated with this normal distribution is $\sigma = \frac{1}{2}\Delta q$. A smoothed histogram, or rather a normalized representation of the probability density function $\phi(q)$ of the observed distribution, is then obtained by adding the individual contributions

$$\phi(q) = \frac{1}{N}\sum_{n=1}^{N}\phi_n(q).$$

By taking $\Delta q = 0.035$ we obtain histograms with 20 bins in the interval $0.6 < q < 1.3$.

An independent indication of the errors in the observed $\phi(q)$ distribution – i.e. other than the indication that is obtained from the individual errors σ_n in the q_n values – may be obtained by means of the Bootstrap technique (Efron 1982). For an introduction to the Jackknife and the Bootstrap, and an extensive list of references, see Nobelis (1990).

Here the Bootstrap technique is used to estimate the standard deviation of the fraction ϕ_i in each bin i of the smoothed $\phi(q)$ histogram. To that end the original sample of q_n measurements $\mathcal{S}(q_1, q_2, \ldots, q_N)$ is resampled B times, producing samples $\mathcal{S}^*_b(q^*_1, q^*_2, \ldots, q^*_N)$. The q^*_n values in each sample $\mathcal{S}^*$ are selected at random (with replacement) from the original q_n in the sample $\mathcal{S}$. Of each sample $\mathcal{S}^*_b$ a new smoothed histogram ϕ^*_b is made. For each bin i, the Bootstrap standard deviation $\widehat{\mathrm{SD}}_i$ is then evaluated as

$$\widehat{\mathrm{SD}}_i = \left\{ \frac{1}{B-1} \sum_{b=1}^{B} \left(\phi^*_{i,b} - \overline{\phi^*_i} \right)^2 \right\}^{1/2} , \tag{5.7}$$

where $\overline{\phi^*_i}$ is the average estimate of the fraction in each bin: $\overline{\phi^*_i} = \sum_{b=1}^{B} \phi^*_{i,b}/B$. For large numbers B, $\overline{\phi^*_i} \rightarrow \phi_i$.

The resulting $\widehat{\mathrm{SD}}_i$ for the ϕ_i histogram of the 237 SBII systems from DAO8 are given in Table 5.4 (section 5.7). They were obtained by taking $B = 100$ Bootstrap samples. In Fig. 5.2 the worst $\widehat{\mathrm{SD}}_i$ error is indicated by an error bar.

5.5 The R-L method applied to the new sample of SBII systems

The Richardson-Lucy method was applied to the sample of 237 SBII systems from DAO8.

In LR79, the average value $\overline{\sigma_q}$ of the individual errors σ_n of the 173 SBII systems from DAO6 is given to be $\overline{\sigma_q} = 0.036$. In a first approach I did not determine the individual errors σ_n of the 237 SBII systems in the new sample, but I took the average value of LR79 to represent the error in each q_n value. The result of the application of the R-L method with these error estimates (i.e. $\sigma_n = 0.036$ in Eq. (5.2)) is shown in Fig. 5.3 for five iterations.

In the application of the method to the SBII systems, the functions and variables in Eqs. (5.1) to (5.5), have the following meaning: $\Psi(\xi)$ is the real distribution of real mass ratios ξ, $\Psi^r(\xi)$ is the estimate of the real distribution resulting from iteration r, x is the observed mass ratio q (notice that we do not follow the convention in LR79, where $x = \ln q$, although the latter has computational advantages), $\phi(x)$ is the observed mass-ratio distribution, denoted below by $\phi(q)$, and $\phi^r(x)$ is the estimate of the observed mass-ratio distribution resulting after iteration r, denoted below by $\phi^r(q)$. The starting distribution was $\Psi^0(\xi) = 1/(q_2 - q_1)$. The distributions Ψ^r, ϕ^r, and ϕ were in the calculations represented by arrays of 1000 elements.

The $\chi^2\{\phi^r\}$ values, evaluated according to Eq. (5.6), were found to be 76, 7.4, 4.2, 3.4, and 3.1, for $r = 1, \ldots, 5$ respectively. For a number of bins $I = 20$, the number of degrees of freedom associated with these χ^2 values is $\nu = I - 1 = 19$. The $P(\chi^2\{\phi^r\}|\nu)$ values are then $9 \cdot 10^{-9}$, 0.99, 1, 1, and 1, respectively (values $P > 0.995$ are rounded off to 1).

