ASTRONOMY & ASTROPHYSICS
SUPPLEMENT SERIES

*Astron. Astrophys. Suppl. Ser.* **140**, 247–260 (1999)



# On the automatic folding of optical rotation curves

**D.F. Roscoe**

School of Mathematics, Sheffield University, Sheffield, S3 7RH, UK
e-mail: D.Roscoe@shef.ac.uk

Received June 7; accepted September 6, 1999

**Abstract.** Mathewson, Ford and Buchhorn (1992, MFB hereafter) published the unreduced data for the optical rotation curves of 967 southern sky spiral galaxies. Recognizing that accurate dynamical modelling of spiral galaxies required the availability of a large data-base of correspondingly accurately folded rotation curves, Persic & Salucci (1995, PS hereafter) undertook to fold the MFB sample in an appropriately meticulous way; of the 967 folded rotation curves, 900 were judged by PS to be of moderate to excellent quality, whilst 67 were judged to be of poor quality and of very limited use for dynamical studies. The folding process used by PS was a time-consuming and labour-intensive one in which the quality of each fold was judged "by eye".

Subsequently, MFB (1996) published the unreduced optical rotation curves for approximately another 1100 southern sky spirals and, undoubtedly, more will follow from various sources. For this reason, and because of the importance of having large numbers of accurately folded rotation curves for dynamical studies, we have developed the automatic folding algorithm described herein.

An uncompiled Fortran program (using NAG routines) and data files are available via http://www.shef.ac.uk/~ap1dfr. Download the text file "ReadMe" and follow instructions.

**Key words:** methods: numerical — galaxies: kinematics and dynamics

## 1. Introduction

Rotation curves are the primary source of information about spiral galaxy dynamics and, as such, they are essential tools for any programme which seeks to determine relationships between dynamics and mass distributions in spiral galaxies - whether this is for dark matter studies, or for testing gravitational theories. However, before they can be used in any dynamical context, they must be accurately folded. Prior to the PS contribution of 900 accurately folded optical rotation curves, there was no large data base available and, so far as we are aware, even though a large amount of unfolded data is available, no other large data base of folded rotation curves has subsequently been made available. The reason for this absence is most probably that rotation curve folding has hitherto been a labour-intensive "by eye" process - fine for small volumes of data, but extremely time consuming for large volumes of data. For this reason, we undertook to attempt the development of an automatic folding process. In the event, it has turned out that a successful development has only been possible because of the availability of the PS data-base of 900 folded rotation curves which has acted as an essential template against which the performance of the auto-folder described here was judged in the various stages of its development. In addition to this, the development has required the recognition, and use, of certain correlations (previously unsuspected, so far as we are aware) which point directly to underlying physical processes, the study of which may lead to a deeper understanding of spiral galaxies and their evolution.

The PS sample of folded optical rotation curves was analysed in Roscoe 1999A, and gave rise to the following primary result: defining $R$ to be the radial displacement from the kinematic centre and $V$ to be the rotational velocity at radial displacement $R$, then, to a very high statistical precision, rotation curves conform to the law

$$\frac{V}{V_0} = \left(\frac{R}{R_0}\right)^{\alpha} \tag{1}$$

$$\log V_0 \approx -0.584 - 0.133\,M - 0.000243\,S,$$

$$\log R_0 \approx -3.291 - 0.208\,M - 0.00292\,S,$$

where $M$ denotes absolute magnitude, $S$ denotes surface brightness and $\alpha$ is a parameter, constant for any given galaxy, but which which varies between galaxies. The high statistical precision which accompanied this result implied, in its turn, that the PS folding process was very

*Send offprint requests to*: D.F. Roscoe

reliable and that their sample of folded rotation curves could be considered as a "model solution" against which the success, or otherwise, of any automatic folding algorithm could be properly judged. Consequently, in all that follows, the developing auto-folder is continually tested by using it to fold the raw MFB data which formed the core of the PS sample, and comparing the resulting solution with the PS solution over the same sample.

In Sect. 2 we describe the general problem of folding rotation curves whilst, in Sect. 3, we describe precisely what quality of the PS model solution is used to judge the automatic folding algorithm under development and in Sect. 4 we briefly discuss a class of methods which were considered, but proved to be insufficiently robust.

In the remaining sections we provide a sequential description of a robust auto-folder development paying particular attention to the difficulties which arose, and to the manner of their resolution.

## 2. The problem of folding

A typical unreduced rotation curve is shown in Fig. 1 and, in this figure, the origin of the angular coordinate axis denotes the measured optical centre, $O_{\rm opt}$ say, of the galaxy; apart from projection corrections (required because spiral galaxies are generally not seen edge-on), a successful folding of this rotation curve requires (a) an accurate estimate of the cosmological redshift, $V_{\rm sys}$ say, that must be subtracted from the velocity measurements along the whole curve; (b) an accurate estimate of the dynamical centre, $O_{\rm dyn}$ say, of the rotation curve. The MFB data-

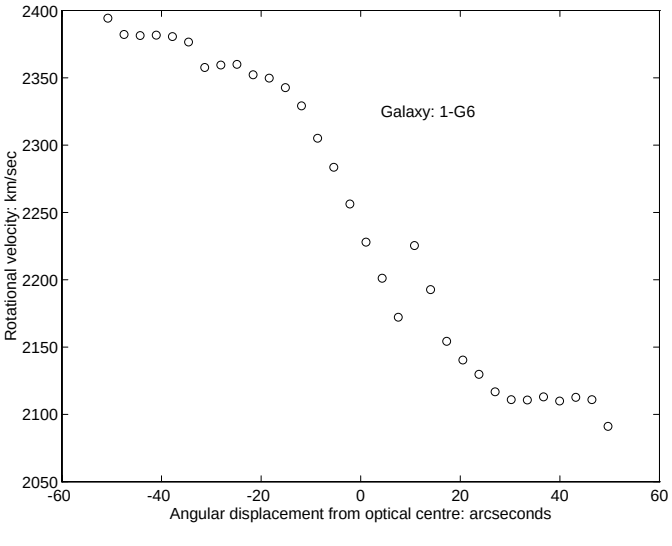


**Fig. 1.** Typical unfolded rotation curve

base for the unreduced curves provides approximations for $V_{\rm sys}$ and $O_{\rm dyn}$ (this latter being identified with $O_{\rm opt}$) which are sufficiently good for the peculiar velocity studies which motivated MFB but not good enough for the dynamical studies which motivated PS. So, the problem resides entirely in the accuracy with which $V_{\rm sys}$ and $O_{\rm dyn}$ can be determined, and the major obstacle to obtaining this accuracy is simply the noisiness of rotation curve data - if the galaxies from which the data was drawn were perfectly rotationally symmetric, and if the measurements were exact, then it would be a trivial task to find the centre of asymmetry of any rotation curve.

### *2.1. Assessment of noise in MFB $H_\alpha$ velocity data*

As a means of determining the level of noise in each velocity measurement, MFB provided a cross-correlation coefficient, $0 < \rho \leq 1$, derived by comparing the $H_\alpha$ line profile with an artificial template line profile (see MFB for details). In their analysis, PS found that selecting those velocities for which $\rho > 0.35$ gave a mean rms of less than about 10 km s$^{-1}$, whilst for $\rho \leq 0.35$ an rms of up to 20 km s$^{-1}$ was to be expected. For this reason, they restricted their data to values for which $\rho > 0.35$.

In our development, we have found that $\rho > 0.4$ is necessary for the accurate operation of the auto-folder being developed.

## 3. The adopted measure of folding quality

The very noisy nature of rotation curve data means that any solution to the folding of any given rotation curve is likely to contain a large element of uncertainty; it follows that, given two reasonable folding techniques, there are likely to be many rotation curves which are best folded by one particular technique, and vice-versa for other rotation curves whilst, at the same time, any objective judgement about which are the best individual solutions is likely to be extremely difficult (if not impossible) at best. It follows that any judgement about the relative merits of any two folding methods must be made by some global statistical method applied to the folded solution of a large number of rotation curves. The method used for the present development is described in the following.

