2. OVERVIEW OF THE TYCHO DATA PROCESSING

An overview is given of the main Tycho data processing up to the completion of the main Tycho Catalogue of astrometric and photometric mean values for one million stars, and the Tycho Epoch Photometry Annex for a selection of fairly bright stars. The simulations before the satellite launch and the test processing during the first part of the mission are mentioned, but the emphasis is on the final main processing. In particular the basic signal processing with digital filtering and signal detection is described in detail.

2.1. Introduction

Some main concepts and methods of the Tycho data reduction have been introduced in Section 1.6 as a complement to the following description of the various work tasks. The Tycho processing described before launch (Høg 1989) has not been changed with respect to the main architecture. The data flow for the main processing is shown in Figure 1.2. Volume 4 is divided as follows:

Chapter 3: The Tycho Input Catalogue of 3 million stars, and its role in Tycho data analysis, is described.

Chapter 4: The satellite data processing divides into two processes: 'prediction' and 'detection'. The 'prediction' process calculates the time of each slit group crossing by means of previously known star positions and satellite attitude. In the 'detection' process a background value is normally determined for intervals of 10 seconds since the background varies quite slowly. The background contains 'spikes' which are eliminated from the raw counts by means of a non-linear filtering.

Chapter 5: In 'star recognition' a mapping of the whole sky is carried out from one year of observations giving the Tycho Input Catalogue Revision.

Chapter 6: The transit identification process determines the final correspondence between each detected transit and the stars on the sky, using the accurate on-ground attitude reconstruction and the stars in the Tycho Input Catalogue Revision. The resulting identified transits serve as input for the final photometry and astrometry processing.

Further chapters describe the astrometric and photometric processing and the final catalogue production and verification.

2.2. Tycho Input Catalogue

The Tycho Input Catalogue (TIC in Figure 1.2) consists of the three million brightest stars on the sky selected at the Centre de Données astronomiques de Strasbourg, to a limit of $B = 12.8$ mag or $V = 12.1$ mag from a merging of the Hubble Space Telescope Guide Star Catalog (GSC), and the Hipparcos Input Catalogue Data Base, INCA (see Chapter 3). The accuracy of positions in the Tycho Input Catalogue at the epoch of Tycho observations is in general 1–2 arcsec rms. The catalogue contains information on stars of different categories: 40 000 astrometric reference stars, some 12 000 good photometric standard stars, double stars, variables, non-stellar objects etc.

The Tycho Input Catalogue greatly facilitated the on-ground analysis of the photon recordings and was used as input to all TDAC processes. It was however not used to control the observations on-board the satellite. The transits found in the photon record by analysis on the ground were identified with real stars by means of the transits predicted from the Tycho Input Catalogue.

Many of the stars in the Tycho Input Catalogue were too faint to be recognized, but the complete Tycho Input Catalogue of three million stars was retained throughout the mission in 'prediction' and 'detection' in order to keep a uniform data set.

2.3. Prediction of Group Crossings

The 'prediction process' was carried out at the Astronomisches Rechen-Institut (ARI) in Heidelberg by means of the Tycho Input Catalogue, the satellite's real-time attitude and position in the orbit, and a description of the star mapper geometry. The transit time for the central position of each slit group was predicted for all stars in time sequence. The resulting 'predicted group crossings' (PGC in Figure 1.2) were used as input to the detection processing of the raw photon counts provided by ESOC, the European Space Operations Centre in Darmstadt.

Three stages of 'prediction' and 'prediction updating' based on successive improvements of the attitude and the Tycho Input Catalogue are distinguished.

First Prediction

The first pass through the data used the Tycho Input Catalogue and the satellite's real-time attitude. It was therefore uncertain by about 1 arcsec.

Updating-2 by Attitude

A better estimate of the attitude was derived by the main Hipparcos data analysis groups, NDAC and FAST, for their own data reduction. The uncertainty of the first version of this attitude was about 0.2 arcsec at the 'chevron' or inclined slits. The attitude relevant for the vertical slits is directed along the scan and was determined from the 'great circle solution'. It had an accuracy of about 0.005 arcsec in all 'prediction updatings'.

The attitude was used to update the prediction. The result was used to update the 'transit summary' (TS in Figure 1.2) from which the recognition process produced the Tycho Input Catalogue Revision of some 1 million stars with accuracy about 0.15 arcsec.

Updating-3 with Tycho Input Catalogue Revision and Final Attitude

The prediction was again updated using this catalogue and the final satellite attitude with an accuracy of about 0.035 arcsec at the inclined slits. This resulted in predictions only for the recognized stars having an uncertainty of about 0.2 arcsec, including the uncertainty of the star positions. These data were used for transit identification which produced the final input for photometry and astrometry.

The satellite attitude used for Updating-2 was delivered by NDAC between December 1991 and February 1992. The final satellite attitude with uncertainty at the chevron slits of about 0.035 arcsec was received from NDAC between May 1993 and June 1994. This final attitude was originally expected after the end of mission, and the much earlier delivery helped to complete the Tycho Catalogue on the same schedule as the Hipparcos Catalogue.

2.4. Detection of Star Transits

A star appears in the data stream as four rather narrow peaks superposed on a practically constant background. An example is shown in Figure 2.1. The Tycho magnitudes B_T and V_T were derived from the signal amplitudes in the B_T and V_T channels respectively, and sometimes a T magnitude was derived from the amplitude of the added signal $B_T + V_T$. In the following, B_T and V_T refer to Tycho specific magnitudes, while B and V may be used where the distinction between the Tycho and Johnson magnitudes is unimportant (B_J and V_J are used for the Johnson B and V magnitudes where the distinction is stressed).

