THE ASTROMETRIC INSTRUMENT OF GAIA: PRINCIPLES

L. Lindegren

Lund Observatory, Lund University, 22100 Lund, Sweden

ABSTRACT

Compared with Hipparcos, Gaia will give an enormous improvement in accuracy, completeness and number of stars: about two orders of magnitude in accuracy, four orders in number, and a completeness limit that is 12 magnitudes fainter. How is all this possible? The answer is: by a combination of many factors, the most important being bigger and more efficient detectors, and bigger optics. The method of astrometric measurements by Gaia is described from first principles, and the fundamental limitations explained in terms of physics (diffraction and photon noise), geometry, temporal sampling and reference frames. Although Gaia is basically a self-calibrating instrument, things have to be stable enough over time scales that are long enough for the calibrations to be carried out, and the corresponding requirements are outlined. To achieve microarcsecond accuracy is technically extremely demanding, but feasible with a clever and careful design of the instrument.

Key words: Gaia; Astrometry; parallax; accuracy; calibration.

1. GAIA COMPARED WITH HIPPARCOS

The goals of Gaia in terms of astrometric measurements are well known (Mignard 2005; de Bruijne 2005): a complete sky survey of point sources to 20 mag ($\sim$1 billion objects) with a target accuracy in parallax of ~ 4 μas at < 12 mag, 10–20 μas at 15 mag, and a few hundred μas at 20 mag. Annual proper motions, and positions at the mean epoch of observation, are obtained with corresponding accuracies and linked to the extragalactic reference frame. Multi-epoch, multi-colour photometry and radial velocities are an integral and essential part of the mission, but here I will exclusively focus on the astrometric measurements and try to review the most important circumstances that allow such accuracies to be achieved.

A natural starting point is the Hipparcos mission (ESA 1997), which clearly demonstrated the feasibility of conducting global astrometry from a continuously scanning space observatory. Gaia builds on this proven concept by extending it as far as current technology reasonably allows within given limits of size and mass. The Hipparcos Catalogue contains nearly 118 000 entries with associated astrometry, for which the median parallax accuracy is 1.1 mas (at the median magnitude $Hp = 8.5$). Thus, in round numbers, Gaia is expected to achieve 100 times higher accuracy for stars that are 100 times fainter than the Hipparcos Catalogue, and in total observe 10 000 times as many objects.

2. BRIEF OVERVIEW OF THE INSTRUMENT

The payload design and measurement principles were described in some detail by Perryman et al. (2001). Since then, many of the basic instrument and mission parameters have changed, in particular as a result of a major technical re-assessment in 2001–2 leading to a significant reduction of size, complexity and cost. The functionality, measurement principles and accuracy goals remain, however, almost unchanged. In the current design (cf. Figure 1), the astrometric instrument features (dimensions are along $\times$ across scan):

- two viewing directions separated by a basic angle of 99.4 deg;
- each viewing direction with a pupil of 1.4×0.5 m^2;
- a focal length of 46.7 m;
- a beam combiner superposing the two fields on a single focal plane;
- a focal plane assembly with 170 identical CCDs operating in TDI mode and covering a solid angle of 0.68 deg^2 per viewing direction (Figure 2);
- CCDs optimized for MTF and quantum efficiency, each with 4500×1966 pixels of size 10×30 μm^2 (44.17×132.5 mas^2);
- mirrors and support structure entirely in SiC;
- ultra-stable thermal environment to ensure opto-mechanical stability on < 10 μas level, including basic-angle variations;
- skymappers for real-time detection of all point sources brighter than 20 mag, including field-of-view discrimination;
- a dedicated astrometric field of 0.66×0.74 deg^2;
- CCDs with broad-band filters for photometry in 4 or 5 colours.

