# A Procedure for Preliminary Reduction of Bandwidth Synthesis Data

George Purcell
Tracking Systems and Applications Section

*The procedure described here provides a fast, flexible, and inexpensive way to reduce and evaluate bandwidth synthesis observations before committing them to a more complex general-purpose fitting program. It enables the user to apply various corrections to the data, to resolve integer-cycle ambiguities, to calculate preliminary values of the baseline vector and source positions, and to assess the quality of the observations.*

## I. Introduction

Reference 1 outlines the first steps in a procedure for analyzing bandwidth synthesis observations in radio interferometry. At present, a typical observing session for this purpose includes perhaps twenty separate observations of ten compact extragalactic radio sources, made simultaneously at two or more stations. Each observation lasts about ten minutes and contains data in two or more (usually three) S-band frequency channels, which are sampled sequentially at intervals of one second. Each channel is nominally 24 kHz wide, and the channel-to-channel separations are at present normally 2, 8, and 10 MHz. The observations at each station are recorded digitally at 48,000 bits per second on standard 1.27-cm (1/2-in.) computer tapes, which are later brought together for analysis.

After the initial steps of cross-correlation, fringe-stopping, and phase-tracking, discussed in Ref. 1, the observer has, for each observation on a given baseline (that is, between two particular stations), the following intermediate results:

(1) From each channel, a measurement of the residual fringe rate

(2) From each channel, a measurement of the residual delay based on the bitstream alignment performed during cross-correlation (in the present system a relatively crude measurement, not used in the subsequent analysis)

(3) From each *pair* of channels, another measurement of the residual delay obtained by bandwidth synthesis. These measurements contain integer-cycle ambiguities (see formula 27 and the following discussion in Ref. 1) that have to be resolved before the data can be analyzed further.

The "residual" delays and fringe rates listed above are in each instance the difference between an observed value and the value predicted for that measurement by a model. This model is an analytic function of the presumed values of the baseline vector, the direction to the radio source, and the properties of Earth's troposphere. Consequently, the residuals can be used to compute corrections to small errors in the original estimates of the model parameters. Although the functions involved are highly nonlinear, the parameters are in general so well known that the residuals are very small. In this case, a linearized model is entirely adequate, and a straightforward application of the method of least squares yields the desired corrections.

To obtain the best results with this method, it is desirable to reduce simultaneously a large number of observations, comprising numerous observing sessions and many sources. A fitting procedure that performs this function will be the subject of a forthcoming report. It can solve simultaneously for an arbitrary number of model parameters including baselines, source positions, tropospheric delays, clock offsets, and clock rate offsets. Changes in the resulting baselines, in turn, can be interpreted in terms of polar motion and fluctuations in the rate of rotation of Earth. However, this sophisticated procedure is necessarily complex, time-consuming, and expensive to use. To use it efficiently, one needs also a simpler program to preprocess subsets of the data: to apply various corrections, to remove integer-cycle ambiguities, and to perform at a less refined level many of the functions of the master program. The following sections describe the capabilities and operation of a program that has been developed for this purpose.

## II. Description of the Program

The program is designed to handle conveniently the observations from a single observing session on one baseline. Its principal functions are:

(1) To resolve the integer-cycle ambiguities in the residual delays

(2) To calculate preliminary least-squares estimates of the baseline, clock parameters, and source positions

(3) To expose irregularities in the data, including bad points, discontinuities, and abnormal distributions of errors.

Table 1 lists the program's important input and output. The optional input allows the user to delete observations temporarily from the fit, to specify corrections to the parameters of the delay model, and to assign statistical weights to his initial estimates of the baseline and source positions.

The flow chart in Fig. 1 shows the principal operations performed by the program. Most of them are routine or self-explanatory, but two are not — the resolution of integer-cycle ambiguities in the synthesized delays, and the estimation of statistical errors.

For the program to resolve the integer-cycle ambiguities correctly for a particular channel pair, it is necessary for that part of each residual delay caused by errors in the model parameters to be substantially less than half the reciprocal of the difference between the two channel frequencies. Therefore the program ordinarily begins by processing the most closely spaced pair of channels (that is, the pair that permits the largest initial uncertainty in the model) and corrects the observations for all *known* errors in the original model before trying to resolve the ambiguities for that first channel pair. Having resolved the ambiguities, the program can calculate a first set of corrections to the model parameters, which it then uses to correct the observations from the channel pair with the next wider separation. Since the errors in the delays decrease as the separation between channels increases, the program can thus proceed iteratively to more and more widely spaced channel pairs, sufficiently refining its estimates of the baseline and source positions at each iteration to allow resolution of the ambiguities in the next iteration. If this scheme should break down, it is also possible to override the automatic procedure by specifying predetermined integer-cycle corrections in the input stream.