According to the stop criterion proposed in L74, the estimate of the real distribution Ψ^1, obtained after the first iteration, is not acceptable, because $P(\chi^2|\nu) = 9 \cdot 10^{-9} \ll 0.05$. At $r = 2$ the probability already enters the realm where $1 - P < 0.05$, but this is the first iteration at which $P > 0.05$, so we accept it as giving the best estimate $\Psi^r(\xi)$ of the real mass-ratio distribution $\Psi(\xi)$. The Ψ^2 estimate shows only a modest maximum at $q \simeq 1$, which is even less pronounced than in the smoothed

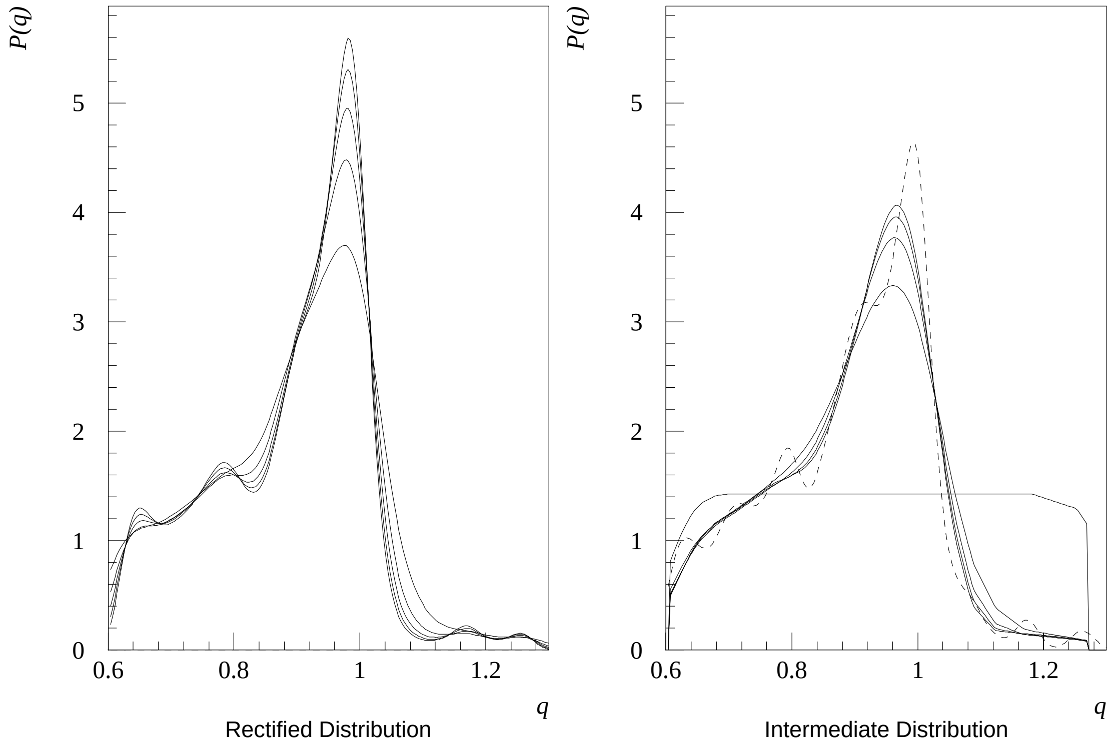

Figure 5.3: The Richardson-Lucy method applied to 237 SBII systems from DAO8, assuming that each individual error $\sigma_n = \overline{\sigma_q} = 0.036$. The left hand panel shows the estimates $\Psi^r(\xi)$ of the real mass-ratio distribution for $r = 1$ (curve with the lowest maximum at $q \sim 1$) to 5 (curve with the highest maximum at $q \sim 1$). The right hand panel shows the estimates of the observed distribution $\phi^r(q)$ for each iteration (solid lines), and the smoothed histogram of the observed distribution $\phi(q)$ (dashed line).

histogram of $\phi(q)$. It should be noticed, however, that the best estimate Ψ^2, resulting from the iterative technique, assumes errors $\sigma_n = 0.036$, while the smoothed histogram is obtained with bin widths $\Delta q = 0.035$, which correspond approximately to 'errors' $\sigma_q = 0.0175$.

The value $\overline{\sigma_q} = 0.036$ given in LR79 is an average value. When Table 1 of recomputed q–values in LR79 is representative of the whole sample of SBII systems, the errors vary around the average value of $\overline{\sigma_q} = 0.036$ with a standard deviation of about $\sigma_{\sigma_q} = 0.018$. In the next experiment, artificial errors were attributed to the q_n values of the 237 SBII systems from DAO8 by random selection from a normal distribution with mean 0.036 and standard deviation 0.018. When this procedure gave values $\sigma_n < 0.015$, they were set to $\sigma_n = 0.015$, just as was done in LR79 for systems with real $\sigma_q < 0.015$. With these error values entered in Eq. (5.2), the R-L method was again applied to the sample of SBII systems from DAO8. Figure 5.4 shows the resulting estimates $\Psi^r(\xi)$, again for five iterations. The peak at $q \simeq 1$ has become very prominent for iterations beyond $r = 2$.