### *3.1. The ln A distribution*

As we have indicated, the basic result (shorn of the fine details) of Roscoe 1999A was that optical rotation curve data is described, to a high statistical precision, by the power-law structure $V = AR^\alpha$ where $(A, \alpha)$ are constants which differ from galaxy to galaxy. Essentially, but with the "hole-cutting" data-reduction process described in Roscoe 1999A, $\ln A$ and $\alpha$ for each rotation curve were determined by linear regression of $\ln V$ data on $\ln R$ data. Thus, for the PS sample of 900 good quality rotation curves we obtain

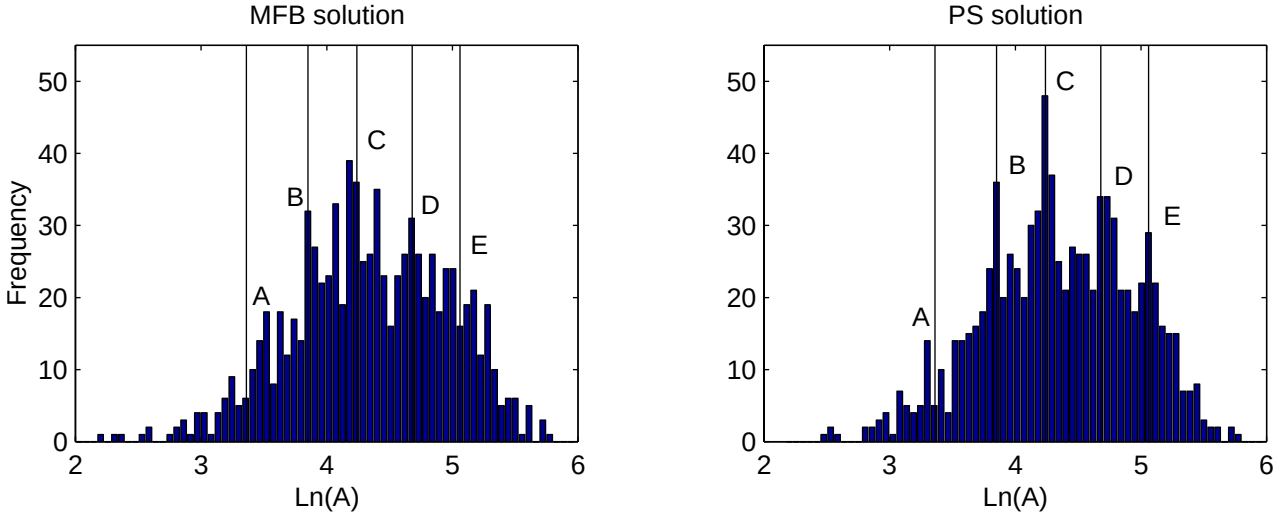


**Fig. 2.** Comparison of MFB and PS solutions. The vertical lines in both diagrams indicate the positions of the A, B, C, D and E peaks in the PS solution

900 pairs $(\ln A, \alpha)$. Figure 2 (left) shows the $\ln A$ distribution for the MFB solution (that is, when the rotation curves are folded using MFB's optical centre and their estimate of $V_{\rm sys}$), whilst Fig. 2 (right) shows the corresponding distribution for the PS solution. A comparison shows that the peaks labelled B, C and D in the MFB solution are reproduced - enhanced and with much reduced noise - in the PS solution. Since the MFB solutions were determined directly from observational estimates of optical centres and systematic redshifts whilst the PS solutions were determined using a "by-eye" folding process without any reference to an underlying power-law structure (and hence without reference to any $\ln A$ distribution), then the foregoing comparison strongly suggests that the B, C, D peak structure reflects objective physical qualities in the underlying distribution of rotation curves. This conclusion was reached independently (with the addition of the A and E peaks), supported by very strong statistics, in Roscoe 1999B.

For these reasons, the $\ln A$ distribution is adopted here as a suitable statistical representation of a folding solution over a large number of rotation curves.

## 4. Non-integral techniques for folding rotation curves

Prior to developing the method described in detail herein, various other methods were tried, and these shared the common property of being non-integral methods. That is, they attempted the process of folding by minimizing functionals defined by various forms of direct comparison of the velocity distributions on the approaching and receding sides of the spiral.

Whilst each of these various methods always folded some rotation curves successfully, they also each had high failure rates. The basic problem with such methods is that functionals defined over noisy data (which rotation curve data is) do not always display the mathematical properties of minimum points - even when in the close neighbourhood of such points. The results obtained with one obvious method in this class are briefly described in Appendix A.

## 5. An overview of auto-folder development

The ideal solution to the problem of minimizing functions defined over noisy data is to use methods based on integral techniques, and we have chosen a method based on the Fourier decomposition of rotation curves. Because of the necessarily detailed nature of the following auto-folder development, it is useful to begin with a section-by-section overview of the process:

- Sect. 6: Define a basic folding method based upon a Fourier decomposition of a rotation curve. This gives the Mk I auto-folder;
- Sect. 7: Refine the basic method by optimizing the number of Fourier modes to be used for the decomposition of any given rotation curve. This gives the Mk II auto-folder;
- Sect. 8: Remember that, by the considerations of Roscoe 1999A (see also Appendix B), the interiors of rotation curves generally behave differently from their exteriors. So, consider the possibility that this differential behaviour might be due to noisy disturbance induced by the proximity of the central bulge, and therefore accountable in the folding process by removing the noisy interior sections - a process termed as "hole-cutting";
- Sect. 9: Modify the Mk II auto-folder by a global implimentation of the hole-cutting process to obtain the Mk III auto-folder;

– Sect. 10: Compare the folding solutions with and without hole-cutting, and note that the effectiveness of the hole-cutting process appears to be $\ln A$ dependent;
– Sect. 11: Investigate the $\ln A$ dependency of hole-cutting effectiveness;
– Sect. 12: Refine the application of the hole-cutting process according the results of the latter invetsigation to obtain the Mk IV auto-folder.

## 6. The Mk I auto-folder

The basic folding technique is developed from the idea that *any* function, $y \equiv f(x), \ -X \leq x \leq X, \ X > 0$ can be decomposed into two parts, one of which is asymmetric on the region about $x = 0$ and one of which is symmetric on the region about $x = 0$.

In the present case, and for ideal data, once $V_{\rm sys}$ has been found and accounted for by subtraction from the measured rotation velocities, then the resulting processed rotation curve should be perfectly asymmetric about $O_{\rm dyn}$, its dynamical centre. However, suppose we have $V_{\rm sys} \approx V'_{\rm sys}$ and $O_{\rm dyn} \approx O'_{\rm dyn}$ then, after subtracting $V'_{\rm sys}$ the curve will be only *approximately* asymmetric about the assumed dynamical origin, $O'_{\rm dyn}$ - and can therefore be considered composed of an exact asymmetric part together with an exact symmetric part. The basic folding technique therefore consists in minimizing the symmetric component with respect to variations in the estimates of $V_{\rm sys}$ and $O_{\rm dyn}$.

### 6.1. The Fourier decomposition

The symmetric component, $f_{\rm sym}(x)$, of any function $y \equiv f(x)$ defined on the interval $-X \leq x \leq X, \ X > 0$ can be calculated using the finite cosine transform:

$$f_{\rm sym}(x) \ = \ \sum_{m=0}^{\infty} A(m) \cos(m\pi x/X),$$

$$A(m) \ = \ \frac{1}{2\,X} \int_{-X}^{X} f(x) \cos(m\pi x/X) \ {\rm d}x. \qquad (2)$$

If the function represents a rotation curve then, in practice, $y \equiv f(x), \ -X \leq x \leq X$ is replaced by a numerical function defined at $N + 1$ points on the discretized interval $-X \leq x \leq X$. Suppose that this numerical function is given as $(x_0, f_0), ..., (x_N, f_N)$. In the ideal case, for which $x_0 = -X$ and $x_N = X$, then the Fourier component, $A(m)$, can be approximated by a direct numerical integration of (2). However, the more usual case will be that the numerical function is defined over a non-symmetric interval (for example, $-1.83 \leq x \leq 2.47$). In this case, the requirements of the cosine transform makes it necessary to discard and/or interpolate points of the numerical function so that it becomes defined over the largest symmetric interval that can be fitted into the original non-symmetric interval; in the case of the example, this would be $-1.83 \leq x \leq 1.83$.

### 6.2. How many Fourier modes should be computed?

In very simple situations, for which the numerical function is defined at equal intervals, $\Delta$ say, in the independent variable, there is no point in calculating Fourier modes which have a wave-length less than $2\Delta$. This leads to the standard rule that the computation of $N$ Fourier modes requires a minimum of $2^N + 1$ data points, when these are equally spaced. Conversely, a given number of points then informs how many Fourier modes it is worthwhile including.

However, rotation curve data is very frequently defined over non-equally spaced intervals and, in this case, the basic rule has to be replaced by the rule that there is no point in calculating Fourier modes with a wave-length less that $2\Delta_{\rm max}$, where $\Delta_{\rm max}$ is the *largest* interval separating adjacent points on the rotation curve.