In order to detect the star signal and accurately estimate its amplitude and location in time it was important to determine also the background and possible non-stellar signals. Disturbing non-stellar signals were the spikes, usually having a width of only one sample.

Background

A 'format' background value was determined (see Chapter 4) from 6400 samples = 10.66... s, defined as a telemetry format. For each format the distribution functions of the counts in the two channels B_T and V_T were determined and these were in fact nearly Poissonian. The background was determined from the median value of the distribution function. Figure 2.2 shows the background in $B_T + V_T$ for 12 hours of data (roughly five revolutions) in the elliptical orbit. The background increased when the satellite was in the van Allen belts, but between the inner and outer belts the background was sometimes as low as around apogee. Information on position in the

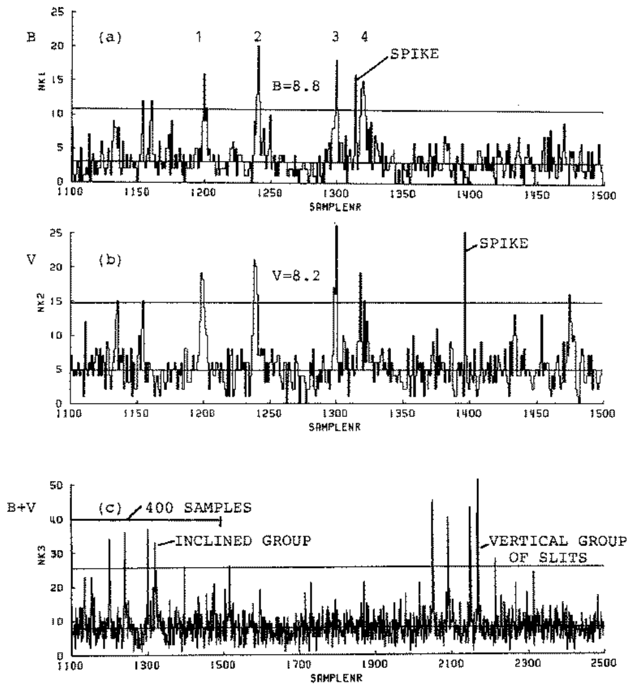

Figure 2.1. *Typical Tycho counts. (a, b) 400 samples in the B_T and V_T channels. The four peaks of a star and a few spikes are marked. Horizontal lines mark the background and the threshold for spikes in the background. (c) 1400 samples in the added counts $B_T + V_T$. The star is in fact seen crossing the 'inclined' group of 4 slits and then the 'vertical' group.*

orbit was contained in the data sets for photometry and astrometry, e.g. for eliminating less reliable observations between the two belts.

The background arose from different sources: sky background of faint stars, Cerenkov radiation, stray light from the Sun, and dark current in the photomultiplier. Nebulosities contributed relatively little because the star mapper slits were very narrow and extended over two times 40 arcmin on the sky. Table 2.1 lists the background as observed around apogee during about 40 per cent of the time, and the background expected from an analysis before launch. It appears that the apogee background was somewhat lower than expected which gave an advantage in the detection of faint stars. Stars brighter than about 8 mag could usually be detected as close to the Earth as the outer radiation belt. Sometimes, however, the outer belt extended so far out that it increased the background even at apogee.

It appears from Figures 2.2(a) and 2.2(b) that the background varied so slowly that one background value per 10 s was normally sufficient. The typical variation was about 0.3 counts per sample per 10 s, and only very rarely was it 3 times larger, even at a background as high as 40 counts per sample. The wings of stars brighter than 7.5 mag were so prominent that a 'local background' was determined based on 0.32 s of data. In this way it became possible to detect companion stars at separations about 5 arcsec and larger.

The high wing of the distribution function of counts is determined by bright stars and spikes, but the vast majority of samples appear to belong to a Poisson distribution. Significant deviations from this typical behaviour were found in the van Allen belts only, but not always. Figure 2.3(a) shows such a large deviation due to spikes. But