The estimates $\phi^r(q)$ of the observed distribution are now no longer continuous curves. This is due to the fact that $\phi^r(q)$ is evaluated at the N values q_n, according

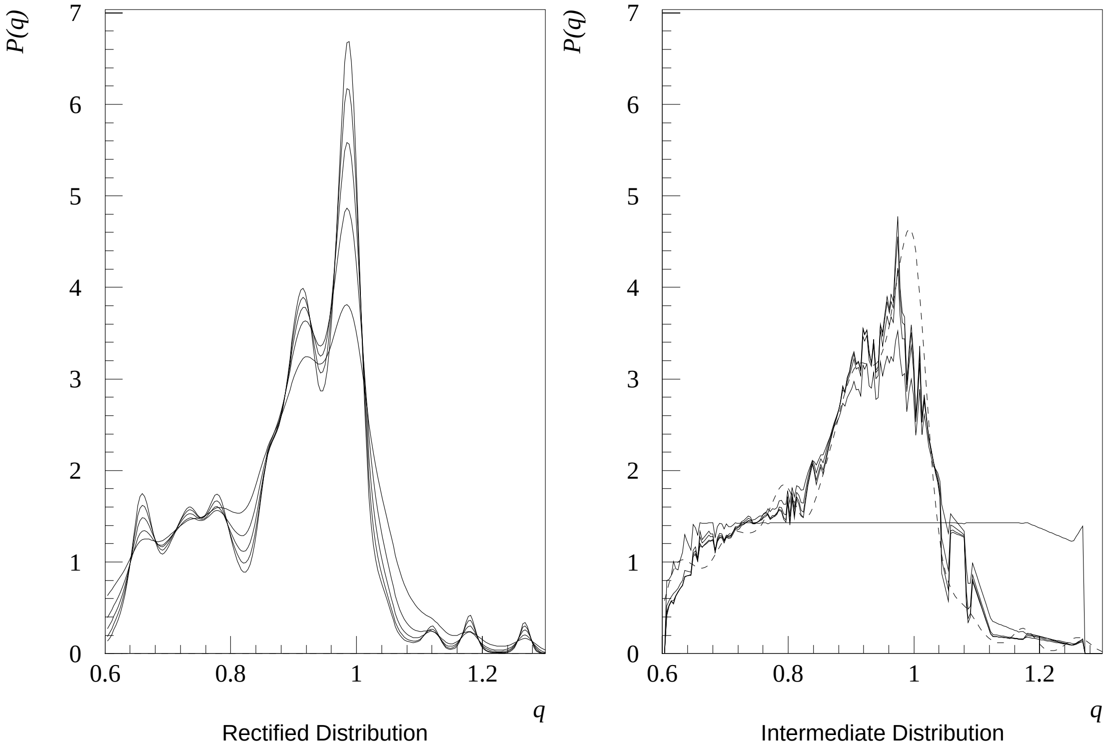

Figure 5.4: The Richardson-Lucy method applied to 237 SBII systems from DAO8, assuming individual errors according to a normal distribution with mean $\sigma_n = \overline{\sigma_q} = 0.036$ and standard deviation $\sigma_{\sigma_q} = 0.018$. The left hand panel shows the estimates $\Psi^r(\xi)$ of the real mass-ratio distribution for $r = 1$ to 5. The right hand panel shows the estimates of the observed distribution $\phi^r(q)$ for each iteration (solid lines), and the smoothed histogram of the observed distribution $\phi(q)$ (dashed line).

to Eq. (5.5). With $\Psi^r(\xi)$ being a continuous function, two adjacent measures q_n, with markedly different values σ_n, are cause for very different values of $P_n(q_n|\xi)$, thus causing the 'noise' in $\phi^r(q)$.

The $\chi^2\{\phi^r\}$ values now evaluate to 76, 9.0, 5.8, 5.1, and 4.8, respectively. The corresponding probabilities $P(\chi^2\{\phi^r\}|\nu)$ are $1 \cdot 10^{-8}$, 0.97, 1, 1, and 1. Interpreting the probabilities P as we did above, the best estimate Ψ^r of the real mass-ratio distribution is again given by the second iteration. The maximum at $q \simeq 1$ is, also in this distribution, not exceptionally prominent.

Although I did not determine the individual errors σ_n from the original papers, as Lucy and Ricco did in LR79, the distribution of the real errors will not be very different from the distribution assumed here. This is confirmed by the shape of the Ψ^5 distribution, which is very similar to the same distribution given in LR79 for the 173 SBII systems from DAO6.

5.6 A synthetic sample

To verify the findings in the previous section, the Richardson-Lucy method was applied to a synthetic sample of elements with a distribution similar to that of the mass-ratio distribution of the SBII systems, and of the same size. The advantage of the synthetic sample is that we know the real distribution, so we are able tp compare the estimates produced by the technique with this real distribution.