The characteristics of the operation are as follows:

- nominal mission length 5 years ($\geq$97% availability

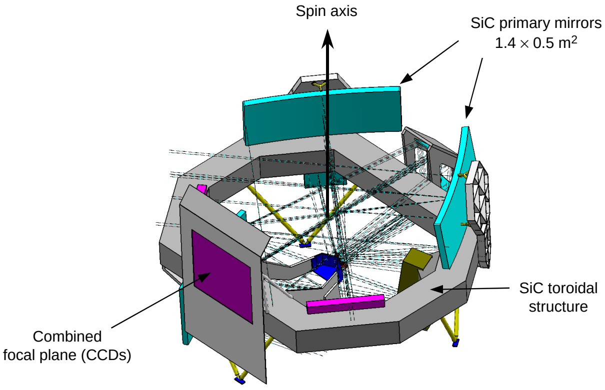

Figure 1. A possible configuration of the scientific instruments, in which all optical components, including mirrors, beam combiner and focal plane assembly, are mounted on a toroidal optical bench and enclosed by a thermal cover (not shown). Courtesy of EADS-Astrium and ESA.

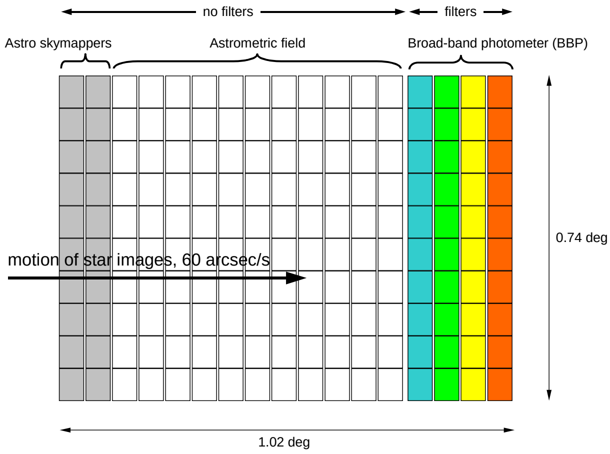

Figure 2. Schematic layout of the astrometric focal plane, comprising 170 identical CCDs in three regions: the skymappers and the astrometric field (without filters), and the broad-band photometer with interference filters in front of the CCDs.

for science);
- revolving scanning with spin axis 50 $\pm$ 0.1 deg from the Sun (solar aspect angle) and spin rate 60 arcsec s^{-1} (4 rev day^{-1});
- integration time 3.3 s per CCD, or 36 s per AF transit (along the astrometric field, AF);
- on average 90 such AF transits for any object over the mission;
- observation window of 6 $\times$ 12 pixels centred on detected objects (12 $\times$ 12 pixels for <16 mag), with on-chip across-scan binning reducing the window to a one-dimensional array of 6 or 12 samples (Figure 3);
- loss-less transmission of these samples to the ground.

The core astrometric data processing includes accurate centroiding on the AF and skymapper samples; determination of the 3-axis instrument attitude as a function of time; determination of broad-band photometry and astrometric parameters for well-behaved (i.e., apparently single and constant) stars; and the calibration of point spread functions and geometric transformations versus field position, time and colour. These core tasks are performed iteratively within a general-relativistic framework taking into account known effects to a level of ~ 0.1 μas.

3. LIMITS TO ASTROMETRIC ACCURACY

Gaia's ability to observe many more stars than Hipparcos, roughly by a factor 10^4, is simply due to the multiplexing advantage of the CCDs: on the average some 20 000 stars are observed simultaneously in the astrometric fields of Gaia, while the detector technology of Hipparcos – an image dissector tube – only allowed observation of one object at a time. In the following subsections I discuss a number of other, perhaps less obvious, issues of relevance for the astrometric accuracy.

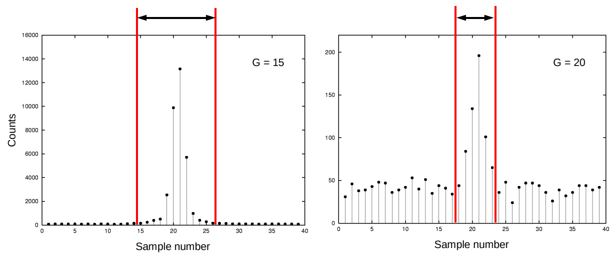

Figure 3. The CCDs in the astrometric field sample the optical image (source brightness convolved with the point-spread function) in the along-scan direction by binning the charges across scan in a small read-out window centred on the image detected in the skymappers. Along scan, these windows (indicated by the arrows) are typically 12 or 6 samples wide; only these samples are transmitted to the ground.

