GAIA: Global astrometric interferometer for astrophysics

L. Lindegren[1] and M.A.C. Perryman[2,3]

[1] Lund Observatory, Box 43, S-22100 Lund, Sweden
[2] Astrophysics Division, European Space Agency, ESTEC, Noordwijk 2200AG, The Netherlands
[3] Sterrewacht Leiden, Postbus 9513, 2300RA Leiden, The Netherlands

Abstract. — GAIA is a preliminary concept for an astrometric mission, recently recommended within the context of ESA's 'Horizon 2000 Plus' long-term scientific programme. In its present form, the experiment is estimated to lead to positions, proper motions, and parallaxes of some 50 million objects, down to about $V = 15$ mag, with an accuracy better than 10 microarcsec, along with multi-colour multi-epoch photometry of each object. The scientific case for such a mission is dramatic: distances and kinematical motions for tens of millions of objects, throughout our Galaxy, would be obtained—the expected accuracy is such that direct (trigonometric) distance estimates to the galactic centre would be accurate to 10%, with transverse motions accurate to about 1 km s^{-1} at 20 kpc. As 'by-products', the global measurements would yield unprecedented information on the space-time metric (γ to a precision of about 1 part in 10^6 or better, close to values which might distinguish currently competing theories of gravity), angular diameters of hundreds of stars, and a vast body of information on double and multiple systems. Perhaps the most dramatic of these subsidiary goals would be the possibility of screening some 100 000 stars within 100 pc for periodic photocentric motions, which would provide the most powerful and systematic method of detecting possible planetary companions proposed to date.

Key words: astrometry — space vehicles — planetary systems — relativity — instrumentation: interferometers — Galaxy: general

1. Introduction

The resurgence of interest in astrometry today is dramatically illustrated by the fact that, in 1989, the European Space Agency (ESA) developed and launched a satellite entirely dedicated to high-accuracy stellar positional measurements. This elaborate mission has been an unprecedented success. The appreciation of the continuing fundamental importance of astrometry, and its potential for the future, has inspired the proposal of three astrometric missions within the former Soviet Union, elaborate (wide-angle) ground-based astrometric facilities and a variety of proposed space programmes in the United States, the recent commencement of studies of an astrometric space mission in Japan, and the proposal for an astrometric mission within the ESA 'M3' round.

The Hipparcos project represented the first step in the acquisition of astrometric data from space. Positions, annual proper motions, and parallaxes have been determined with accuracies in the range 1−2 milliarcsec, and the concept of carrying out a global astrometric survey from space, resulting in a rigid positional reference frame, and

providing *absolute* trigonometric parallaxes and proper motions, has been fully demonstrated, with the final results expected to be available during 1996 (Kovalevsky et al. 1995; Lindegren et al. 1995; Perryman et al. 1995).

The European Space Agency's 'Horizon 2000 Plus' scientific programme is now being formulated. Covering the period 2005−16, it contains three 'cornerstone' missions: one of these will be an interferometric observatory (Bonnet 1995). Assuming that an appropriate target accuracy can be demonstrated, the Survey Committee has recommended that such an interferometric mission be dedicated to astrometry (Battrick 1994), building on the results of the Hipparcos mission.

A future astrometry mission could be enhanced, with respect to Hipparcos, in terms of accuracy, number of objects, and limiting magnitude. The GAIA concept (Lindegren & Perryman 1994a, b; Lindegren et al. 1993; Lindegren et al. 1993b) provides orders of magnitude improvement in all of these parameters. In addition, consideration is being given to the parallel determination of the sixth astrometric parameter, the radial velocity. This would greatly enhance kinematical studies, improve the identification and understanding of multiple systems, and

furthermore allow the effects of non-linear (orbital) photocentric motion to be decoupled from the effect of perspective acceleration (Favata & Perryman 1995).

The Hipparcos satellite's photometric observations have also demonstrated the power of an all sky-photometric survey from space. Optimisation of the GAIA photometric system could lead to a further profound improvement in the study of stellar variability, and also to spectral classification through intermediate-band multicolour observations at the instrument's focal surface. Determination of spectral type, luminosity class, metallicity, and spectral peculiarities may be feasible, and this would lead to GAIA performing the most dramatic census of the galactic stellar population contemplated to date.