The synthetic sample was obtained by first making a linear least squares fit to the observed q–distribution of the SBII systems at the interval $0.6 < q < 1$. At this interval, the distribution is described by $g(q) = a + bq$, with $a = -4.938$ and $b = 8.71$. A distribution g, with $\int g(q)\,dq = 1$ at $q_1 < q < q_2$, with $q_1 = 0.6$ and $q_2 = 1.3$, is obtained when at the interval $1 < q < q_2$ the function has a constant value $g(q) = c = 0.627$. A sample of 237 elements, distributed according to g, was obtained by determining the inverse G^{-1} of the cumulative distribution G of g:

$$G(Q) = \int_{q_1}^{Q} g(q)\,dq.$$

With

$$g(q) = \begin{cases} a + bq, & \text{for } q_1 < q \leq 1, \\ c, & \text{for } 1 < q \leq q_2, \end{cases}$$

where $a = -4.938, b = 8.71$, and $c = 0.627$, we find

$$G(Q) = \begin{cases} a(Q-1) + \frac{1}{2}b(Q^2 - 1), & \text{for } q_1 < Q \leq 1, \\ G(1) + c(Q-1), & \text{for } 1 < Q \leq q_2. \end{cases}$$

Notice that $G(Q = q_1) = 0$, and $G(Q = q_2) = 1$. Let $G(Q) = R$, with R a real number at the interval (0,1). Then the inverse function $Q = G^{-1}(R)$ is found to be

$$Q = G^{-1}(R) = \begin{cases} \dfrac{-a + \sqrt{a^2 + 2b(aq_1 + \frac{1}{2}bq_1^2 + R)}}{b}, & \text{for } 0 < R \leq G(1), \\ \frac{1}{c}RG(1) - 1, & \text{for } G(1) < R \leq 1. \end{cases} \tag{5.8}$$

When R is a random number in the interval $0 < R \leq 1$, the expression produces a random sample of Q_n values, distributed according to $g(q)$ (Banks and Carson II 1984).

When the expression of Eq. (5.8) is evaluated for $R_n = n/N$, with $n = 1, \ldots, N$, the resulting Q_n are distributed according to $g(q)$ in a homogeneous way, with no statistical fluctuations. For $N = 237$, Eq. (5.8) was used to produce a synthetic sample with a distribution similar to the mass-ratio distribution of the SBII systems from DAO8. Figure 5.5 shows the normal and the smoothed histogram of the $g(q)$ distribution of the sample. Errors in the smoothed histogram were again obtained by means of the Bootstrap method. They are given in Table 5.1.

The theoretical distribution $g(q)$ may be regarded as the real $\Psi(\xi)$ distribution in the R-L method. $\Psi(\xi)$ and $g(q)$ are therefore assumed to be equivalent in the discussion below. The Q_n values produced by Eq. (5.8) comprise the real sample,

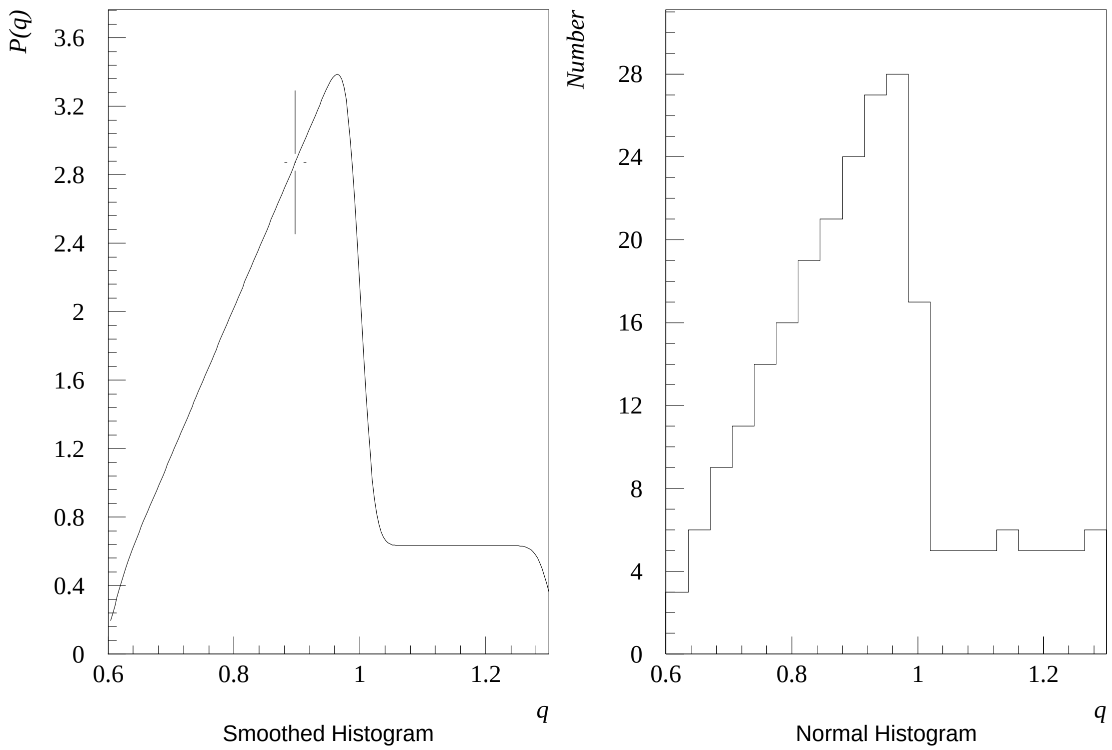

Figure 5.5: Normal (right hand panel) and smoothed histogram (left hand panel) of the q–distribution of a synthetic sample of 237 elements. The error bar in the smoothed histogram is the worst $\widehat{\mathrm{SD}}_i$ value obtained by means of the Bootstrap (see text).