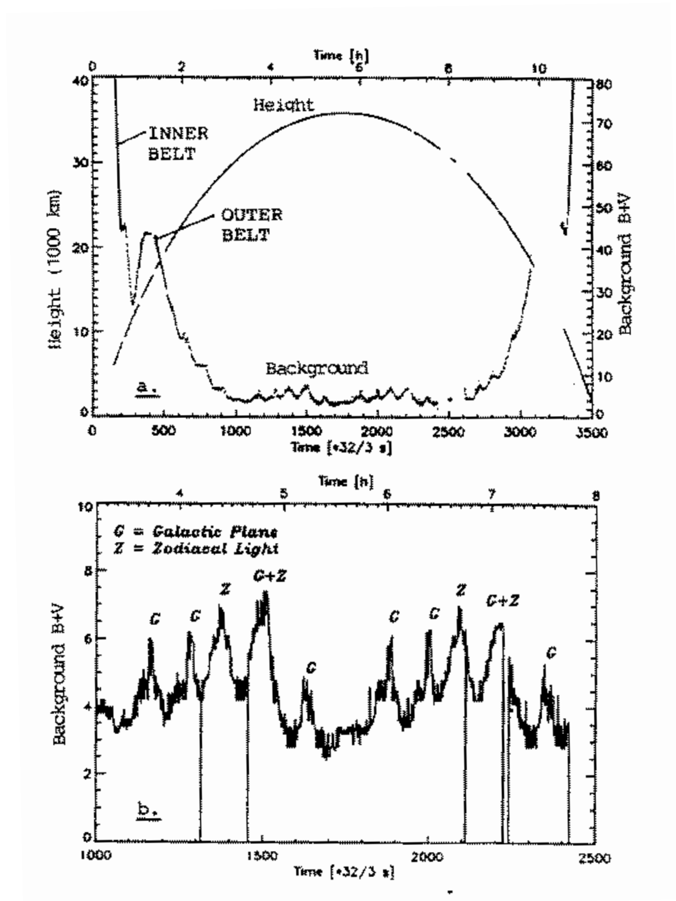

Figure 2.2. *Typical Tycho background as a function of time. (a) Top: the height above the Earth, and the background in $B_T + V_T$, during one orbit = 10.667 hours. Note the high background in the Earth's radiation belts and the missing signal when the telescope is pointing at the Earth, or when the background is high. (b) Bottom: magnified view of a small section. The background was low for about 40 per cent of the orbit around apogee. Some features are marked: G = galactic plane, Z = zodiacal light. Unit: 1 format = 10.66... s.*

Figure 2.3(b) observed 70 s earlier, also between the inner and outer belt, shows much less deviation from a Poissonian, and most of it is probably due to the presence of a dozen bright stars up to 6 mag.

Spikes

A spike was defined as a count which was significantly higher than could be expected from photon statistics in the light coming from background plus stars. In the Tycho detection process (see Chapter 4) the threshold for spike detection has been chosen so that the nominal probability was about 1/6000 per sample for a 'false spike', i.e. a spike due to pure photon statistics. The definition of the threshold took the actual, sometimes non-Poissonian distribution, into account so that the nominal false spike probability was obtained. With the typical background of 4 in B_T and V_T the spike threshold became about 12 counts per sample.

It was characteristic for spikes that they did not occur simultaneously in B_T and V_T which shows that they must be due to light generated behind the dichroic beam splitter.

Table 2.1. Background counts b. Average values observed around apogee, and expected before launch. Unit of b: counts per sample. 1 sample = 1/600 s $\simeq$ 0.281 arcsec of scan. The last two lines give, in addition, the observed maximum and minimum values at apogee.

	Observed		Expected	
	B_T	V_T	B_T	V_T
Sky background	1	1	<4	<4
Cerenkov radiation	1	1	<2	<2
Straylight, max	<1	<1	<1	<1
Dark current	<1	<1	0.5	0.5
Total	3	4	6	6
Total, minimum	0.8	1.2	–	–
Total, maximum	10	10	–	–

The duration of low spikes was usually one sample, but higher spikes may be wider, even up to 5 samples, requiring a physical phenomenon lasting a few milliseconds. This excluded Cerenkov radiation of a single electron because such an event may well generate many photons, but within a time shorter than the dead time of the detector, which was about 300 ns. The result was therefore only one count, no matter how many photons hit the cathode. The origin of the light is assumed to be fluorescence in the glass of the photomultiplier tubes caused by high-energy cosmic ray protons.

Statistics of spikes in the background between stars has been determined. Figure 2.4 shows the number, $N(H)$, of spikes in channel B_T which are higher than H counts per sample during the observing period illustrated in Figure 2.2. Thus, e.g. N(16) = 1.7 spikes per second. The statistic for V_T was very similar. The function appears to follow a power law distribution:

$$N(H) = aH^b \qquad [2.1]$$

with two different values of the exponent b in the interval $1.2 < \log(H) < 2.8$.

With the frequency distribution of spikes given in Figure 2.4 it can be shown that only a few per cent of the transits of a star were disturbed significantly and that more than 90 per cent did not suffer any disturbance at all. Thus, the spikes were not a very serious problem for the Tycho mission, but serious enough not to be neglected. Therefore, a method described in the following subsection was developed at Tübingen to eliminate significant spikes in the detection process and in the estimation of transit time and photon flux.

Detection and Estimation

The 'detection' process (see Chapter 4) treated the photon records of the B_T and V_T channels in parallel with the output from the 'prediction' process containing the predicted group crossing epochs and other information. The 'detection' process included both the detection of statistically significant transits and the estimation of their epochs and signal amplitudes.