3.1. Physics

According to the Heisenberg uncertainty principle, one cannot simultaneously measure the position $\mathbf{r}$ and momentum $\mathbf{p}$ of a photon with infinite precision. More precisely, with Δx and Δp_x denoting the *standard deviations* of the components of these vectors in the measurement direction x, we have (Landau & Lifshitz 1977)

$$\Delta x \cdot \Delta p_x \geq \frac{h}{4\pi} \qquad (1)$$

where h is the Planck constant. We assume that the x direction is nearly perpendicular to the momentum. Since $|\mathbf{p}| = E/c = h/\lambda$, the standard deviation of the angular coordinate $\theta_x = p_x/p$ will be given by

$$\Delta x \cdot \Delta\theta_x \geq \frac{\lambda}{4\pi} \qquad (2)$$

The astrometric telescopes of Gaia have rectangular pupils with dimension $D_x = 1.4$ m in the along-scan di-

Table 1. Key quantities for comparing the astrometric potentials of Hipparcos and Gaia. The photon fluxes (in ph s^{-1} m^{-2} nm^{-1}) are representative for solar-type stars of magnitude $V = 10$ (for Hipparcos) and $V = 15$ (for Gaia). For Hipparcos, $D_x = 0.25$ m is equivalent to the actual semi-circular aperture of diameter 0.29 m. EBW is the equivalent bandwidth defined in the text. The bottom line gives the potential astrometric accuracies from Equation (4), neglecting geometrical factors, etc (see text for further comments).

Quantity	Hipparcos	Gaia
photon flux	10^4 (10 mag)	10^2 (15 mag)
aperture size, D_x	0.25 m	1.4 m
aperture area	0.03 m^2	0.7 m^2
EBW	4 nm	300 nm
total time on object	500 s	3000 s
no. of photons, N	6×10^5	6×10^7
$(\sqrt{3}/2\pi)(\lambda/D_x\sqrt{N})$	0.2 mas	3 μas

rection (the main measurement direction of Gaia; Section 3.2). At pupil level, a detected photon is consequently localized in x with constant probability density over an interval of length D_x. The corresponding standard deviation is $\Delta x = D_x/\sqrt{12}$, from which

$$\Delta\theta_x \geq \frac{\sqrt{3}}{2\pi}\,\frac{\lambda}{D_x} \qquad (3)$$

This formula is for a single detected photon; if a total of N photons are used the angular accuracy is limited by[1]

$$\sigma_x \geq \frac{\sqrt{3}}{2\pi}\,\frac{\lambda}{D_x\sqrt{N}} \qquad (4)$$

An equivalent expression was derived by Lindegren (1978).

If we are restricted to optical or near-infrared wavelengths the ultimate astrometric accuracy that can be achieved by any mission is thus limited by the linear size of the pupil (D_x) and the total number of detected photons (N) from a given object. Table 1 gives some key numbers for comparing the astrometric potentials of Hipparcos and Gaia. The 'equivalent bandwidth' is defined as EBW $= \int T(\lambda)Q(\lambda)\mathrm{d}\lambda$, where T is the total transmittance of the instrument and Q the quantum efficiency. This is a good measure of the overall efficiency of the optics plus detector, at least for solar-type stars whose spectra are relatively flat from 400 to 1000 nm when expressed in photon flux per unit wavelength. It is seen that Gaia collects, over the mission, a hundred times more photons for a 15 mag star than Hipparcos did for a 10 mag star. The total gain by a factor 10^4 is achieved through a combination of bigger aperture area (23×), larger EBW (75×), and longer integration time (6×). Together with the improved optical resolution due to the linear size of the aperture this explains the overall gain in terms of accuracy and limiting magnitude.

[1] Strictly speaking, $\Delta\theta_x$ is infinite for a rectangular pupil since the probability density function for θ_x, proportional to $\mathrm{sinc}^2(\pi\theta_x D_x/\lambda)$, only falls off as $|\theta_x|^{-2}$ for large angles. However, using a maximum-likelihood estimator for the location of the diffraction peak one asymptotically approaches equality in (4) as $N \to \infty$.