Table 5.1: Errors $\widehat{\mathrm{SD}}_i$ in the smoothed histogram of the synthetic sample, obtained by means of the Bootstrap, with the original sample resampled $B = 100$ times. Bin numbers are indicated by i, the center of each bin is given by q_i.

i	q_i	$\widehat{\mathrm{SD}}_i$	i	q_i	$\widehat{\mathrm{SD}}_i$
1	0.6175	0.15	11	0.9675	0.41
2	0.6525	0.23	12	1.0025	0.37
3	0.6875	0.25	13	1.0375	0.23
4	0.7225	0.29	14	1.0725	0.19
5	0.7575	0.31	15	1.1075	0.19
6	0.7925	0.37	16	1.1425	0.19
7	0.8275	0.37	17	1.1775	0.20
8	0.8625	0.40	18	1.2125	0.20
9	0.8975	0.42	19	1.2475	0.21
10	0.9325	0.40	20	1.2825	0.20

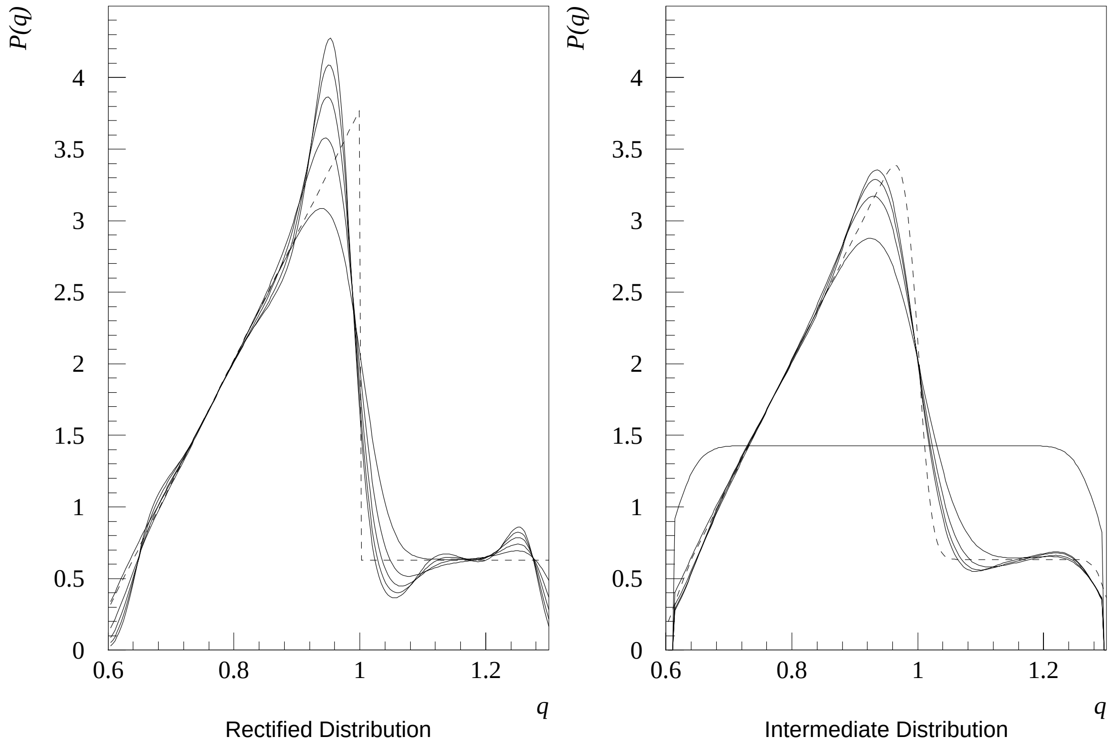

Figure 5.6: Left hand panel: estimates Ψ^r (solid lines) of the real distribution $\Psi(\xi)[= g(q)]$ (dashed line) of the synthetic sample produced by the Richardson-Lucy method for $r = 1, \ldots, 5$. Right hand panel: estimates ϕ^r (solid lines) of the observed distribution ϕ (dashed line) as they occur in the procedure at the same iterations r. The simulation assumes fixed errors $\sigma_n = \overline{\sigma_q} = 0.036$.

and are equivalent to ξ_n. From the real sample, we may obtain an 'observed' sample by simply taking $q_n = Q_n$.

To each q_n value in the synthetic sample, artificial errors $\sigma_n = \overline{\sigma_q} = 0.036$ were assigned. To this sample, the R-L method was applied. Table 5.2 gives the $\chi^2\{\phi^r\}$, $P(\chi^2|\nu)$, and $\chi^2\{\Psi^r\}$ values for the comparison of ϕ^r with ϕ, and Ψ^r with Ψ. The $\chi^2\{\Psi^r\}$ were calculated according to

$$\chi^2\{\Psi^r\} = N \sum_{i=1}^{I} \frac{(\Psi_i - \Psi_i^r)^2}{\Psi_i}. \tag{5.9}$$

This formula differs from Eq. (5.6) in the denominator, in which Ψ_i^r does not occur. This is because the theoretical Ψ_i are known exactly, contrary to the observed ϕ_i.