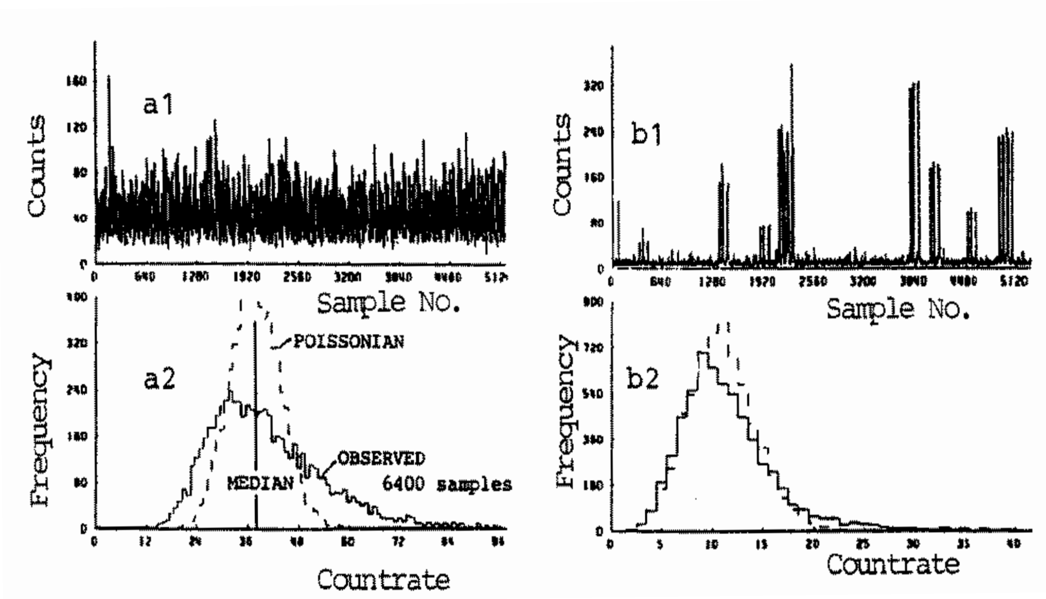

Figure 2.3. *Extreme cases of Tycho counts obtained between the inner and outer van Allen belts are shown in the upper figures, with the distribution functions below from 6400 samples. (a) Strongly non-Poissonian distribution due to spikes. (b) Many bright stars and low background, but nearly Poissonian distribution.*

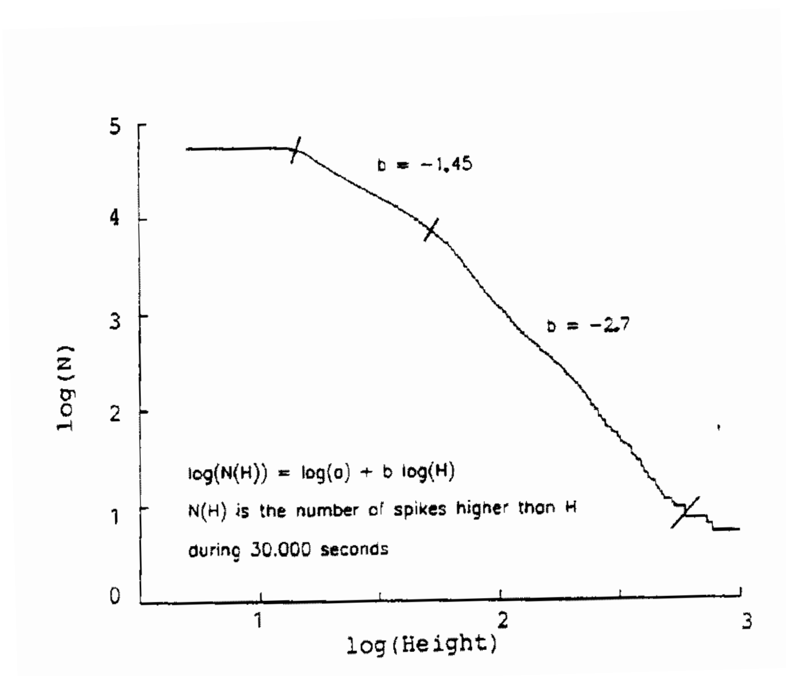

Figure 2.4. *Distribution of spikes higher than the height H for channel B_T in the background shown in Figure 2.2(a). The levelling-off at small spikes with less than 12 counts is due to the present definition of a spike rather than to the absence of low-amplitude disturbances.*

A 'predicted group crossing interval' is an interval of time corresponding to 12 arcsec centred on the predicted group crossing epoch. The true group crossing for a star will nearly always lie within the predicted group crossing interval of 12 arcsec, even if the accuracy of the Tycho Input Catalogue is 2 arcsec rms. If the detection of transits in the photon record was limited to such a small time interval, detections down to a signal-to-noise ratio of about 1.5 could be retained without getting too many false detections caused by photon noise. Simulations had shown that this limit gave about one false detection in each predicted group crossing interval of 12 arcsec, and the observations produced an average total of 1.5 detections in such an interval.

The bias in the flux of faint stars resulting from the neglect of transits with a signal below the detection limit was corrected in the photometric analysis, see Chapter 9.

The detection was carried out on the complete star mapper data stream, using the $B_T + V_T$ counts, i.e. on the combined counts from the B_T and V_T channels, which were sampled simultaneously. If a detection in $B_T + V_T$ with a signal-to-noise ratio larger than 1.8 was found, an estimation in the B_T, V_T and $B_T + V_T$ channels was made, irrespective of whether or not the detection lay within a predicted group crossing interval. This will only give a relatively small number of false detections. If the signal-to-noise ratio was smaller, but still larger than 1.5, an estimation was carried out only if the detection was inside a predicted group crossing interval. These latter estimates were recorded as 'raw transits' (RT in Figure 1.2) under the heading of the appropriate star. In cases of overlapping predicted group crossing intervals, the estimates were carried out and recorded under two or more stars of the Tycho Input Catalogue. Such multiple or double recordings occurred in about 30 per cent of cases with a Tycho Input Catalogue of three million stars.

The detections outside predicted group crossing intervals were referred to as 'serendipity detections', and may be due to photon noise or to true stars missing from the Tycho Input Catalogue, the latter referred to as 'serendipity stars'. With the signal-to-noise ratio limits given above and a Tycho Input Catalogue of three million stars, the total number of false detections was about equal to the number of true detections from stars, and this was acceptable.

