NDAC/LO/132
May 31, 1990

HIPPARCOS reductions for multiple stars, XIII
by S. Söderhjelm, Lund Observatory

*This note gives a more concrete description of the programs for the CHF-generation than the previous outline in NDAC/LO/121. The whole process is now rather completely specified, with most of the programs finished.*

1. Introduction

As described in NDAC/LO/121, the input to the CHF-generation has three large components. First, there are the saved FOVC-data ($\beta$-values with information matrices) from the whole mission or a considerable part (~1 yr) of it. Secondly, the final OGAR attitudes are used, including the spin-phases derived in the Great-circle reductions and the set zero-points from the Sphere solution. The third major input is the instrument-parameters (normally determined once per set in the Great-circle reductions), describing the large-scale field-to-grid transformations. Also needed are a photometric calibration of the IDT-data over the whole field-of-view (varying with time), a subset of the Input Catalogue, and the goodness-of-fit histograms from the IDT preprocessing. These latter ultimately define which objects will have their observations saved in Case History Files for further double-star analysis.

The present note concerns mainly the 'preparatory' software needed before the main CHFDAT and CHFGEN modules may be used. First, a special 'system' catalogue (ICSPEC) has to be created in a direct-access disc file. The main purpose is to associate those IC-entries together that will form a double or multiple system in the DS reductions, and it defines for each such system the 'reference point' astrometric parameters. This seemingly trivial task is made complex by the great variety of multiple star systems in the Input Catalogue, and some of the clarifying work was reported in NDAC/LO/131. In practice, a preliminary version of this catalogue is created by the program(s) MKCAT1/2. A 'final' version is then made by the SELSYS1 program, using suitable limits in sep/$\Delta$m (for the known doubles) and in the histogram parameters (to define suspected new doubles). The SELSYS2 program then uses this system-selection in copying the relevant parts of the FOVC-data. Finally, the MKCAL-programs collect and reformat the necessary calibration-data, including the frame-by-frame attitudes.

2. The MKCAT programs

In 'mode 1' of MKCAT1, the following data are copied from the (tape) Input Catalogue to a direct-access disc-file (with recno=IC #): CCDM#, IC multiplicity/orbit/variability codes, H-mag, B-V, astrometric parameters (with epoch). The IC associated star(s) are used to make the a priori division into double or multiple systems, and the reference point for each system is given simply by the IC astrometric parameters for its primary. (In NDAC/LO/121, I suggested using the photocentre of the system as reference point, but this is an unnecessary
complication. The difficult part in any case is to estimate each component's offset from the reference point). In 'mode 2', data from the Annex1 catalogue (put in a direct access file by the program MKANX1) are used to add (mainly) the separation and magnitude-differences for 'close' systems (single IC entries). (A 'mode 3' changes the multiplicity-codes for a very few (9) systems where three IDT-pointings are definitely superfluous. For practical reasons, this is done before the main SELSYS-selection). The MKCAT1 program may be run at any time, because it uses only data from the Input Catalogue.

Later, when all the observations are available, the MKCAT2 program reads the histogram statistics. These frame-wise goodness-of-fit parameters are collected in eight bins (with approximately equal counts expected for single stars), and the program transforms the 8 integer counts into two or three 'histogram-parameters', which are then stored in ICSPEC. So far, I use the parameters \( \log(C_H) \) and \( t \) as defined in NDAC/LO/114, but any better choice may be substituted. Because these parameters depend slightly on the total number of frames observed, this number is also stored in ICSPEC.

In MKCAT2, the photometric data are also updated with data from the SM-observations, because these are available on the same 'STAR'-files as the histogram data. The astrometric data are not updated, because they define the nominal IDT-pointings actually used in the observations. (The SM astrometric updates may be useful later, however, in providing start positions in the actual double star solutions).

3. The SELSYS programs

The ICSPEC file now contains data for the 'known' doubles and multiples listed in the Input Catalogue, but several of these systems are (for HIPPARCOS) 'observably single'. Also, some of the IC 'singles' are in reality double, and part of these may be identified using the histogram parameters. A new selection-programme (SELSYS1) has to be used in order to give the 'final' list of systems needed to create the Case History Files.

The histogram-parameters \( (C_H \text{ and } t) \) do not follow theoretical distributions even for simulated data (cf. NDAC/LO/113), and the double-star discovery-criteria have to be adjusted using the real parameters. For this purpose I have written two simple statistical programs (HISTAD/HISCRT) which will be run primarily on the data from the known doubles. The result of such runs will be a set of suitable parameter-limits indicating the non-single stars. For some typical (simulation) results, see the Appendix.

The main part of SELSYS1 consists of routines recognizing the different categories of doubles/multiples as outlined in NDAC/LO/131. For category 1a) and 1b) (a priori singles), the 'final category' is 1 (suspected double) if the histogram-parameters exceed the chosen limits, else 0 (single). For categories 2a) and 2c), a \( \Delta m \)-limit is introduced (cf. eq. (1) in NDAC/LO/131), and only those systems with a smaller \( \Delta m \) go in final category 2 (known doubles) immediately. The rejected systems are passed through histogram-test, and may thereafter go to category 0 or 1. (A close double rejected due to its faint secondary may have an undetected closer companion). Similarly, for categories 2b) and 2d), the rejection-criterion will be the separation. Typically, systems with above 30 arcsec separation will count as 2 singles, and again they go to category 0 or 1 depending on the
histogram-parameters. The close multiples (3a) as well as those with three IDT-pointings (3c) are certain for final category 3 (multiples), while those in category 3b) are checked in more detail. According to separation, magnitude-difference and histogram-parameters, these components or systems may get any final category. The output from SELSYS1 is a single 'final category'-number for each IC-entry, and it is written in the ICSPEC-file.