A few things should be kept in mind when interpreting Table 1. First, the potential accuracy given by Equation (4) refers to a hypothetical observation where all the photons are spent on determining a *single* coordinate x. In reality (at least) five astrometric parameters must be determined for every object, which implies an increased error on each parameter even in the ideal case. The relevant factor for the parallax will be discussed in Section 3.2. Secondly, we have assumed ideal sampling of the signal, while in reality the finite detector resolution (pixel size, electronic MTF, etc) may significantly degrade the performance. In this respect, the grid modulation technique used by Hipparcos was much inferior to the CCD detectors that will be employed by Gaia. Finally, there is of course a host of other error sources, completely ignored here, that contribute to the total astrometric error (de Bruijne 2005).

3.2. Geometry

Parallax accuracies at sub-mas level are now fairly routinely obtained in ground-based optical astrometry, e.g. Dahn et al. (2002), and much higher precisions are soon expected to be reached by interferometric techniques (Paresce et al. 2003). However, these observations are *differential* within a small field. In order to reach μas precision, e.g. for exoplanet searches, one is in practice even confined to the isoplanatic patch, i.e. within 5–10 arcsec of a suitable reference object. Using a space observatory for astrometry offers four important advantages: absence of atmosphere, weightlessness, full-sky visibility, and a thermally stable environment. All of these are crucial for *global* astrometry, based on wide-angle measurements.

Why are wide-angle measurements scientifically important? It is easy to see that they are very useful in building a coherent reference frame over the whole sky: only by directly bridging large angles (of order 1 rad) is it possible to avoid the piling up of many small errors that would result in a less rigid reference frame. Non-rigidity would for instance manifest itself in systematic proper-motion errors correlated over large scales, which could easily be mistaken for dynamical effects. However, wide-angle measurements are also essential for getting *absolute* parallaxes, as opposed to the traditional differential ground-based technique (Figure 4). This becomes more critical as we enter the μas regime, because of the lack of suitable background objects.

How do we go about making large-angle measurements with practically the same accuracy as on small angular scales? Here is the recipe tried out by Hipparcos and adopted also for Gaia:

1. Make simultaneous observations in two fields separated by a large angle.
2. Scan roughly along a great circle through both fields.

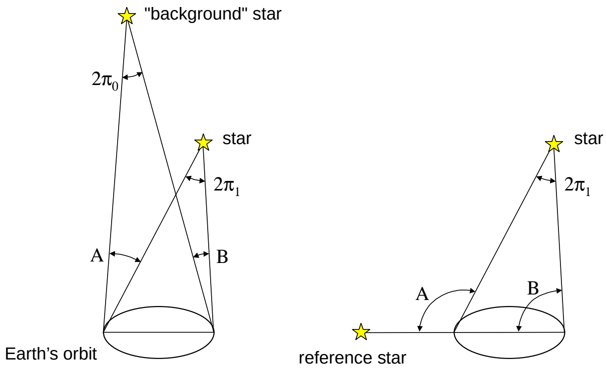

Figure 4. Illustrating the principle of absolute parallax measurement: in the left diagram, the measurement of the (small) angles A, B only allows to determine the relative parallax $\pi_1 - \pi_0 = (A - B)/2$. By contrast, measuring the large angles in the right diagram allows to obtain the absolute parallax $\pi_1 = (A - B)/2$ independent of the distance to the reference star.

3. Make (mainly) one-dimensional measurements, viz., along scan.
4. The 'basic angle' between the two fields must be stable and known (Section 3.4).
5. Repeat measurements as many times as required and with scans in varying orientations in order to cover, eventually, the whole sky.