The other unusual feature of the program is the way in which it assigns errors and statistical weights to the data. (For a more thorough discussion see Ref. 2.) The process of phase-tracking produces along with each residual delay or fringe rate an estimate of its uncertainty due to sources of error that operate at relatively short time scales, up to a minute or so. These errors normally are entirely dominated by system noise with an autocorrelation time much less than a second. However, there are other sources of error, including some instrumental drifts and changes in the propagation media, that operate on time scales from minutes to hours and can contribute significantly to the scatter in data over the length of an observing session. The program attempts to account for these errors by assigning an additional fringe rate error and an additional delay error to each channel pair, and adding them in quadrature to the system noise errors. Then, after fitting the observations in a particular channel pair, the program uses a chi-square test to determine whether the residual scatter is consistent with the assigned total errors. If not, the program adjusts its estimates of the additional errors accordingly, recomputes the statistical weights, and repeats the fit.

The program uses the data from the last channel pair — that is, the one with the widest channel separation — to calculate its best estimates of the baseline vector, clock parameters, and whatever source positions it was asked to find, along with estimates of the errors in those quantities. By comparing these results with other measurements, and by examining the printer plots of the residuals after fitting, the user can then gauge the quality of the observations and determine possible sources of difficulty for the master fitting program. If he wants to, he can run the program again with dubious data deleted, or with different constraints on the parameters of the model.

## III. A Sample Experiment

As an example of the use of the procedure, consider the observations made in Spain at DSSs 61 and 63 between UT 0500 and 0900 on January 9, 1976. There were eleven observations of five sources, and the channel frequencies were 2285, 2287, and 2295 MHz.

Figure 2 shows the residual delays at successive stages of the reduction. (The residual fringe rates, not shown, follow a similar progression except for the resolution of the integer-cycle ambiguities.)

Figure 2a shows the raw data for the most closely spaced pair of channels, at 2285 and 2287 MHz. The large scatter is the result of an intentional error of about 50 meters in the z component of the a priori baseline, along with smaller errors in the x and y components. Despite the large size of this baseline error, no compensatory corrections were necessary, and the program proceeded directly to resolve the integer-cycle ambiguities. For this pair of channels, 2-MHz apart, the cycle time is 500 nanoseconds, and so an observed difference of at least 250 nanoseconds between successive observations would be required to induce a correction. Since the largest change turned out to be only about 115 nanoseconds, however, no corrections were made. (Notice that if the error in the baseline had been a little more than two times as large, the largest difference would have exceeded 250 nanoseconds, and a "correction" would have been applied — incorrectly. For such a large baseline error it would have been necessary to apply at least a partial baseline correction to the data before resolving the integer-cycle ambiguities.)

The least-squares fit of the data in Fig. 2a gave a clock offset of 247 nanoseconds and corrections of 6.0, 2.6, and -48.7 meters to the x, y, and z components, respectively, of the baseline. Figure 2b shows the residuals after fitting of the data in Fig. 2a. The error bars here (and in Fig. 2f) are the estimates of system noise derived from phase tracking. Since the residuals are small, no additional sources of error had to be invoked.

The results for the 2287 and 2295 MHz channel pair are not shown. Figure 2c gives the raw data for the last channel pair to be analyzed, 2285 and 2295 MHz, with a separation of 10 MHz and a cycle time of 100 nanoseconds. Here again, the large scatter is due almost entirely to the incorrect a priori baseline. Now, however, the program can use the baseline corrections computed for the previous channel pair (2287 and 2295 MHz) with the results shown in Fig. 2d. Here the total scatter is even larger (note the change in scale between plots), but the distribution of delays is almost discrete, with values separated by 100 nanoseconds. The program then resolves the integer-cycle ambiguities and obtains the final residuals before fitting shown in Fig. 2e. Finally, Fig. 2f shows the residuals after fitting. The improvement between Figs. 2e and 2f is not dramatic, because the baseline corrections computed for the previous channel pair were already quite accurate.

# References


1. Thomas, J. B., "An Analysis of Long Baseline Interferometry, Part III," *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. XVI, pp. 47–64. Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1973.

2. Thomas, J. B., et al., "A Demonstration of an Independent-Station Radio Interferometry System with 4-cm Precision on a 16-km Baseline," *J. Geophys. Res.*, Vol. 81, pp. 995–1005, Feb. 1976.