The FOVC-data (on high-density magnetic tapes) are available for all stars, and the sole purpose of SELSYS2 is to take out the 'interesting' data from these tapes. Reading through the ICSPEC-file, the category-numbers given by SELSYS1 are stored as a large array in primary memory before the FOVC-tapes are read through sequentially. The majority of the data (the '0'-systems) are simply skipped, but categories 1-3 are labelled as they are transferred to a new sequential disk-file. After this file has grown large, it may be saved on tape, and a new disc-file is then started. (To avoid a possible break in the middle of a GCR time-interval, each new 'LOFOVCnn'-file starts with a repetition of about 1 day of data).

4. The MKCAL programs

The programs for reading attitude- and calibration-data from different sources and making the files needed by CHFDAT/CHFGEN are in various states of completion. The main MKCAL1-program reads tapes (from CUO) containing both the 'geometric' instrument-parameters (determined once per set), the satellite orbit data and the final attitude data. The first two types are put in a direct-access disc-file. Because the Great-circle reductions are not necessarily chronological, and because more than one reduction may be done for a given time-interval, this 'SETCALD'-file has one record per satellite (spin) revolution, and the data for one GCR are repeated in several records. (The final chronological list of sets is defined later in the MKCAL3-program). The attitude-data are given for each frame, and it is natural to put them in a direct-access file. Because they take a lot of disc-space (32 bytes per frame = 3-400 Mb per year of obs), each CUO-tape is stored in a separate 'LOATTnnn'-file with a different zero-point for the frame-numbering. (These 'LOATT'-files are saved back on tape, and CHFDAT is run several times, with a new 'LOATT'-file on disc each time).

The time-varying photometric calibration of the IDT (including estimates of the background contribution) will be determined as part of the general photometric reductions, but no agreements about parametrisation or formats have so far been made. A fairly simple MKCAL2-program for transfer of these data to the 'SETCALD'-file is planned.

In the simple MKCAL3-program, the direct-access 'SETCALD'-file is rewritten to the sequential 'SETCAL' with one record per set (=final GCR). Also, the set zero-points determined in the Sphere solution are added from a LO-file, as are the LOATT and LOFOVC file-numbers for each time-interval.

5. Overview and conclusions

The CHF-generation process is rather involved, but the diagram below shows the major steps. The program-sizes (in lines of Fortran code) should give some indication about their complexity, while the file-sizes are for a nominal 2.5-year mission. For a revised mission, the LOFOV-files are smaller in proportion to the effective time of observation, but the LOATT-files are not diminished (one direct-access record per frame-interval even if there are no observations).
First versions of the the MKCAT(1+2)-, SELSYS(1+2)- and MKCAL(1+3)-programs are all ready, but they may not be tested and finalized until some real data is available. The MKCAL2-program awaits a better specifications of the IDT photometric calibrations. No conceptual changes have been made in CHF-DAT/CHFGEN as described in NDAC/LO/121, but the input-files and 'administration' have been changed significantly to conform to the outputs from the above preparatory programs. Partial tests of the programs have been successful, but the full-scale testing will have to be done with real data. (The existing double-star simulation programs 'shortcut' the chronological CHF-generation by calculating at once all FOV-passages for each particular object. More realistic simulations are not planned).
Appendix

While experimenting with the histogram selection-criteria, I made a series of large-scale simulations of the first year (december 1989-nov 1990) of HIPPARCOS-observations, for about 8350 'close' (single IDT-pointing) doubles. The separations and magnitudes were given their IC values, and the masses, distances and orbital elements were given 'plausible' random values as in my previous simulations. The satellite orbit and observation-restrictions were as in NDAC/LO/128, including the previously anticipated March 1990 'eclipse' gap. No detailed realism was attempted, and I added simply a few more random exclusions (40% of the time Dec-Feb, 20% after the March gap), ending up with a (conservative) 'effective' observation time of about 45%. The resulting distribution of the number of FOV-passages at different positions over the sky is rather uneven, but with the 'standard' shape with respect to the ecliptic latitude (Fig. 1).

The histogram-counts were accumulated, and the $C_H$ and $t$ parameters derived for the 8350 doubles. For practical reasons, I use $\log C_H$, and a plot versus $\Delta m_{eff}$ (Fig. 2) shows that 'single-star' values are found for $\Delta m_{eff}$ greater than about 4.5.
Using about 1800 systems with $\Delta m_{eff} > 4.0$, there are clear systematic variations with magnitude and number of observations which should be calibrated away for the real data. Roughly, however, we simply introduce a limit in $C_H$, declaring all a priori singles exceeding this limit as probably double. Qualitatively very similar behaviour is obtained for the $t$-parameter, and with a suitably chosen limits, we may have approximately equal proportions of 'true singles' falsely flagged by both criteria. After some experimenting, I have found a 'combined' criterion preferrable in most cases: only when both criteria flag a suspected double it is accepted.

Quantitative data are rather meaningless at this stage, but the simulations do give some typical discovery rates. Accepting e.g. a flagging of 2% of the true singles, some 8% of the systems with $\Delta m_{eff} = 4.4.5$ were flagged, rapidly rising to 20% at 3.5-4 and 50% at $\Delta m_{eff}$ in the range 3-3.5. These figures are very encouraging, bearing in mind the limited amount of observations. For a full mission, they will certainly be higher.