The Filtering Process

The raw photon counts were subject to a filtering process which was basically linear, but became non-linear when significantly bright stars or spikes were present in the counts. The non-linear filtering had the desirable effect that significant spikes and prominent side lobes of bright stars were removed in the filtered output, whereas they would be quite disturbing for a purely linear filtering.

Figure 2.5 illustrates the detection process by means of numerical filters. The left-hand side of Figure 2.5 shows the raw counts N_k in $B_T + V_T$ of two cases where one or more bright stars are involved. The 4 peaks corresponding to the 4 slits were in the case of a linear filtering superposed by means of a 4-peak filter w_j which was equal to 1 at four correctly spaced values, corresponding to the spin velocity of the satellite.

Thus, strictly linear filtering is defined as:

$$P_k = \sum_j w_j N_{k+j} \qquad [2.2]$$

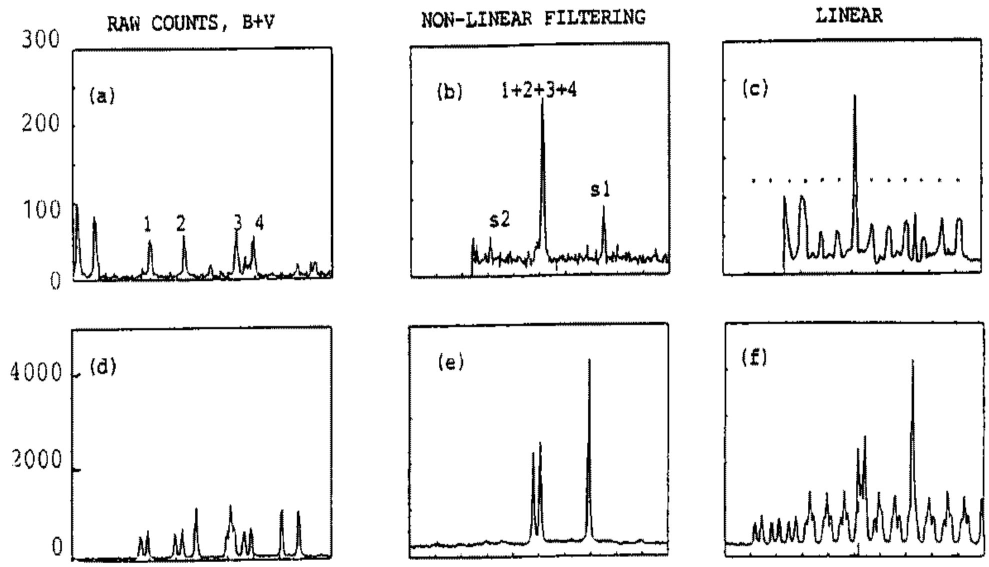

Figure 2.5. *Filtering of raw photon counts. Unit: counts per sample. Spikes and side lobes of bright stars were eliminated with a non-linear filter, but not with a linear 4-peak filter. Faint stars were in fact only subject to linear filtering. (a) Raw counts with 4 peaks of a bright star (7.5 mag). (b) Output from a non-linear filter. (c) Output from a linear filter giving a main lobe as in (b) and 12 side lobes at positions marked by dots (.) which may disturb main lobes of other stars. (d, e, f) Similarly for a complicated triple system.*

where the index of N, here $k+j$, is the number of a sample.

The result shown in Figure 2.5(c) contains the main lobe of the bright star as it should, but surrounded by several disturbing features. At the 12 positions marked by dots are side lobes, each arising when one of the filter peaks passed one of the lobes in N_k. These 1+12 lobes result from all stars in N_k.

The devastating effect of the strictly linear filtering on a bright triple star is illustrated in Figure 2.5(f) where the many side lobes disturb the detection of transits.

The improvement by non-linear filtering (see Section 4.3) in these two cases is illustrated in the Figure 2.5(b) and 2.5(e), respectively. In Figure 2.5(b) the main lobe of the bright star appears in the middle, 4 times higher than each of the lobes in N_k. The lobes from two fainter stars appear at $s1$ and $s2$. The non-linear filtering in Figure 2.5(e) clearly reveals the three stellar components. They are in fact due to ϵ Lyrae with four components: the fainter D and C are the first from left, while A and B are superposed because they are oriented along the slits.

This non-linear filtering, described in Chapter 4, is a simple modification of the linear filtering described above. It was carried out by a comparison of the 4 samples selected by the filter. If some of the samples were significantly larger than the smallest one, they were replaced by the mean value of the three other samples before they were added. The resulting bias of the amplitudes is very small due to the ensured low probability of a false spike (see above). A count may be too high for two reasons other than pure statistics: it may be a spike, or it may belong to one of the 4 lobes of a bright star. No distinction can be made between these two cases.

Table 2.2. The amplitude of a star, i.e. the counts at the centre of a slit, observed by TDAC, and expected before launch, for a star of magnitude $B_T = V_T = 10$ mag. The ratio of inclined/vertical is given, and of observed/expected counts.