Perhaps some explanation is needed as to why measurements must be (mainly) one-dimensional. Ideally the angle between any two objects could be obtained by orienting the instrument such that both objects are precisely on the instantaneous scanning great circle (through the two field centres). This is not practical and in any case we want to measure not just these two objects but a large number of them that are simultaneously inside the fields. What is measured then is the projection of the desired angle onto the scanning circle (Figure 5). It is easy to see that when the angle between the two objects is of order 90 deg, then the difference between that angle and its projection on the scanning circle is of order $\phi_1\phi_2$, where ϕ_i are the across-scan angles. Since $|\phi_i| < 10^{-2}$ rad for a one-degree field, it follows that the tolerance for across-scan errors may be up to 100 times more relaxed than along scan.

The astrometric instrument is therefore optimized for measurements in the along-scan direction. This explains the shape of the telescope pupil (elongated along scan) and of the pixels (elongated across scan), the method of CCD readout (drift-scanning or TDI mode), as well as many other aspects of the design.

The basic one-dimensionality of observations also motivates the scanning law. In order to obtain parallaxes as accurately as possible it is desirable to maximize the along-scan component $p_{||}$ of the parallax. With reference to Figure 6 we have a total parallactic displacement of $p = \pi R \sin\theta$, where $R \sim 1.01 \pm 0.017$ is the Sun–spacecraft distance in au, and θ the angle between the star and the Sun. But $p_{||} = p \sin\eta$ and from the sine theorem

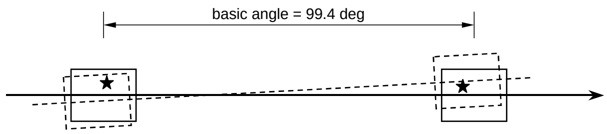

Figure 5. Why the astrometric measurements are mainly one-dimensional, namely along the scan (arrow). Imagine two stars simultaneously observed respectively in the preceding and following astrometric fields (rectangles). To first order, the measured angle between the stars, projected along-scan, is independent of the orientation of the instrument (solid versus dashed lines).

$\sin\theta \sin\eta = \sin\xi \sin v$, so

$$p_{||} = \pi R \sin\xi \sin v \tag{5}$$

where ξ is the solar aspect angle. ξ should therefore be made as large as possible. In practice it is limited by the size of the sun shield in relation to the total height of the service module plus payload, since the latter must be in permanent shade. For thermal reasons, and to maximize the flow of parallax information, the solar aspect angle should moreover be kept at this maximum value all the time, which naturally leads to a revolving scanning law. $\xi = 50°$ will be used for Gaia.

Since ξ is fixed, while objects are in the mean uniformly distributed in the along-scan angle v, we expect that the standard error in parallax on average obtains a geometrical factor $\langle (R \sin\xi \sin v)^2 \rangle^{-1/2} \simeq 1.40/\sin\xi$ or 1.83 for Gaia. Detailed numerical simulations, taking into account the simultaneous determination of position and proper motion, results in a marginally larger factor of 1.93 (de Bruijne 2004).

3.3. Time

The scanning law gives full sky coverage in six months and in fact guarantees that every point on the sky is scanned at three or more distinct epochs during each semester, at varying position angle. Thus, already after 1.5–2 years there is good ability to separate all five astrometric parameters and after the nominal mission length of $L = 5$ years, on average 90 astrometric field transits are obtained of a given object. This allows to determine several more parameters, if needed, including complete orbits for astrometric binaries (or exoplanet systems) having periods less than ~ 5 years.

Three aspects of the temporal distribution of measurements are essential: (i) that the total time span is long enough for proper motions; (ii) that the total number of measurements is significantly redundant even for rather complex objects – think of an extra-solar system with multiple periods; and (iii) that the distribution is quasi-irregular which reduces frequency aliasing and extends the frequency interval that can be searched for periodic phenomena. Concerning the proper motions, it can be noted that the standard errors improve as $L^{-3/2}$, where a factor $L^{-1/2}$ comes from the accumulated number