**Table 1. Input and output of fitting program**

| |
|---|
| **Required input (punched cards)** |
| A priori baseline (the one used for cross-correlation) |
| Channel frequencies and pairings |
| Intermediate results for each observation (punched by previous step of the reduction): |
| Identification, date and time, source name |
| Observed residual delays and associated errors |
| Observed residual fringe rates and associated errors |
| Partial derivatives of observed delays and rates with respect to model parameters (for least-squares design matrix) |
| Program-control options (control use of optional input and output) |
| **Optional input (punched cards)** |
| Corrections to a priori baseline |
| Corrections to a priori source positions |
| Corrections to a priori estimates of polar motion and UT1 – UTC |
| Corrections to tropospheric delays at zenith |
| Arbitrary corrections to the observed residual delays |
| Estimates of errors in the a priori baseline and source positions |
| Initial estimates of scatter in the data due to sources of error other than system noise |
| Predetermined integer-cycle corrections to the observed residual delays |
| List of sources for which corrected positions are to be calculated |
| List of data to be deleted from the fit |
| **Output for each channel pair (line printer)** |
| List of active program-control options |
| Estimates of observational errors other than system noise both before and after fitting |
| Summary of the observations |
| List of data deleted from the fit |
| Table of raw data, corrections, and corrected data |
| List of phase-turn corrections, predetermined or computed by the program |
| List of least-squares adjustments to the fitted parameters |
| Table of residuals before and after fitting |
| Printer plots of the delay and rate residuals after fitting |
| Fitted values of the baseline, clock parameters, and source positions, and their errors |