Quantity	Slits	Observed		Expected	
		B_T	V_T	B_T	V_T
Amplitude (counts per sample)	inclined	3.56	2.55	3.97	1.98
Amplitude (counts per sample)	vertical	4.75	3.62	4.88	2.79
Inclined/Vertical	–	0.75	0.71	0.81	0.71
Observed/Expected	vertical	0.97	1.30	–	–

On the typical background of 4 counts per sample an additional count of 8 is 'significant' in the present context. This required a star brighter than about $B = 9.3$ mag or $V = 8.9$ mag at the centre of a vertical slit in the B_T and V_T channels, as may be deduced from Table 2.2. In $B_T + V_T$ the limit was somewhat fainter. Stars fainter than these magnitudes in B and V were consequently subject to a strictly linear filtering. The same is true for the fainter parts of the wings of brighter stars.

The net effect of the non-linear filter was to use only the undisturbed parts of lobes of a star in the observations N_k. Spikes and bright side lobes, (see Figure 2.5), did not appear in the result. Fainter spikes and side lobes which were border cases may be included in the result, and act as a slight increase of the photon noise as a result of the statistical criteria defining the filter. Another non-linear (multiplicative) filtering for star mapper data reduction was developed by F. van Leeuwen for the NDAC Consortium (Section 3.6 of Canuto *et al.* 1989). The FAST consortium determined the four transit times through each slit independently and combined them afterwards (Volume 3, Section 6.5).

The Tycho detection processing consisted of the following steps. First, derive P_k from the $B_T + V_T$ counts by non-linear 4-peak filtering. Second, locate the peaks by maximum cross correlation with a slit response function. Third, for each of these peaks or detections the signal-to-noise ratio was determined and the estimation was carried out in P_k of B_T, V_T and $B_T + V_T$ as described above.

The accurate estimation of position and amplitude was carried out with maximum likelihood estimators, though sometimes simplified in order to save computing time. The simplifications introduced a certain loss of accuracy, but never more than 3 per cent in the mean error.

Using values for the background, b, and the amplitude, a, from Tables 2.1 and 2.2 a theoretical limiting magnitude for the Tycho processing was derived, defined as the magnitude giving an average signal-to-noise ratio in $B_T + V_T$ of 1.5. For stars of the typical colour index $B - V = 0.7$ mag we obtained limits of $B = 11.6$ mag expected before launch of Hipparcos, and $B = 12.1$ mag expected from the observed count rates. This agreed with the practical limit reported by Halbwachs *et al.* (1992). There are 1.2 million stars on the sky brighter than $B = 12.0$ mag, and in fact most of these are contained in the final Tycho Catalogue.

The improvement by 0.5 mag is due to three equally important reasons. First, the more precise format background, replacing the 9-peak background (see Høg 1989). Second, the lower observed background of 7 counts per sample in $B_T + V_T$, instead of the expected 12. Third, the 14 per cent higher sensitivity of the detector system for $B_T + V_T$, which appears from the last line of Table 2.2.

Based on the inspection of the first observed data it was safe to state in June 1990 (Høg & Wicenec 1991) that the original expectations could be surpassed with the revised mission extending for 3 years, in spite of the fact that Tycho observations on a low background of less than 10 counts per sample were only obtained during about 40 per cent of the time, instead of the 90 per cent expected before launch.

2.5. Recognition of Stars

The 'recognition' processing, see Chapter 5, was based on one year of transit data. It combined the observed detections for each Tycho Input Catalogue position and performed a mapping of a small area in order to 'recognize' one or several stars in the area. In most cases no sufficiently bright star was found because a rather faint limiting magnitude was chosen for the Tycho Input Catalogue in order to account for the uncertainty of, for example, the magnitudes in the Tycho Input Catalogue.

The data were delivered to the recognition process as the 'transit summary' (TS in Figure 1.2) containing records with the transit times from 'prediction' and 'detection'. The transit times from the prediction process were at first based on the real-time satellite attitude having an accuracy about 1 arcsec. Before these transit times were used for the recognition process they were updated by means of the improved attitude from the on-ground processing of the main Hipparcos data, see 'updating-2' in Section 2.3. The attitude accuracy at this stage was expected to be better than 0.2 arcsec.

Recognition means that a mapping was performed for a circular area of 40 arcsec diameter centred on each Tycho Input Catalogue position for the first 12 months of mission data. In many cases several components were found, with a resolution of 2–3 arcsec. Eventually, however, many of these new components were found to be false, and were eliminated by use of proper measures of significance.

An additional process was the serendipity recognition mapping which searched for stars outside these small circles that were bright enough to give a rather high signal-to-noise ratio. The serendipity recognition was performed in maps of half a square degree arranged in a network covering the whole sky uniformly. In order to save computing time and to limit the amount of data, only the detections with a signal-to-noise ratio higher than 3.5 and which were detected outside small circles around the bright stars of the Tycho Input Catalogue were taken into account in this process.

The result of the recognition process was a Tycho Input Catalogue Revision of 1 270 000 stars, completed in September 1992. The accuracy of positions was about 0.15 arcsec. This catalogue contained many thousands of suspected serendipity stars and new components, but most of them were later found to be insignificant or redundant with other stars so that the final Tycho Catalogue includes only about 160 serendipity stars and 6600 new components found within 20 arcsec of the Tycho Input Catalogue positions.

2.6. Transit Identification

The ‘transit identification’ process determined which detected transits in the photon records corresponded to which stars on the sky, given in the Tycho Input Catalogue Revision.

For this purpose the predicted group crossing epochs were improved by means of the accurate on-ground attitude reconstruction and an improved geometric calibration of the slit system. Also, the more accurate positions of the stars in the Tycho Input Catalogue Revision were taken into account for the resulting ‘predicted group crossing updates-3’, see Section 2.3, to be used for transit identification.