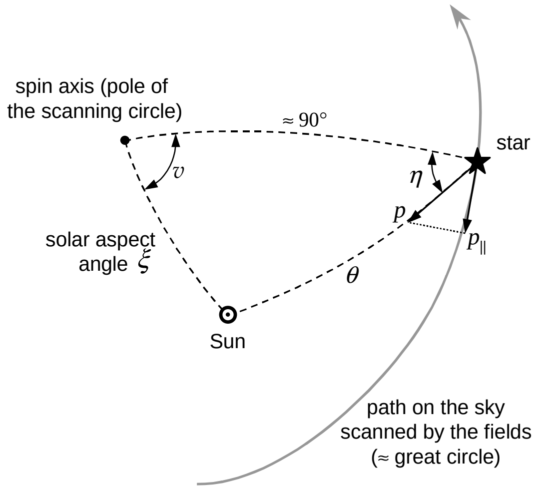

Figure 6. Geometry of parallax determination through one-dimensional scans: the apparent displacement (p) of a star due to its parallax is always directed towards the Sun. Consequently, in order to have a significant fraction of this displacement along the scan ($p_{||}$), one needs a significant solar aspect angle (ξ).

of measurements, and another factor L^{-1} from the increased temporal baseline.

Although the scanning law for Gaia is of course optimized primarily for the basic astrometric and photometric needs, it has sufficient redundancy and built-in quasi-irregularity to cope with rather complex objects.

3.4. Reference Frames

The astrometric data analysis uses *forward modelling* from a very large set of model parameters representing the celestial reference frame, the astronomical objects, the celestial orientation (attitude) of the instrument, geometric transformations through the optical instrument as well as the geometric, photometric and electronic characteristics of the detectors. In principle all these parameters have to be adjusted to agree as well as possible with the sample data (Figure 3) collected for all the objects throughout the mission.

Alternatively, one can think of this process as an inverse mapping from sample data to a series of reference frames including abstract pixel coordinates, angular coordinates in the field of view, angular coordinates on the sky, and finally the astrometric parameters. Each stage of this mapping must be calibrated by adjusting the corresponding mapping parameters by means of a subset of well-behaved stars. Gaia is therefore essentially a *self-calibrating instrument*: no special observation mode and no special kind of data are needed to establish this mapping.

Of course, this process will only work if things are sufficiently simple and stable to be adequately modelled with a limited set of parameters, and if there are enough data available to determine them. To ensure that this is really the case at every working level of the instrument may be one of the most difficult aspects of the design. Usually it entails a compromise between the frequency of physical variations and the frequency of available data. I will illustrate this by a few examples.

The celestial frame. If objects moved more or less arbitrarily on the sky there would be no hope to fix a celestial reference frame. Fortunately, the great majority of stars have very simple motions that are adequately described by the standard set of five astrometric parameters (α_0, δ_0, π, $\mu_{\alpha*}$, μ_δ). As Hipparcos demonstrated, this is true even though most of the stars in reality might be double. Perhaps 20–30% of the objects will be unsuitable for this purpose (and possibly many more of the nearby stars), but they can be iteratively sifted out by a simple examination of the residuals. The remaining 'well-behaved' stars form a sufficiently large body of reference objects for most calibration purposes, including the definition of the celestial frame. A subset of extragalactic objects will be used to attach this to the International Celestial Reference Frame.

The instrument attitude. The scanning motion of the astrometric telescopes is modelled by means of a set of continuous functions of time representing the celestial orientations of the instrument axes. (These axes are implicitly defined by the geometric calibration model of the instrument, which is not further discussed here.) Numerically, the attitude model could for instance consist of a spline representation of the components of the quaternion $\mathbf{q}(t)$ (Wertz 1978) relating the scanning reference frame to the celestial frame. This model implies a degree of smoothness of the motion, e.g., that the quaternion components are adequately described by a certain polynomial over a certain time interval. In technical terms this puts a requirement on the high-frequency part of the power spectrum of attitude perturbations, including perturbations caused by the attitude control system itself. On the other hand, it also requires that the complementary low-frequency part of the motion can be adequately determined from observations. In practice the maximum break frequency is $\sim$0.15 Hz, as set by the CCD integration time of 3.3 s. Assuming that the high-frequency requirement can be met by suitable design, the critical question is then whether there is a sufficient number of stars to determine the 3-axes instrument orientation in any given 3.3 s interval. The answer is yes: even when both viewing directions are at high galactic latitudes there will be some 200 stars brighter than 15 mag in the combined astrometric field of view. Even if only half of them can be used for attitude determination, the equivalent along-scan precision is <25 μas, almost negligible compared with the precision from an individual CCD transit, about 250 μas at 15 mag. Clearly the important fact is simply the large number of suitable reference objects (>100) present in the field at any instant.