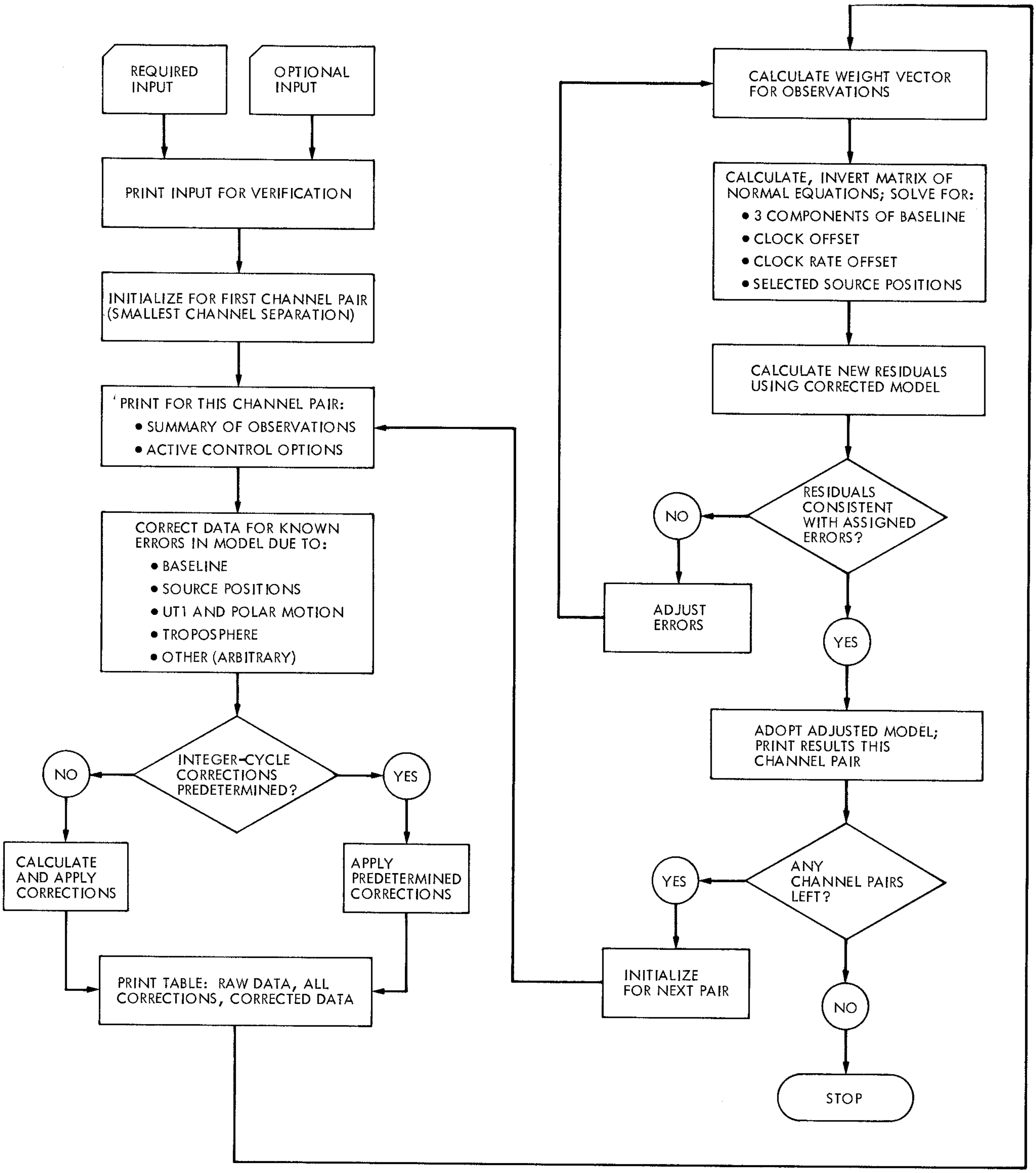


Fig. 1. Flow chart of fitting program

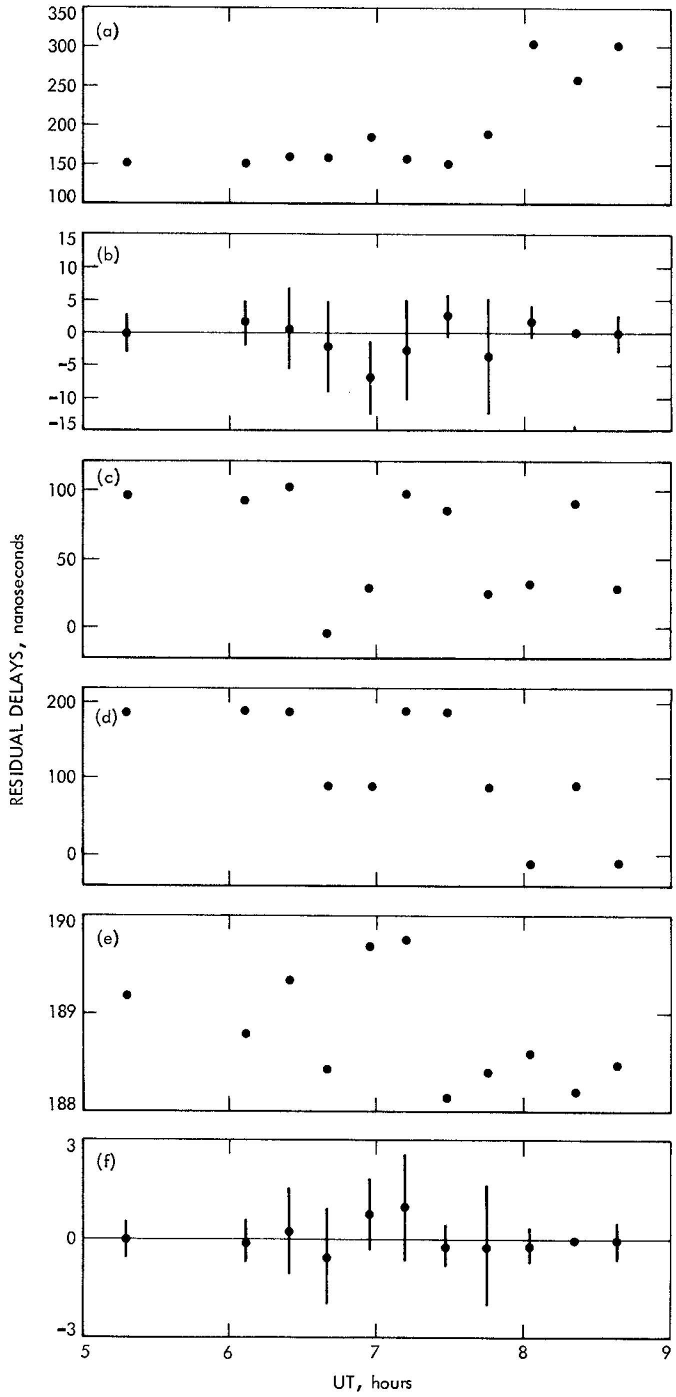


Fig. 2. Residual delays at successive stages of a sample reduction: (a) raw data for channel separation of 2 MHz; (b) residuals after fitting for channel separation of 2 MHz; (c) raw data for channel separation of 10 MHz; (d) data after baseline correction, channel separation of 10 MHz; (e) data after integer-cycle correction, channel separation of 10 MHz; (f) residuals after fitting for channel separation of 10 MHz.