### 6.3. The minimization procedure

The foregoing describes how to compute the symmetric Fourier modes for given estimates of $V_{\rm sys}$ and $O_{\rm dyn}$. Suppose $M$ of these are computed, then we form the functional

$$F(M) = \sum_{m=0}^{M} [A(m)]^2$$

and minimize this with respect to variations in the estimates of $V_{\rm sys}$ and $O_{\rm dyn}$. The very noisy nature of rotation curve data requires that a very robust minimization procedure, using no derivatives, should be employed. A routine based on the Simplex algorithm (Nelder & Mead 1965) was found to be very effective in the present case. When applied to the PS sample, this Mk I auto-folder gives the $\ln A$ distribution shown in Fig. 3, right. We immediately see that peaks B, C and D of the PS solution (reproduced in Fig. 3 left) are excellently reproduced by the Mk I auto-folder; however, peak E is lost in noise.

It is worth emphasizing that the reproduction, in Fig. 3 right, of the major part of the peak structure of the PS solution confirms that this structure is *not* an artifact of the PS procedure but is, at the very least, inherent to the sample.

## 7. The Mk II auto-folder

Whilst the logical development of the Mk I auto-folder appears to preclude any possibility of refinement, the noisiness of the data raises the possibility that the best fold of any given rotation curve is not necessarily the fold which uses the maximal number of Fourier modes possible for the rotation curve. Thus, for example, the algorithm of Sect. 6.2 might indicate the use of five Fourier modes whilst, in practice, the noisiness of the data might allow a better fold with three Fourier modes.

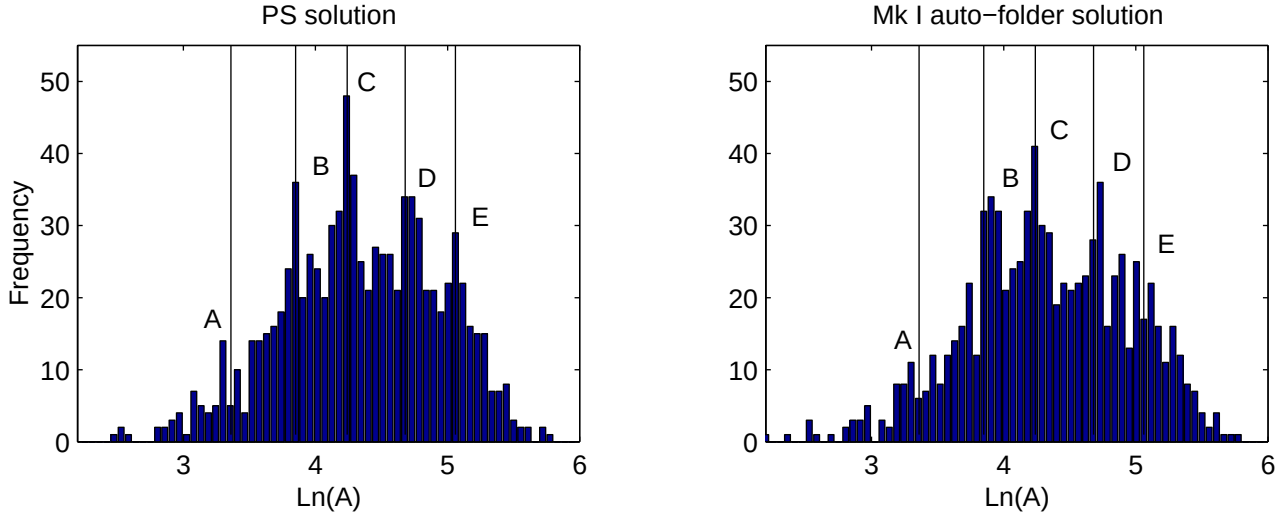


**Fig. 3.** PS solution compared with the Mk I auto-folder solution. The vertical lines indicate the positions of the A, B, C, D and E peaks in the PS solution

Thus, given a rotation curve for which a maximum of $N$ Fourier modes are indicated by the data, then the logic of this latter argument forces us to consider a set of potential solutions consisting of the 1-mode fold, the 2-mode fold, ..., the $N$-mode fold; we must then choose the "best" solution from this set of $N$ possibilities. Naturally, since the objective quality of the folding process over the whole PS sample is to be judged against the PS solution represented by Fig. 3 left, then the means by which we select between these $N$ folds must be *independent* of this latter figure. The means by which this is done is described in the following.

### *7.1. Choosing between Fourier modes*

The logic of the mode-choosing strategy is rooted in the result of Roscoe 1999A that optical rotation curves are described by the power law $V = A\,R^{\alpha}$ so that $\ln V$ and $\ln R$ are in a linear relation: Suppose that, for any given rotation curve, we have a choice between $N$ folds, consisting of the 1-mode fold, the 2-mode fold, ..., the $N$-mode fold. For each of these $N$ folds we compute $\ln A$ as described in Sect. 3 (and applying the hole-cutting algorithm described Roscoe 1999A) and, at the same time, record the residual mean square (rms) arising from the regression. We then simply choose the mode which has the *least* rms associated with it.

In other words, we simply choose the mode that provides the tightest linear fit between $\ln V$ and $\ln R$ *after* the hole-cutting algorithm has been applied.

### *7.2. The Mk II auto-folder applied to the PS sample*

Applying the Mk II auto-folder described above to the PS sample we find that it gives the $\ln A$ distribution of Fig. 4 right. Comparison with the Mk I solution (Fig. 4 left), shows that the B-peak has strengthened considerably, the C-peak is more-or-less unchanged, the D-peak has weakened slightly whilst the E-peak has strengthed considerably. Thus, the overall impression is that the Mk II auto-folder represents an improvement over the Mk I auto-folder.

## 8. Hole-cutting as an active component of the folding process

The "hole-cutting" technique was introduced in Roscoe 1999A on the basis of the qualitative argument that the interior parts of rotation curves must be significantly disturbed by the presence of the bulge. This argument is given quantitative support in Appendix B, where it is explicitly shown how hole-cutting has a very powerful effect on the quality of the power-law fits to rotation curves; furthermore, it is shown in Appendix C, how the cut-out sections can be interpreted as dynamical transition regions on the interiors of rotation curves.

So far, however, the hole-cutting process has played no active part in the actual folding process - that is, we have so far assumed that rotation curves should be folded using all of the available data. This latter assumption is justified if the transition regions are as asymmetric about the dynamical centres as the exterior parts of rotation curves are assumed to be.

But, by the very nature of "dynamical transition regions" in the generality of physical systems, the possibility exists that the dynamics in such regions is, to some

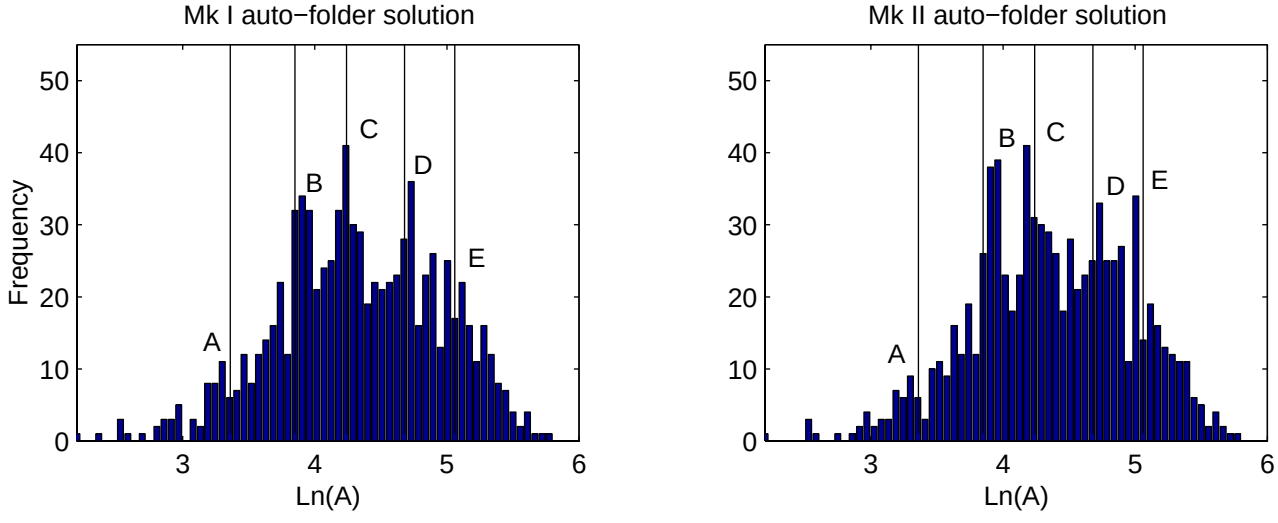


**Fig. 4.** Comparison of Mk I and Mk II auto-folder solutions. The vertical lines indicate the positions of the A, B, C, D and E peaks in the PS solution

extent, chaotic. In this latter case, the shapes of the interior parts of rotation curves could be then significantly disturbed from the asymmetry assumed for the exterior parts of rotation curves - and this would lead to systematic inaccuracies in the final solutions due entirely to the inclusion of the transition regions in the folding process. It would then follow that using only the exterior parts of rotation curves for the folding process will lead to more accurate folding.