The basic angle. One of the many geometric calibration parameters is the basic angle, i.e., the angle between the two sky projections of the centre of the combined astrometric field. The basic angle is by itself not more critical than other geometric calibration parameters – they are all needed to map pixel coordinates into the scanning reference frame – but to ensure its stability is certainly technically more demanding. The value of the basic angle is obtained from the closure condition on the great circle approximating a 360 deg spin of the satellite; in principle, therefore, it is independently calibrated four times per day. On all time scales shorter than this (< 6 h) it has to be stable to μas level. On the other hand, on longer time scales the stability requirement can be quite relaxed.

Geometric calibration of CCDs. Since the CCDs are operated in TDI mode (continuous read-out in synchrony with the satellite spin), the geometrical calibration of the CCDs is not so much concerned about the position of individual pixels as with the mean along-scan coordinates of the whole pixel columns. These are expected to changed over the mission mainly due to progressive changes in the photometric characteristics of the CCDs, which affect the definition of the mean column coordinates. On which time scale can this calibration be carried out? Since the mean coordinate will have a colour term, we need a substantial number of well-behaved stars to cover a sufficient range in colour index as well as being bright enough to contribute significant weight to the calibration. A single CCD column covers a sky area of $2\times$ (0.1325 arcsec)$\times$60 arcsec s^{-1} or 0.1 deg^2 per day. If well-behaved, <15 mag stars have a mean density of 300 deg^{-1}, it follows that observations need to be accumulated over a month or so in order to allow a reliable calibration of individual CCD columns. This sets a corresponding requirement on the geometric/photometric stability of the instrument.

4. CONCLUSIONS

Taking into account the various aspects discussed in Section 3 we conclude that Gaia is designed to perform within a small factor (~ 2) of the theoretical limits set by physics, geometry of the observations, and calibration requirements. This presents a great technical challenge, but fortunately one that European industry is eager to accept and where it has demonstrated great competence.

The basic design concept of Gaia has remained practically unchanged over a number of years, during which numerous details of the design have been intensively studied, modified, re-optimized, and studied again by industry, ESA and the scientific working groups. Considerations similar to those in Section 3.4 may convey some feeling for the basic soundness of the concept: there is something almost uncanny about how well it is adapted to satisfy the various stability requirements on every level. I suspect however that this is not a coincidence but a legacy of the great ideas behind Hipparcos.

ACKNOWLEDGMENTS

Although this paper gives my rather personal interpretation of ‘how Gaia works’, it is based on ideas and contributions from numerous individuals over a period of at least 30 years. The bold and original idea to make astrometry from space was the singular contribution of P. Lacroute; however, it would have come to nothing without the practical and creative minds of people like E. Høg and M.A.C. Perryman.

REFERENCES

Dahn, C.C., Harris, H.C., Vrba, F.J., et al., 2002, AJ 124, 1170

de Bruijne, J.H.J., 2005, ESA SP-576, this volume

de Bruijne, J.H.J., 2004, Gaia technical report Gaia-jdb-012

ESA, 1997, The Hipparcos and Tycho Catalogues, ESA SP–1200

Landau, L.D., Lifshitz, E.M., 1977, Quantum Mechanics, Pergamon Press

Lindegren, L., 1978, in F.V. Prochazka & R.H. Tucker (eds.), Modern Astrometry, IAU Coll. 48, Univ. Obs. Vienna, p. 197

Mignard, F., 2005, ESA SP-576, this volume

Paresce, F., Delplancke, F., Derie, F., et al., 2003, in W.A. Traub (ed.), Interferometry for Optical Astronomy II, Proc. SPIE 4838, p. 486

Perryman, M.A.C., de Boer, K.S., Gilmore, G., et al., 2001, A&A 369, 339

Wertz, J.D. (ed.), 1978, Spacecraft attitude determination and control, ASSL Vol. 73, D. Reidel