According to the $P(\chi^2\{\phi^r\}|\nu) > 0.05$ criterion, the best estimate Ψ^r of the real distribution is obtained for $r = 2$. This is confirmed by the $\chi^2\{\Psi^r\}$ values, which indicate a best comparison of Ψ^r with Ψ after the same iteration.

The R-L method was applied again to the synthetic sample, this time assuming random individual errors σ_n according to the normal distribution with mean 0.036 and standard deviation 0.018, that was also used to simulate the errors in the q_n values of the real SBII systems. Again, when this procedure gave errors $\sigma_n < 0.015$,

Table 5.2: χ^2 and $P(\chi^2\{\phi^r\}|\nu)$ values for the comparison of ϕ^r with ϕ, and Ψ^r with Ψ for $r = 1, \ldots, 5$ as they occur in the application of the Richardson-Lucy method (Fig. 5.6) to the synthetic sample of Fig. 5.5. Errors σ_n are all equal to $\overline{\sigma_q} = 0.036$. The number of degrees of freedom involved in the calculations of the $P(\chi^2|\nu)$ is $\nu = 19$. Values $P > 0.9995$ were rounded off to $P = 1.000$.

r	$\chi^2\{\phi^r\}$	$P(\chi^2\{\phi^r\}\|\nu)$	$\chi^2\{\Psi^r\}$
1	41.2	0.002	4.2
2	2.9	1.000	1.7
3	1.6	1.000	2.6
4	1.3	1.000	3.9
5	1.1	1.000	5.4

Table 5.3: χ^2 and $P(\chi^2|\nu)$ values for the comparison of ϕ^r with ϕ, and Ψ^r with Ψ for $r = 1, \ldots, 5$ as they occur in the application of the Richardson-Lucy method (Fig. 5.7) to the synthetic sample of Fig. 5.5. In this alternative application of the method, errors σ_n are randomly distributed according to a normal distribution with mean 0.036 and standard deviation 0.018. The number of degrees of freedom involved in the calculations of the $P(\chi^2|\nu)$ is $\nu = 19$. Values $P > 0.9995$ were rounded off to $P = 1.000$.

r	$\chi^2\{\phi^r\}$	$P(\chi^2\{\phi^r\}\|\nu)$	$\chi^2\{\Psi^r\}$
1	40.6	0.003	4.6
2	3.5	1.000	3.5
3	2.1	1.000	5.6
4	1.8	1.000	8.3
5	1.8	1.000	10.5

they were set to 0.015. Table 5.3 gives the χ^2 and $P(\chi^2|\nu)$ values for the comparison of ϕ^r with ϕ, and Ψ^r with Ψ, as they result for the sample with the new errors. According to the table, the first iteration at which $P(\chi^2\{\phi^r\}|\nu) > 0.05$ is $r = 2$. At $r = 2$, $1 - P < 0.05$, so this is at the same time the iteration at which we should stop. This is in agreement with the $\chi^2\{\Psi^r\}$ values, which show a minimum at $r = 2$.

5.7 Discussion

The results found in the previous sections from the application of the Richardson-Lucy method to the sample of SBII systems and to the synthetic sample, are to a large extent independent of the histogram ϕ_i that represents the observed distribution $\phi(q)$ in the stop criterion. The experiments were repeated with ϕ being represented by histograms consisting of 4 and 35 bins respectively. Even in the case of 4 bins iteration $r = 2$ was, according to the $P(\chi^2\{\phi^r\}|\nu) > 0.05$ criterion, indicated as giving the best Ψ^r estimate of the real distribution. However, it is best to represent the observed