## 9. The Mk III auto-folder

The Mk III auto-folder is based on the working hypothesis that the folding process should only use the exterior parts of the rotation curve, so that hole-cutting becomes an active part of this process. However, there is one immediate practical difficulty: the hole-cutting algorithm requires rotation curves to be *already* folded. The resolution of this problem is embodied in the algorithmic definition of the Mk III auto-folder, described in Sect. 9.1, and the solution arising from it applied over the PS data is shown in Fig. 5 right.

### *9.1. Algorithmic definition of the Mk III auto-folder*

The Mk III auto-folder is defined by the following algorithm:

- For any given rotation curve, fold using the Mk II auto-folder;
- record the number of Fourier modes used for this fold, say $N_{\rm f}$, and the estimates for $V_{\rm sys}$ and $O_{\rm dyn}$;
- apply the hole-cutting algorithm defined Roscoe 1999A to this folded rotation curve, and record the removed points;
- take the original *unfolded* rotation curve, and reduce it by removing these recorded points;
- fold this reduced rotation curve using the Fourier method already described, but taking care to redefine the Fourier coefficients in the manner described in the next subsection;
- for this second-stage folding process, use either the $N_{\rm f}$ modes recorded at the first-stage folding process, or the number of modes defined by the algorithm of Sect. 6.2, whichever is the least;
- use the recorded estimates for $V_{\rm sys}$ and $O_{\rm dyn}$ as the initial guess for the minimization process of this second-stage folding.

### *9.2. Partition of Fourier region*

The application of the hole-cutting algorithm of Roscoe 1999A cuts a hole out of the rotation curve which is centred on the estimated dynamical centre of the galaxy; consequently, instead of representing a continuous curve, the reduced rotation curve data represents two separated continuous sections with a gap in between. This means that the Fourier cosine coefficients cannot be computed using (2), but must be computed by an integral of the form

$$A(m) = \frac{1}{2\,X}\int_{-X}^{-X'} f(x)\cos(m\pi x/X)\ \mathrm{d}x\ + \frac{1}{2\,X}\int_{X'}^{X} f(x)\cos(m\pi x/X)\ \mathrm{d}x, \qquad (3)$$

where $X'$ is some positive number representing the boundaries of the cut-out hole. In practice, of course, the hole-cutting strategy will generally remove a section which is non-symmetric about the current estimate of $O_{\rm dyn}$, and so the data on one of the two sections of the reduced rotation curve must be interpolated/extrapolated to ensure

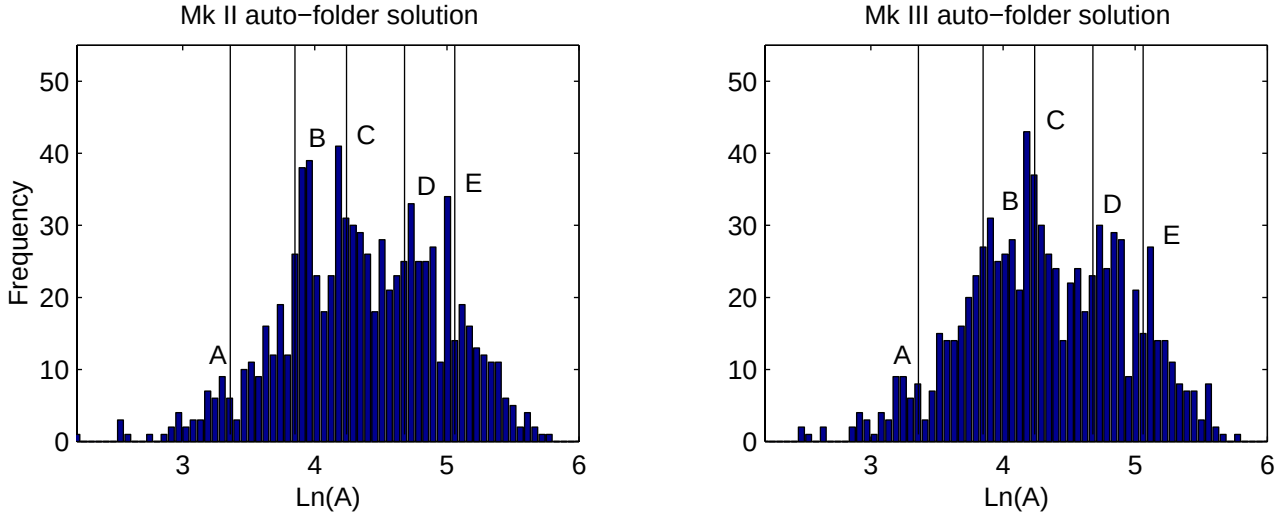


**Fig. 5.** Comparison of Mk II and Mk III auto-folder solutions. The vertical lines indicate the positions of the A, B, C, D and E peaks in the PS solution

that the hole which has been effectively cut out of the rotation curve is centred exactly on the current estimate of $O_{\rm dyn}$. Once the removed section is centred in this way, and the Fourier modes computed using (3), rather than (2), the process is as before.

## 10. The Mk II and Mk III auto-folders compared

Comparison of the results of the Mk II auto-folder (Fig. 5 left) with the results of the Mk III auto-folder (Fig. 5 right) show unambiguously that the B peak is stronger in Mk II whilst the C peak is stronger in Mk III. The status of the D peak is slightly ambiguous in that whilst it reaches a higher level in Mk II, it appears more strongly differentiated from the background on Mk III. The status of the E peak is similarly ambiguous in that, in both solutions, it has a "tuning fork" shape relative to the position of the E peak in the PS solution (vertical line) but, overall, has about equal strength in each diagram.

To summarize, contrary to the working hypothesis made in Sect. 9, that the innermost parts of rotation curves should always be cut out, circumstances appear to suggest that the application of a combination of the Mk II and Mk III auto-folders, each applied according to estimated value of $\ln A$, might result in an algorithm of greater effectiveness. However, since the precise final value of $\ln A$ for any given rotation curve will depend on the folding process used, then it is clear that defining a strategy for choosing between the Mk II and the Mk III auto-folders will not necessarily be a simple process. In the following two sections, we describe the means by which we are able to deduce an effective working strategy.

## 11. A detailed investigation of the effects of hole-cutting

The considerations of the previous section indicate the possibility that a selective application of the Mk II and Mk III auto-folders might be beneficial. But the only possible *objective* justification for using the active hole-cutting process of the Mk III auto-folder selectively is that the transition regions of the selected rotation curves are intrinsically noisier than the transition regions of the non-selected rotation curves. This raises the question of how we can determine which rotation curves have intrinsically noisy transition regions.

### *11.1. A working hypothesis*

A consideration of the relative strengths of the B and C peaks in each of the diagrams of Fig. 5 provides prima facie evidence for the idea that C-peak rotation curves have intrinsically noisy transition regions (since hole-cutting has increased the strength of the C peak), whilst B-peak rotation curves have relatively quiet transition regions (since hole-cutting has decreased the strength of the B peak).

Since rotation curves with intrinsically noisy transition regions will necessarily be subject to more inaccurate folding than other rotation curves then, by the observations of Appendix D, it might be expected that C-peak rotation curves will be associated with correspondingly greater changes in $(R_{\rm min}, R_{\rm opt})$ correlations through hole-cutting than B-peak rotation curves.

Generalizing, this leads to the working hypothesis that an analysis of changes in $(R_{\rm min}, R_{\rm opt})$ correlations through the hole-cutting process as a function of $\ln A$ might reveal significant non-uniformities indicating corresponding non-uniformities in the distribution of rotation curves with intrinsically noisy transition regions. This eventuality would

then provide the required objective rationale of when to employ hole-cutting as an active component of the folding process, and when not to. The following subsections describe this analysis.

### *11.2. Plan of analysis*

We require a detailed assessment of the effects of the hole-cutting strategy on $(R_{\min}, R_{\text{opt}})$ correlations as a function of $\ln A$. Since the potential objective is to determine for what values of $\ln A$ the Mk III auto-folder should be used and since, in practice, these values of $\ln A$ will be obtained via the prior use of the Mk II auto-folder (cf. Sect. 9), it follows that the $R_{\min}$ values used in the proposed analysis should likewise be drawn from Mk II auto-folder solutions.