The transit identification process started in 1993 and converted the ‘raw transits’ into ‘identified transits’ (totalling about 15 Gigabytes) to be used for the final astrometric and photometric data processing. In the process each raw transit was considered. It was included as an identified transit if it was within 3.0 arcsec of a predicted group crossing position of a star in the Tycho Input Catalogue Revision, no matter whether it agreed or not in magnitude. The further decisions to use a particular transit for astrometry and photometry were taken by these processing tasks and are discussed in more detail in the corresponding chapters. The decision in astrometry was based entirely on astrometric arguments, not on the amplitude, or magnitude, of the transit, but a statistical weight was assigned in accordance with the amplitude. During the photometric processing it was known whether a particular transit was accepted for use in astrometry by way of information supplied by the ‘All-Transits’ data set (AT in Figure 1.2). A detection was flagged as false if it was far from any star, and in this case it was not passed on to the final astrometric and photometric data processing. If other stars were close enough to disturb a transit the amplitude of these were recorded in the process ‘parasite recording’ giving information to be utilized by astrometry and photometry.

2.7. Photometry

The signal amplitudes obtained in the estimation process are measures of the stellar photon flux in the (Tycho) B_T and V_T bands, which resemble the Johnson B and V bands. The estimated amplitudes in the $B_T + V_T$ counts gave the Tycho T magnitude. This was useful for faint stars since T may be available even if B_T or V_T is too faint to be estimated. A calibration of the observed amplitudes (Scales *et al.* 1992) was required to transform to an internally consistent magnitude system, taking into account the small sensitivity variations with time, with field coordinate, etc. This was done by means of some 20 000 photometric standard stars provided by the Geneva Observatory as measured from the ground, but with calculated magnitudes in the passbands of B_T and V_T; see Section 1.5.

Provisional photometric calibrations, but not astrometric, were derived from the raw transits, which became available shortly after the observations. The final calibration required the identified transits because they were less affected by false detections and nearby transits of other stars.

An average star obtained 170 valid group crossings or 'transits' during the revised mission of 37 months, though this number was highly dependent on e.g. ecliptic latitude, due to the special scanning law of the satellite. A valid transit is a predicted transit where the satellite attitude was well known, the detector background was not too high, and the star was not too close to the ends of the slits, (see Section 1.6 for further explanation). This resulted on average in 130 accepted transit times and magnitudes per star in T, B_T and V_T for stars bright enough to be detected at every crossing. Thus, on average 25 per cent of the valid transits ($\simeq$ 40 transits per star) were rejected for various reasons: the transit time residual exceeded certain limits in the astrometric processing ($\simeq$ 10 per cent); the transit was disturbed by a parasite, i.e. another fairly bright star crossing the slits nearly at the same time ($\simeq$ 10 per cent); or the background was not determined. The Tycho T magnitude was used during the astrometric processing since transit times from the $B_T + V_T$ signal were used for astrometry. But T magnitudes are given in only 1333 cases in the final Tycho Catalogue of mean magnitudes, and only when V_T is not available. A median precision of 0.06 mag was obtained at V_T = 10.5 mag for the mean value of the measurements for each star, see Table 16.3.

The 'mean' values of B_T and V_T for each star were computed as median values for bright stars or by a special 'de-censoring' analysis in order to take into account that many transits for a faint star were censored, i.e. not detected because its signal-to-noise ratio was below the adopted limit of 1.5. The detected transits were the brightest ones, so that a simple mean or median value of these would give a biased (too bright) magnitude for the star. The de-censoring analysis, described in Section 1.6 and in detail in Chapter 9, took into account the known number of non-detections for each star and was able to reduce the bias to a negligible amount for the faintest stars.

The two independent photometric observations at the vertical and inclined slit groups were separately calibrated and reduced, and were not combined in the final Tycho Catalogue. The Tycho Epoch Photometry Annexes A and B contain magnitudes of individual transits for about 480 000 selected stars, (see Section 2.6 of Volume 1). The individual transits may be used to detect and to study variability of the stars.

2.8. Astrometry

The transit time at a slit group is a one-dimensional measure of the star position in the direction perpendicular to the slit (either vertical or inclined), provided that the satellite attitude and the grid geometry are known. In the 37 months an average of 170 valid transits were obtained, and some 130 of these were accepted for astrometry using very similar criteria of acceptance as the photometric processing, explained in the preceding section. These observations were used to derive the five astrometric parameters for each star, and their covariance matrix (see Chapter 7). A typical (median) precision of 26 mas was obtained for the position components for stars of V_T = 10.5 mag, see Table 16.2.

The astrometric quality of the results for a given star is characterized by various numbers. The formal standard error is one such number which has been shown to agree within a few per cent with the external standard error obtained by comparison with the much more accurate Hipparcos astrometric results. But for faint stars the formal error may be up to 40 per cent too small (see Chapter 18). A rather small formal standard error is however not enough to ensure that especially the faint objects are real and are undisturbed by the background or by possible neighbouring stars. Therefore a signal-to-noise ratio, F_s, of the star image was computed. F_s is defined in Section 7.4 and is a combined measure of the number of detected transits and their concentration within a distance of 1.4 arcsec of the central position. Finally, three numbers were used to define an astrometric quality index Q: the largest of the five formal errors $\sigma_{\max}$, the F_s, and the formal standard error of the single observation. This integer value from 1 to 9 (with $Q = 1$ as the best) was used to divide the Tycho stars into quality classes, see Section 7.4.