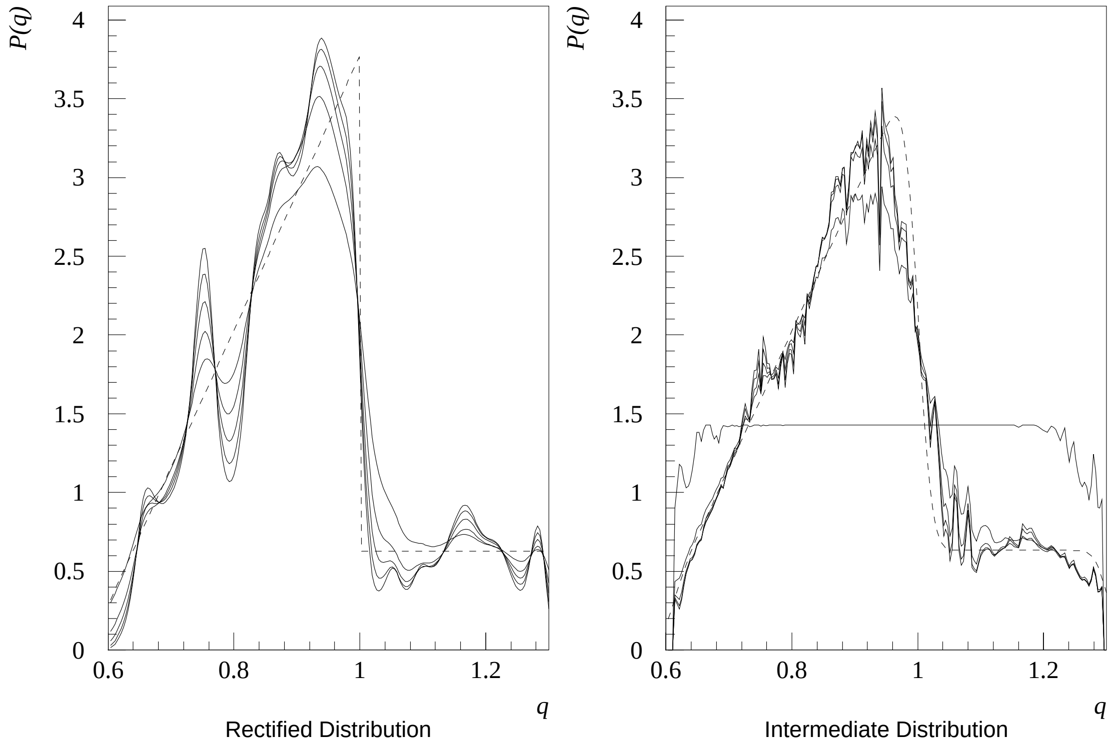

Figure 5.7: Left hand panel: estimates Ψ^r (solid lines) of the real distribution Ψ (dashed line) produced by the Richardson-Lucy method for $r = 1, \ldots, 5$. Right hand panel: estimates ϕ^r (solid lines) of the observed distribution ϕ (dashed line) as they occur in the procedure at the same iterations r. This alternative simulation assumes random errors σ_n according to a normal dsitribution with mean $\overline{\sigma_q} = 0.036$ and standard deviation $\sigma_{\sigma_q} = 0.018$ (see the text).

distribution ϕ by a histogram with bin widths comparable in size to the average error in the data (i.e., $\Delta x = 2\overline{\sigma_x}$), or narrower. By taking too few bins, one might cause the iterations to stop too early with respect to any (real) high frequency features in the distribution.

The significance of the maximum in the mass-ratio distribution of the SBII systems at $q \simeq 1$, as produced by the R-L method, was tested by applying the Bootstrap method in the following way. The original data set $\mathcal{S}(q_1, q_2, \ldots, q_N)$ was resampled again to obtain B samples $\mathcal{S}^*_b(q^*_1, q^*_2, \ldots, q^*_N)$. To each of these samples, the R-L method was applied, to obtain estimates $\widehat{\mathrm{SD}}_i$ of the standard deviation of the Ψ^2 and Ψ^5 estimates of the real distribution. Table 5.4 gives the Bootstrap $\widehat{\mathrm{SD}}_i$ errors for the observed distribution ϕ_i and the estimates Ψ^2_i and Ψ^5_i of the real distribution. Notice that the $\widehat{\mathrm{SD}}_i$ values depend on the number of bins that is chosen to represent the ϕ and Ψ^r distributions (cf. Eq. (5.7)). However, when the number of bins is the same for each distribution, the $\widehat{\mathrm{SD}}_i$ values may be compared directly. The errors in ϕ_i and Ψ^2_i are very similar, while the errors in Ψ^5 are on average twice as large as in the other distributions.

In bin $i = 12$, with central value $q_i = 1.0025$, the $\widehat{\mathrm{SD}}_i$ errors in the distributions

Table 5.4: Results of the application of the Richardson-Lucy method to 100 Bootstrap samples of the sample of 237 SBII systems from DAO8. Bin numbers are indicated by i, the q_i values indicate the center of each bin.

		$\widehat{\mathrm{SD}}_i$ in:					$\widehat{\mathrm{SD}}_i$ in:		
i	q_i	ϕ_i	Ψ_i^2	Ψ_i^5	i	q_i	ϕ_i	Ψ_i^2	Ψ_i^5
1	0.6175	0.26	0.24	0.26	11	0.9675	0.41	0.43	0.73
2	0.6525	0.24	0.36	0.59	12	1.0025	0.49	0.53	0.93
3	0.6875	0.24	0.29	0.43	13	1.0375	0.26	0.32	0.38
4	0.7225	0.30	0.30	0.48	14	1.0725	0.20	0.15	0.16
5	0.7575	0.30	0.29	0.45	15	1.1075	0.13	0.12	0.15
6	0.7925	0.41	0.35	0.56	16	1.1425	0.08	0.08	0.08
7	0.8275	0.27	0.27	0.33	17	1.1775	0.15	0.20	0.36
8	0.8625	0.35	0.45	0.73	18	1.2125	0.02	0.05	0.06
9	0.8975	0.46	0.49	0.82	19	1.2475	0.10	0.05	0.06
10	0.9325	0.40	0.51	0.86	20	1.2825	0.08	0.08	0.06

are 0.49, 0.53, and 0.93, respectively. The large error in Ψ^5 renders the peak at $q \simeq 1$ in this distribution very uncertain.