Comparing the Mk II and Mk III auto-folder solutions of Fig. 5, we were able to identify an approximate partition of the range $2.2 \leq \ln A \leq 6$ in which to study the effects of the hole-cutting strategy and, with a little experimentation, were able to refine this into the four cases, $2.2 < \ln A \leq 4.2$, $4.2 < \ln A \leq 4.5$, $4.5 < \ln A \leq 4.8$ and $4.8 < \ln A \leq 6$.

It is to be emphasized that, although the foregoing ranges were identified by numerical experimentation, the changes in the behaviour of the $(R_{\min}, R_{\text{opt}})$ correlations between these ranges are so strong, and the sample sizes so large, that the ordinary processes of random statistical fluctuation as a source of the variations can be ruled out with virtual certainty; consequently, considerable reliance can be placed on qualitative deductions made from the results summarized in Table 1 which lists the pre-hole cutting and the post-hole cutting values of the indices of determination, $R^2$, for each of the four $\ln A$ ranges:

**Table 1.** Effects of hole-cutting on coefficient of determination

| $\ln A$ range | $R^2$ before hole-cutting | $R^2$ after hole-cutting | $N$ |
|---|---|---|---|
| (2.2, 4.2] | 20.7% | 26.3% | 348 |
| (4.2, 4.5] | 6.4% | 34.6% | 150 |
| (4.5, 4.8] | 1.9% | 8.4% | 141 |
| (4.8, 6.0] | 17.8% | 40.2% | 224 |

It is clear that there are two distinct modes of behaviour for changes in the $(R_{\min}, R_{\text{opt}})$ correlation through the hole-cutting process: there are very strong changes in $\ln A$ ranges (4.2, 4.5] and (4.8, 6.0], and relatively weak changes in $\ln A$ ranges (2.2, 4.2] and (4.5, 4.8] Recalling our working hypothesis (cf. Sect. 11.1) that significant change in the $(R_{\min}, R_{\text{opt}})$ correlation is potentially indicative of rotation curves with intrinsically noisy transition regions, then Table 1 provides an objective rationale for employing the Mk III auto-folder on the sub-intervals $4.2 < \ln A \leq 4.5$ and $4.8 < \ln A \leq 6.0$, but not on the remaining intervals.

## 12. The Mk IV auto-folder

The considerations of the previous section have led us to conclude that the Mk IV auto-folder is to be defined as a composite of the Mk II auto-folder used in the ranges $2.2 < \ln A \leq 4.2$ and $4.5 < \ln A \leq 4.8$, and the Mk III auto-folder in $4.2 < \ln A \leq 4.5$ and $4.8 < \ln A \leq 6$. Bearing in mind that these ranges were defined via an analysis based on the sole use of the Mk II auto-folder, the implication is that the folding process begins by using the Mk II auto-folder for all rotation curves and, when the resulting $\ln A$ lies in one of the appropriate ranges, the solution is repeated using the Mk III auto-folder.

After some marginal fine-tuning of the given $\ln A$ ranges, a workable definition of the Mk IV auto-folder can be given as follows:

- Use the Mk II auto-folder over the whole sample;
- Calculate the resulting $\ln A$ for each rotation curve in the manner described in Sect. 3 and Appendix B;
- Whenever $4.22 < \ln A \leq 4.55$ or $4.8 < \ln A \leq 6$, then continue with the Mk III strategy; see Sect. 9. Occasionally, this secondary process fails to re-fold the rotation curve, in which case the original Mk II solution is used.

### *12.1. Results and implications*

The results of applying this to the PS sample are shown in Fig. 6. It is at once apparent that the Mk IV auto-folder provides a huge improvement over the Mk II and Mk III auto-folders and, on the basis of its ability to produce a high fidelity copy of the peak-structure of the PS solution (Fig. 6 left), the Mk IV auto-folder is accepted as the best achievable with present resources.

The Mk IV auto-folder strategy was based on the working hypothesis (cf. Sect. 11.1), that significant changes in $(R_{\min}, R_{\text{opt}})$ correlations through the hole-cutting process are indicative of rotation curves with intrinsically noisy transition regions. This strategy has been manifestly so successful that its success must be considered as very strong circumstantial evidence supporting the view that the working hypothesis on which it is based is correct. Thus, cross-referencing Table 1 with Fig. 6 (right), we see that strong circumstantial evidence exists to suggest that the rotation curves associated with the peaks C and E have intrinsically noisier transition regions on their interiors than do those associated with the peaks B and D.

Finally, we note from Table 1 that the transition region behaviours of the two "quiet transition region" peaks, B and D, are mutually distinct: specifically, for the B peak, the coefficient of determination changes from 20.7% to

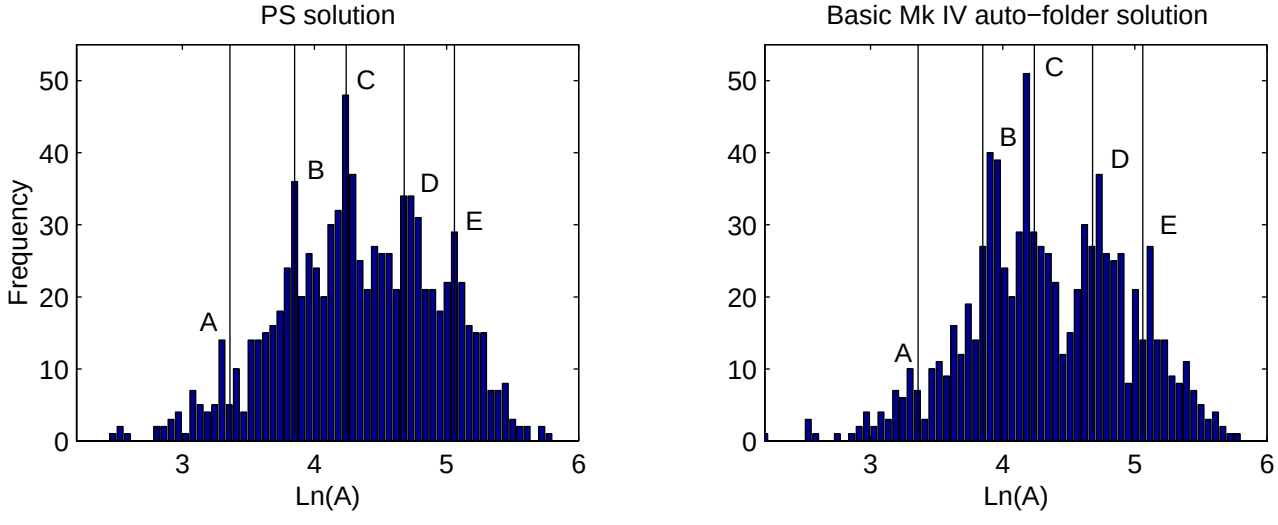


**Fig. 6.** Left: PS solution; Right: Mk IV Auto-folder applied to PS sample. The vertical lines indicate the positions of the A, B, C, D and E peaks in the PS solution

26.3% through hole-cutting - and each of these values signifies a very strong $(R_{\min}, R_{\text{opt}})$ correlation; by contrast, for the D peak, the coefficient of determination changes from 1.9% to 8.4% - the first value signifies a very weak $(R_{\min}, R_{\text{opt}})$ correlation whilst the second signifies a moderately weak $(R_{\min}, R_{\text{opt}})$ correlation.

As we have already noted, the sample sizes are sufficiently great that we can discount random statistical fluctuations as the source of these differences between the B and D peak rotation curves, and so the question arises of why strong $(R_{\min}, R_{\text{opt}})$ correlations are present for B peak rotation curves, but are more-or-less absent for D peak rotation curves? Since the existence of strong correlations implies the existence of some coherent physical process, the question becomes, why are the correlating effects of this physical process not present for the D peak rotation curves?

These considerations all point to the need of further research.

### *12.2. Some examples of folded rotation curves*

Figures 7 and 8 give examples of ten rotation curves folded using the auto-folder, and the corresponding PS solution for these rotation curves - the examples were chosen as representative of the whole set, and not for any particular qualities they display. Whilst there are differences in detail for the two folds of any given rotation curve, there is no obvious sense in which either of the solutions is the superior one.