2.9. Double and Multiple Stars

Stars with separations smaller than 3 arcsec were usually not separated in the one year transit data processing (see Chapter 5), although this limit is almost ten times wider than the diffraction limit of the telescope. In particular the integration by the 0.9 arcsec wide slits limited the practical resolution. In the reprocessing, a particular effort was made in applying an adequate treatment to the double stars and to suspected double stars of an input catalogue containing 22 000 stars. The close double stars were eventually detected in different ways, according to their separations: between 1.5 and 3 arcsec, both components may be detected separately, and the individual positions were derived. In summary descriptions it is stated conservatively that pairs with separations of 2 arcsec and about equal magnitudes are resolved in the Tycho Catalogue. For separations between 0.4 and 1.5 arcsec, a duplicity could often be detected from an analysis of the magnitude versus the slit angle. In the case of a significant detection the position angle of the system was determined, modulo 180°, but is not published within the Tycho Catalogue. This method was applied to the $\simeq$ 500 000 brightest stars of the Tycho Catalogue. The special treatment of double and multiple stars is described in Chapter 14.

2.10. Five Stages of Processing

The whole analysis of Tycho data can logically be divided into five 'stages' or 'passes', starting before launch with the analysis of simulated observations. After launch the analysis began with test processing of selected pieces of data. This led to software improvements, and by March 1991 sufficient confidence had been obtained to start the mass processing of all observations obtained since December 1989.

Simulations

Realistic simulated observations were produced before launch to test all processes from reception of the satellite data, prediction, detection and recognition to photometry and astrometry. They were based on the expected satellite performance and orbit. The detection software in particular had to be almost completely rewritten after the problems of background and spikes were found in the real observations obtained in the unintended orbit. Between 1981 and 1989 there were at first simulations for each of the five processes, but later on the simulations were based on simulated data in downlink format. These were produced at the Royal Greenwich Observatory with a simulation program developed for the NDAC consortium and used on a realistic map of the sky, including high star densities, and covering a few hours of satellite scanning. Such observations, called Tycho 'data streams' were passed through all processes to ensure as far as possible the correctness of processes and the interfaces between these. The simulated data were produced by an 'instrument model' from 'true' values on which a noise model was applied. The 'true' values were provided for comparison with the 'estimated' values from each process.

Test Processing

The test processing started in TDAC in November 1989 when the first real observations were received from ESOC and it lasted for about two years, partly overlapping with the main mass processing. These and many following 'provisional ESOC tapes' contained selected pieces of data and were treated almost immediately by prediction and detection. The results, reported soon after, were relevant for improving the software of TDAC and the interfaces with ESOC. The output tapes containing 'raw transits' have been used for photometry and astrometry since December 1989.

A qualitatively new test began in October 1990 when attitude from NDAC was used to produce 'update-2' prediction. These data were needed for recognition in the mass processing (Figure 1.2), but they were also well suited for 'provisional identified transits' in the test processing. Many stars in the Tycho Input Catalogue had a positional accuracy of better than 0.5 arcsec, sufficient for most tests, but the resolving of double stars could only be tested with a Tycho Input Catalogue Revision by which final predictions (update-3 and 'real identified transits') were produced.

The provisional ESOC tapes contained 12 weeks of observations spread over 8 months of mission data, and they were the basis for the first results reported by Høg *et al.* (1992). All of these data had been processed by prediction and detection, whereas only one to three weeks of observations had been analyzed by the other tasks. About December 1990 the first 'provisional identified transits' for photometry and astrometry were produced and the recognition task received the first 'transit summaries'.

Similar test processing was carried out by NDAC and FAST and extensive comparisons between the results were made, including TDAC. Finally, in January 1991 it was decided to start ESOC's mass production for all consortia of the first 6 months of the mission.

Main Mass Processing

The main mass processing of prediction and detection lasted from March 1991 to April 1994, after the end of the 37-month mission in August 1993. The Tycho Input Catalogue Revision was completed in September 1992. This catalogue was needed for transit identification from which astrometry and photometry produced the final catalogue in early 1996.

Before launch it was planned to handle the observations in weekly batches. With the backlog in April 1991 of more than one year it was practical to work with batches of several months.

Reprocessing

The main processing of the Tycho raw data was a compromise between the wish to complete an output catalogue within a given (short) time and with the given resources of computing power and manpower on the one hand, and the wish for optimum scientific exploitation of the data on the other. Some of the drawbacks on data quality stemming from this compromise were eliminated for some groups of celestial objects by the first reprocessing, as described in Chapter 10.

For this purpose, a special list of stars to be treated again, the 'Tycho Input Catalogue Update' was compiled from the Tycho Input Catalogue Revision and other sources. It was used as input to analogues of the prediction, detection/estimation and transit identification steps of the main processing. This reprocessing of all the original raw data was a moderate effort, since the star list contained only 300 000 entries (less than 10 per cent of the Tycho Input Catalogue), and since the iterative parts of the main processing (Predicted Group Crossing updating, star recognition and most of transit identification) could be omitted. It started in September 1994 and was completed in January 1995. The subsequent astrometric and photometric reductions were done in the same way as for the main processing. End results from the main processing and reprocessing were then merged into the Tycho Catalogue, as appropriate.

The reprocessing included all solar system objects as well. The result was included in the transit identification, astrometry and photometry processing.

E. Høg