The R-L method requires information about the individual errors σ_n in the q_n values for the determination of its estimates of the real distribution. It should be noticed, however, that the errors that were attributed to the q_n measurements in this paper are artificial errors, which are assumed to provide an (reasonable) approximation of the real errors. Therefore nothing can be decided about the real mass-ratio distribution based on the distributions Ψ^r as they are presented here. The experiments with the artificial errors only serve to investigate the workings of the R-L method. They do, however, allow us to decide that, for any likely real distribution of errors, the iterations should be stopped earlier than was assumed in LR79, thus giving no evidence for an enhanced number of binaries with mass ratios near unity.

The Bootstrap method does *not* require information about the individual σ_n errors to determine the $\widehat{\mathrm{SD}}_i$ errors in the smoothed histogram of the $\phi(q)$ distribution.

The stop criterion has not been investigated for its theoretical merits. However, in the cases of the synthetic samples studied in this paper, it proved to be very effective in determining after which iteration r the best estimate Ψ^r of the real distribution is obtained.

5.8 Conclusions

The Richardson-Lucy method for the rectification of observed distributions is applied to a sample of SBII systems from DAO8, in order to test its capability to restore the mass-ratio distribution of the SBII systems. The method is also applied to a synthetic sample of elements, to allow verification of the results obtained for the SBII systems.

In the application of the R-L method, the criterion proposed in L74 for stopping the iterations was used. For different assumptions about the errors in the data it was found, for both the real and the synthetic sample, that the process should be stopped after the second iteration.

For the SBII systems from DAO8 this means that the number of systems with $q \simeq 1$ is not enhanced as much as was found in LR79, where five iterations were performed to obtain a best estimate of the mass-ratio distribution of 173 SBII systems from DAO6.

In a separate paper (Hogeveen 1991, this thesis Chapter 3), it is shown that the mass-ratio distribution of the SBII systems in DAO8 is seriously affected by several selection effects. Measurement errors are only part of the problem. There is no doubt that the R-L method is one of the best techniques to correct for these errors. However, the determination of the individual errors σ_n, as it was carried out by Lucy and Ricco in LR79 for the 173 SBII systems from DAO6 in their sample, is a laborious and time consuming enterprise. Although such an undertaking would, for the new SBII systems in DAO8, certainly be worthwhile, it is beyond the scope of the research project in the course of which this work was carried out.

In the case of the SBII systems, the severity of the selection effects other than measurement errors, justifies the approximation of the observed mass-ratio distribution by a smoothed histogram, especially when the errors in this distribution are estimated by means of a reliable method such as the Bootstrap.

Acknowledgements

I wish to acknowledge stimulating discussions with Prof. E. P. J. van den Heuvel, and the valuable comments provided by Prof. J. van Paradijs and Dr. R. Wijers. I gratefully acknowledge the use of the computer facilities of, and support from, NIKHEF–H (Netherlands Institute for Nuclear and High Energy Physics, section H) in Amsterdam. I wish to thank Drs. A. H. Batten and J. M. Fletcher of the Dominion Astronomical Observatory, Victoria, B.C., for providing me with a machine-readable version of the *Eighth Catalogue of the Orbital Elements of Spectroscopic Binary Stars* (1989). My work is supported through the Netherlands Foundation for Research in Astronomy (Astron) by the Netherlands Foundation for the Advancement of Research, grant 782–371–026.

References

Adorf, H. M.: 1990, *ST–ECF Newsletter* **14**, 8

Banks, J. and Carson II, J. S.: 1984, *Discrete Event System Simulation*, Prentice Hall, New Jersey

Batten, A. H.: 1967, *Sixth Catalogue of the Orbital Elements of Spectroscopic Binary Stars*, Dominion Astrophysical Observatory, Victoria, B.C., (DAO6)

Batten, A. H., Fletcher, J. M., and MacCarthy, D. G.: 1989, *Eighth Catalogue of the Orbital Elements of Spectroscopic Binary Sytems*, Dominion Astrophysical Observatory, Victoria, B.C., (DAO8)

Efron, B.: 1982, *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania

Hogeveen, S. J.: 1991, *Astrophys. Space Sci.*, submitted

Lucy, L. B.: 1974, *Astron. J.* **79**, 745

Lucy, L. B. and Ricco, E.: 1979, *Astron. J.* **84**, 401

Nobelis, P. P.: 1990, in C. Jaschek and F. Murtagh (eds.), *Errors, Bias and Uncertainties in Astronomy*, pp 89–100, Cambridge University Press, Cambridge

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: 1986, *Numerical Recipes*, Cambridge University Press, Cambridge