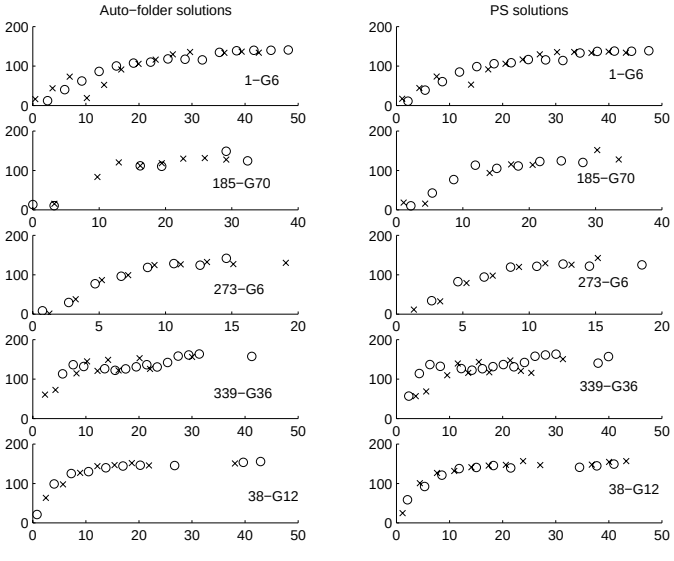


**Fig. 7.** Five folded rotation curves: Left auto-folder soluton; Right PS solution

## 13. Folding of simulated rotation curves

So far, the efficacy of the developed auto-folder has been judged by means of a comparison of the PS folding solution and the auto-folder's solution defined on the same sample; the assumption is, of course, that the PS solution over the sample can be considered "correct" in an overall statistical sense. However, in the absence of objective certainties, it is useful to see how the auto-folder deals with the folding of simulated rotation curves for which the folded solution is known a priori.

The following describes the results of applying the auto-folder to three sets of five simulated rotation curves per set, where each set is chosen to represent a particular

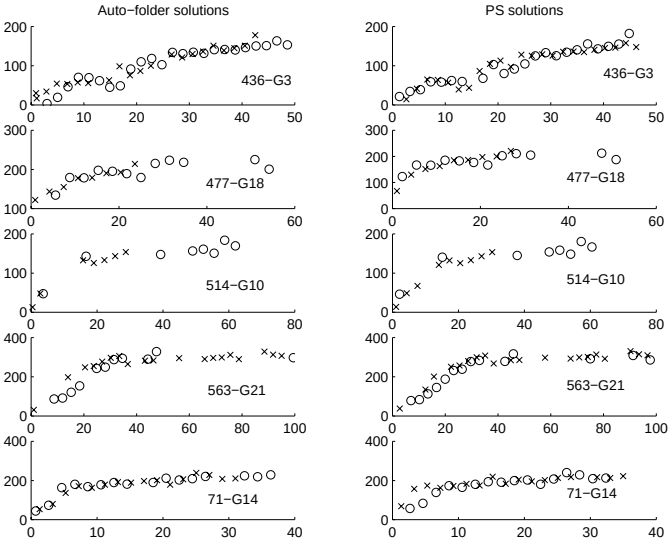


**Fig. 8.** Five folded rotation curves: Left auto-folder soluton; Right PS solution

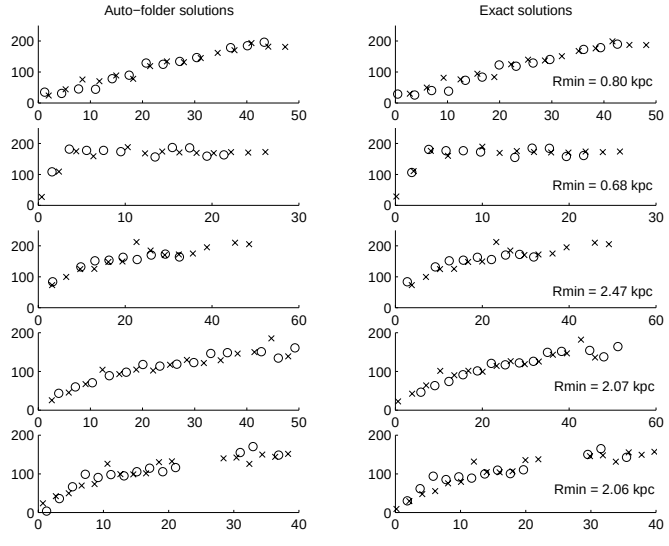


**Fig. 10.** Folding of simulated medium-density sampling rotation curves

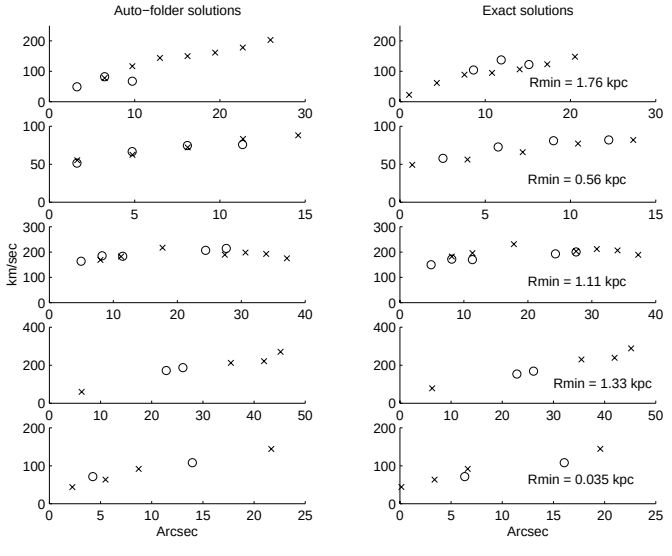


**Fig. 9.** Folding of simulated low-density sampling rotation curves

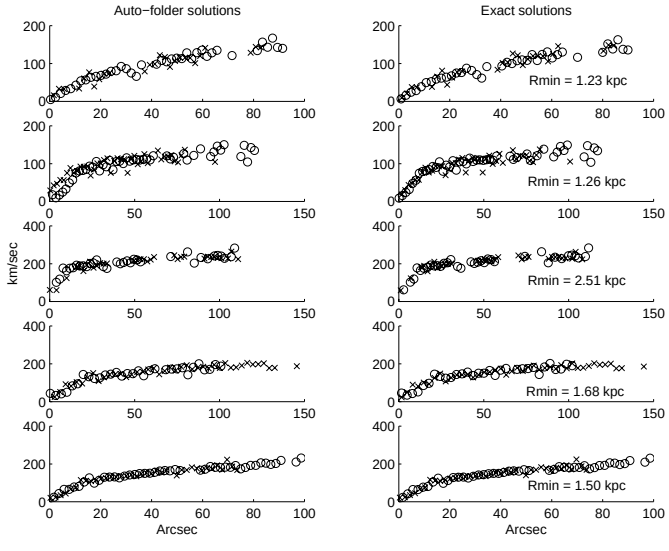


**Fig. 11.** Folding of simulated high-density sampling rotation curves

sampling regime - low-density sampling, medium-density sampling and high-density sampling.

### *13.1. Simulation details*

The considerations of Roscoe (1999A) show that, to an extremely high level of statistical precision, optical rotation curves conform to a power-law, $V = A\,R^{\alpha}$ for $R > R_{\rm min}$ where $R_{\rm min}$ defines the exterior limit of the dynamical effects of the bulge on the disc.

Given this, and in order to preserve maximum authenticity otherwise, we used the existing PS sample as the basis for generating simulated unfolded rotation curves as follows:

- fold the PS sample using the auto-folder and record the set of parameters, ($\Delta R$, $V_{\rm sys}$), which define each folding solution. Here, $\Delta R$ is the offset between MFB's estimated dynamical centre and the auto-folder's estimated dynamical centre, whilst $V_{\rm sys}$ is the auto-folder's estimated systematic velocity for the galaxy concerned;
- record the sampling radii, $R_i$, $i = 1, .., N$ on each rotation curve, and note which is on the approaching arm and which is on the receding arm;
- at each radial position, $R_i$, $i = 1, .., N$, on any give rotation curve, record the value of the cross-correlation coefficient, $\rho_i$, used by MFB to indicate the quality of velocity measurement at $R_i$ - see Sect. 2.1;
- calculate $A$, $\alpha$ and $R_{\rm min}$ for each rotation curve;

- for $R_i > R_{\min}$, form $V(R_i) = A\,R_i^{\alpha} + \epsilon_i,\ i = 1,..,N$ where $\epsilon_i$ is a normal random variable defined so that 95% of its values lie within the interval $[-a, +a]$ km s$^{-1}$ when $\rho_i = 0.4$, and within the interval $[-a/6, +a/6]$ km s$^{-1}$ when $\rho_i = 0.9$. Here, $a$ is a positive parameter with a typical value of 20 km s$^{-1}$;
- in the absence of any quantitative knowledge in the region $R < R_{\min}$, we simply form $V(R_i) = B\,R_i^{\beta} + \epsilon_i$ in this region, where $B$ is chosen to ensure continuity of velocity across $R = R_{\min}$, and $\beta$ is a normal random variable chosen so that 95% of its values lie within the interval $[0, 1]$;
- unfold the rotation curve by assigning to each velocity, $V(R_i)$, $i = 1,..,N$, a positive or negative signature according to whether the sampling radius, $R_i$, is on an approaching or receding arm;
- make a random selection of ($\Delta R$, $V_{\rm sys}$) from the previously computed list, and transform the unfolded rotation curve according to $R_i \rightarrow R_i + \Delta R$ and $V(R_i) \rightarrow V(R_i) + V_{\rm sys},\ i = 1,..,N$;
- apply the auto-folder to the folding of these simulated rotation curves.

### *13.2. Simulation results*

When the velocity noise parameter, $\epsilon_i$, is set to zero, then the auto-folder folds all of the simulated rotation curves (864 of them) exactly, to within rounding errors. This demonstrates that the general logic of the underlying algorithm is working correctly.

Typically realistic simulated rotation curves resulted when the velocity noise parameter, $\epsilon_i$, was defined by setting a = 20 km s$^{-1}$ in the foregoing prescription. The results of applying the auto-folder to the folding of a sample of these noisy curves are shown in Figs. 9, 10 and 11 respectively. In all cases, it is seen that the auto-folder solution has the appearance of being reasonable in the sense that it is not obvious that the exact solution is necessarily better. However, a detailed comparison of the exact solutions with the auto-folder's solutions shows that, in general, the auto-folder's solutions improve as the sampling rate increases; this is a reasonable and expected result.

## 14. Conclusions

We have successfully developed a robust and accurate automatic rotation curve folder, and the crucial resources underwriting this success were the PS data base of 900 folded rotation curves and the corresponding raw data of MFB. Additionally, there were five key points in the development:

- The knowledge, derived from an analysis of the PS sample (Roscoe 1999A) that the exteriors of optical rotation curves conform to power laws, $V/R^{\alpha} = \text{const}$, to a very high statistical precision. This information is fundamental to the mode-choosing strategy (see Sect. 7.1) and the hole-cutting process;
- The availability of a measure of internal consistency in velocity measurements in MFB's data base of unfolded rotation curves;
- The identification of an appropriate, Fourier-based, minimization principle;
- The recognition that the hole-cutting algorithm effectively partitions rotation curves into two dynamically distinct regions - the interior one of which has been identified as a dynamical transition region;
- The recognition that the class of transition regions on rotation curve interiors appears to be partitioned into intrinsically quiet and intrinsically noisy subclasses.

Furthermore, the intrinsically quiet and intrinsically noisy transition regions referred to above appear to be preferentially associated with distinct regions along the $\ln A$ axis of Fig. 6; this suggests that a study which compares the archived CCD images of the galaxies observed by MFB with the positions of those galaxies in Fig. 6 might be a very worthwhile undertaking.

Finally, Appendix E shows that, in practice, there is no advantage to be gained by using more than five Fourier modes for the folding process, whilst the final Appendix makes various miscelleneous comments of a detailed nature.

*Acknowledgements.* I would like to thank Vince Ford, of MSO, for his help in making available the original unreduced MFB rotation curve sample without which this work would not have been possible.

## Appendix A: A non-integral method, and its problems

An obvious alternative method for folding rotation curves can be described as follows: Suppose that $U(r)$ denotes the estimated magnitude of velocities on the approaching arm whilst $V(r)$ denotes the estimated magnitude of velocities on the receding arm, where $r$ is the radial displacement from the estimated dynamical centre of the galaxy. Then, for an exactly known systematic velocity, $V_{\rm sys}$ and dynamical centre, $O_{\rm dyn}$, of a given galaxy, there is:

$$\sum_{i=1}^{N} \left(U(r_k) - V(r_k)\right)^2 = 0.$$

It follows that, in principle, rotation curves can be folded by minimizing the functional

$$F \equiv \sum_{i=1}^{N} \left(U(r_k) - V(r_k)\right)^2$$

with respect to variations in $V_{\rm sys}$ and $O_{\rm dyn}$.

However, this method makes *direct* use of velocity and radial displacement measurements which are both intrinsically very noisy. By contrast, the Fourier method described in this paper is based on minimizing a functional

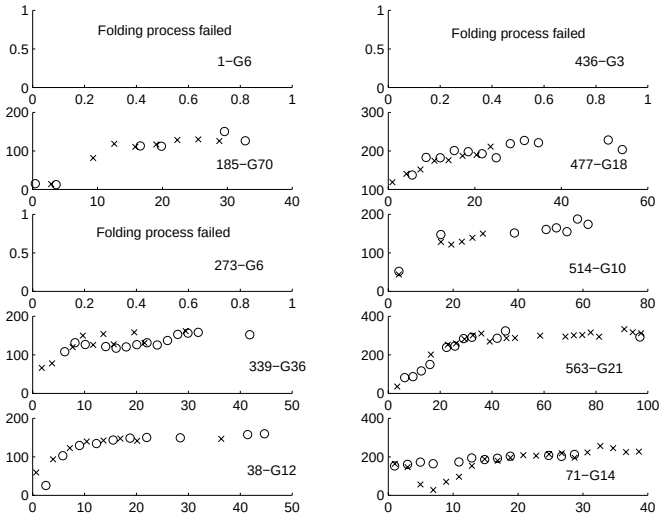


**Fig. A1.** Solutions for the rotation curves of Figs. 7 and 8

defined over the Fourier coefficients $A(m)$ which are estimated by integrating over the noisy velocity data - in other words, the method has an intrinsic smoothing mechanism incorporated within it. A recognized consequence of defining functions directly over noisy data is that they frequently fail to exhibit the mathematical properties of minimum points, even where such points are known to exist. In such a situation any minimizing algorithm designed to locate minimum points will fail. In the present case it is therefore to be expected that the functional $F$ will frequently fail to exhibit the mathematical properties of a minimum point in the region of the actual $(V_{\rm sys}, O_{\rm dyn})$ solution.

In practice, when applied to the PS sample, the algorithm failed to find a sensible solution about 30% of the time. To illustrate this point, Fig. A1 shows the results of applying the algorithm to the ten rotation curves shown in Figs. 7 and 8. There are three total failures and one (71-G14) very poor solution. The remaining six solutions compare favourably with both the PS and the MK IV auto-folder solutions.

## Appendix B: The hole-cutting data reduction process

The basic tool which we use to measure the success, or otherwise, of our developing auto-folder is the distribution of $\ln A$ over the whole PS sample. However, an accurate estimate of $\ln A$ for each rotation curve requires the application of the "hole-cutting" technique, which aims to remove that interior part of the rotation curve which is dominated primarily by the bulge. The details of this technique are given in Roscoe (1999A), and its effect on the data of the PS solution are summarized in Table B1 which lists the average and standard deviation of the 900 mean square residuals which arise from linearly regressing $\ln V$ on $\ln R$ for each of the folded rotation curves of the PS data base, before and after the application of hole-cutting.

**Table B1.** Effects of hole-cutting on power-law fits

| Hole-cutting | Mean rms | Std Dev |
|---|---|---|
| Before | $8.4\ 10^{-2}$ | 0.16 |
| After | $2.7\ 10^{-2}$ | 0.10 |
| | $N\ =\ 900$ | |

Table B1 makes it clear that the deviation of rotation curve data from the simple power law model is very much concentrated in the innermost regions of the rotation curves, and thereby provides the empirical justification for the hole-cutting technique.

## Appendix C: The cut-out regions as physically coherent transition regions

The hole-cutting technique is based upon the assertion that the interior parts of rotation curves are transition regions between bulge-dominated and disc-dominated dynamics, and Table B1 provides strong evidence in support of this.

Additional strong evidence is provided by Sect. 4 of (Roscoe 1999A) Table 1 there shows that the pre-hole cutting values of $R_{\rm min}$ act, with near certainty, as noisy traces of optical radii. Similarly, Table 2 there shows how the post-hole cutting values of $R_{\rm min}$ also act, with similar near certainty, as noisy traces of optical radii. Since the optical radii themselves are very likely dependent upon bulge radii, then a natural conclusion is that the pre-hole cutting *and* post-hole cutting values of $R_{\rm min}$ provide estimates of the boundaries of bulge radii effects.

In other words, there is good evidence to support the conclusion that the two values of $R_{\rm min}$ provide estimates to the interior and exterior boundaries of physically coherent *transition regions* between bulge-dominated and disc-dominated dynamics.

## Appendix D: An uncalibrated indicator of folding accuracy

One striking, and so far uncommented, feature of the results of the Sect. 4 (Roscoe 1999A) analysis is the very large increase in the degree of $(R_{\rm min}, R_{\rm opt})$ correlation that arises from the hole-cutting process: specifically, we see that the $t$-ratio statistic for the coefficient of $R_{\rm min}$ in the model $R_{\rm opt} = b_0 + b_1 R_{\rm min}$ increases from 14 to 24 whilst the coefficient of determination increases from 19.2% to 38.6%. The significance of these numerical changes is given

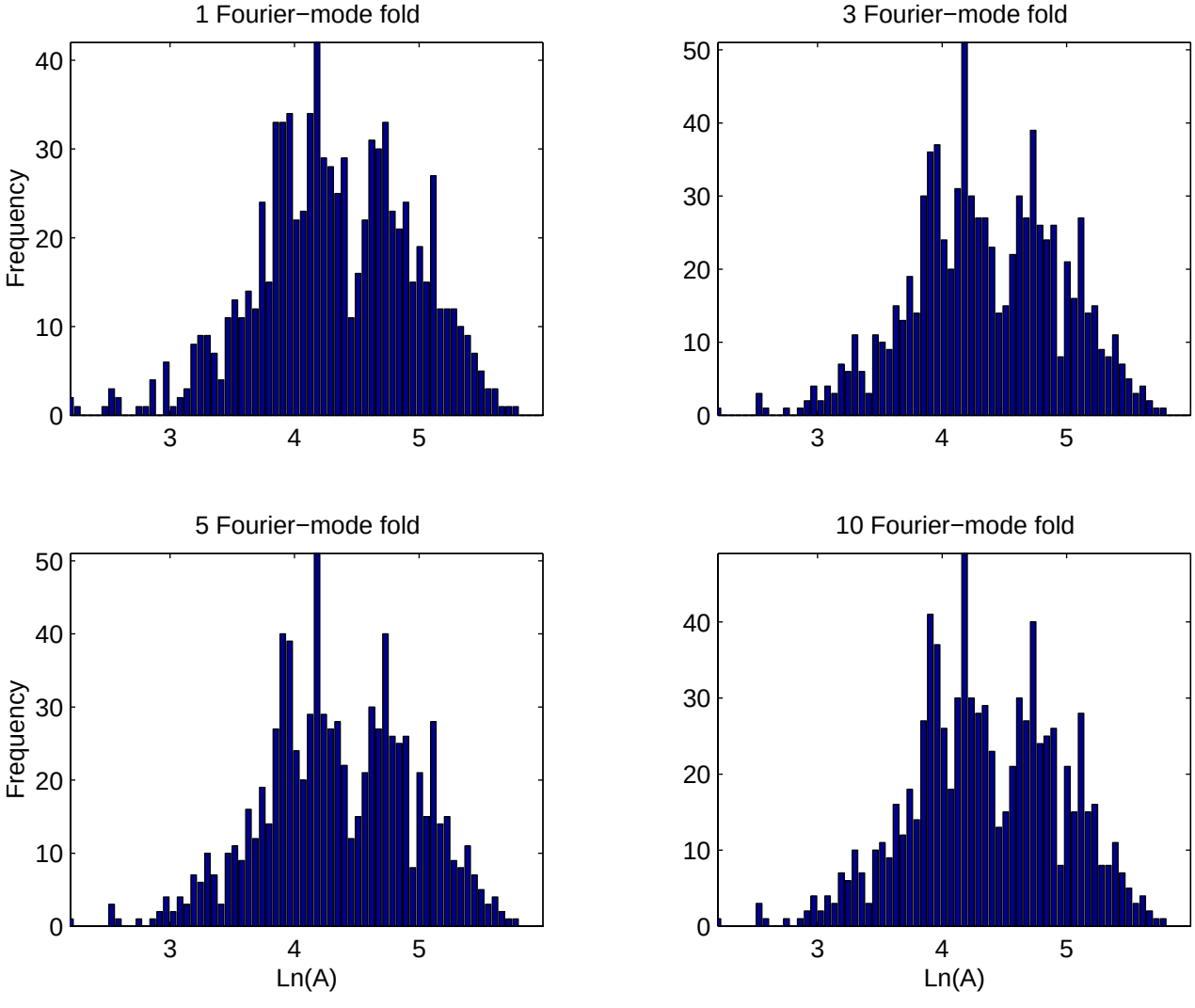


**Fig. E1.** Convergence of Fourier modes in the Mk IV auto-folder applied over the whole PS sample

a startling visual impact by the "before and after" diagrams of Figs. 2 and 3 in that paper.

In a qualitative sense, this result is to be expected for the following reasons: given that, after any folding process, *absolute* errors will inevitably exist in the estimated position of the dynamical centre of any given galaxy, then there will be absolute errors of similar magnitudes in the pre-hole cutting and post-hole cutting estimates of $R_{\min}$. Since the pre-hole cutting estimate of $R_{\min}$ is less than the post-hole cutting estimate of $R_{\min}$, it follows that the corresponding percentage errors in the pre-hole cutting estimates of $R_{\min}$ will be greater than those in the post-hole cutting estimates of $R_{\min}$. Consequently, given that $(R_{\min}, R_{\rm opt})$ correlations exist in the first place, these correlations must automatically *increase* through the hole-cutting process.

The real significance of the foregoing results is therefore not that the $(R_{\min}, R_{\rm opt})$ correlations increase at all through the hole-cutting process, but is rather that the magnitude of this increase acts potentially as an uncalibrated statistical measure for the magnitude of the original *absolute* errors in the estimated positions of the dynamical centres taken over the whole sample.

Thus, we might expect that relatively large changes in $(R_{\min}, R_{\rm opt})$ correlations through the hole-cutting process will indicate the presence of relatively large errors in the estimates of $O_{\rm dyn}$ over the set of folding solutions, and vice-versa. These considerations play a significant part in the developments from Sect. 8 onwards.

## Appendix E: Convergence of auto-folder solutions as the numbers of included Fourier modes increase

The actual number of modes used for the folding of any given rotation curve is determined by the mode-choosing strategy of Sect. 7, and this strategy assumes that this number has a maximum limit determined by Sect. 6.2.

However, we find that the folding solution over the whole set of rotation curves (defined as the $\ln A$ distribution) can be considered converged for all practical purposes when a maximum of only *five* Fourier modes are included. The illustration of this fact is given in Fig. E1, which displays the solutions for the four cases of the 1-mode solution, the 3-mode solution, the 5-mode solution and the 10-mode solution. It is quite clear that the solution converges extremely rapidly.

In practice, all the Mk II, III and IV auto-folder solutions shown in the earlier figures have been computed using just five Fourier modes.

## Appendix F: Miscellaneous comments

### *F.1. The damped Mk IV auto-folder*

The basic Mk IV auto-folder gives excellent folds for most rotation curves, and all of the Mk IV solutions displayed have been generated using from basic model. However, for a very few individual rotation curves, it gives folds which are manifestly poor, and the problem always manifests itself in the form of allowing the estimated position of $O_{\rm dyn}$ to be greatly in error. It is found that replacing the standard function-to-be-minimized,

$$F(M) = \sum_{m=0}^{M} [A(m)]^2$$

by

$$F(M) = \sum_{m=0}^{M} [A(m)]^2 + 4(\Delta O_{\rm dyn})^2,$$

where $\Delta O_{\rm dyn}$ is the calculated correction (measured in arcseconds) to the initial estimate for the position of $O_{\rm dyn}$, damps the overgrowth of $\Delta O_{\rm dyn}$ in the few pathological cases, but without having any marked effect on the well behaved majority.

### *F.2. Missing rotation curves*

Although the PS sample contains 900 folded rotation curves, the corresponding data base used here to develop the auto-folder only contains data for 877 (unfolded) rotation curves. The discrepancy of 23 rotation curves has arisen because the corresponding galaxies were supplied with incomplete information in one of the source databases provided by MFB.

### *F.3. Failure to fold thirteen rotation curves*

Of the 877 rotation curves, the auto-folder failed to fold thirteen. Investigation showed that this was because, after choosing only those velocities for which MFB's cross-correlation coefficient exceeded 0.4 (see Sect. 2.1 concerning data quality) eleven of the thirteen rotation curves concerned had no velocity measurements on one side of the rotation curve, whilst two had only one measurement.

## References

Mathewson D.S., Ford V.L., Buchhorn M., 1992, ApJS 81, 413
Nelder J.A., Mead R., 1965, Comp. J. 7, 308-313
Persic M., Salucci P., 1995, ApJ 99, 501
Roscoe D.F., 1999, A&A 343, 788 (= Roscoe 1999A)
Roscoe D.F., 1999, A&A 343, 697 (= Roscoe 1999B)