A comparative summary of the performances of GAIA and Hipparcos is given in Table 1. The historical development of astrometric accuracies, and the role played by this potential future mission are summarised in Fig. 1.

Table 1. A comparison between the scientific capabilities of GAIA with those of Hipparcos. Note that about 1 million objects are included within the Tycho Catalogue derived from the Hipparcos satellite. A parallax accuracy of 10 μas (targetted at $V = 15$ mag) corresponds to a 10% distance accuracy at 10 kpc, while a proper motion accuracy of 10 μas/year corresponds to about 1 km s^{-1} at 20 kpc

	Hipparcos	GAIA
Limiting magnitude (V)	~ 12	$> 15 - 16$
Completeness (V)	7.3–9.0	$> 15 - 16$
Number of objects	120 000	50 000 000
Accuracy (all five parameters)	1–2 mas	5–20 μas

2. General scientific objectives

Astrometric measurements provide model-independent estimates of basic geometrical and kinematical properties of astronomical sources. Traditionally the most important applications are the determination of stellar distances, motions and masses. GAIA will provide an immense quantity of extremely accurate astrometric and photometric data from which ultimately all branches of astrophysics will benefit. In the areas of the physics and evolution of individual stars and of the Galaxy as a whole the impact will be immediate and profound. On a much more modest scale this process will begin already with the availability of the Hipparcos results. However, while Hipparcos can probe less than 0.1% of the volume of the Galaxy by direct distance measurements, GAIA will encompass a very large fraction of the Galaxy within its parallax horizon, including much of the halo, and even touching on the nearest companion galaxies such as the Magellanic Clouds. Figure 2 provides a schematic overview of some objects and

phenomena that can be studied by optical astrometry at various levels of accuracy and sensitivity.

Another goal of astrometry is the establishment of an accurate set of reference directions for dynamical interpretation of the motions of the Earth and other solar system objects, and of stars within our Galaxy as a whole, i.e., an optical reference system. The importance of having access to a dense and accurate inertial reference frame is easily overlooked in comparison with the more direct benefits of, say, parallaxes. However, the ability to establish small systematic deviations in the patterns of motions, whether it concerns the search for transneptunian planets or dark matter in our Galaxy, crucially depends on the accuracy of the reference system. Through its global survey nature, GAIA is ideally suited to provide an extremely dense and accurate reference system. In contrast to the situation prevailing with Hipparcos, where the inertial frame must be materialised indirectly (Lindegren & Kovalevsky 1995), the limiting magnitude of GAIA is sufficiently faint to allow a direct link with the extragalactic system by observation of quasars.
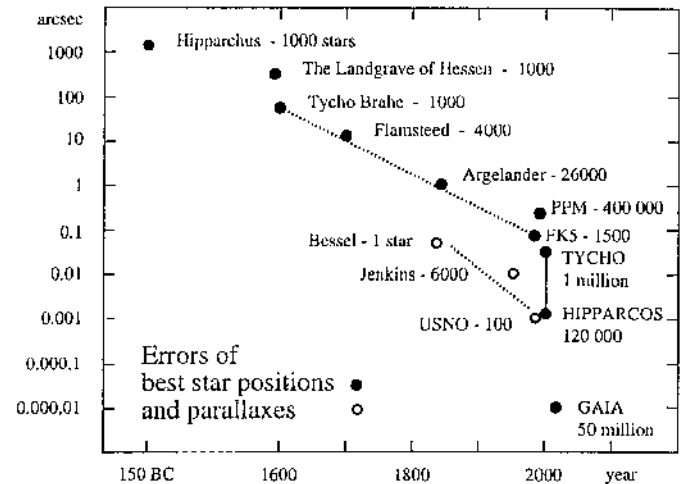

Fig. 1. The development of astrometric accuracies over the past two thousand years, illustrating the advance made by the Hipparcos mission, and the potential accuracies achievable by GAIA (courtesy of E. Høg)

In this section, we provide a concise (and evidently non-comprehensive) scientific survey, emphasising topics where the estimated astrometric accuracies achievable with GAIA lead to clearly identifiable applications. In addition to these, the multi-colour, multi-epoch, sub-millimagnitude photoelectric photometry will open up vast areas for study related to stellar stability over the entire Hertzsprung-Russell diagram. Other possible topics have not yet been carefully assessed, but appear promising—amongst these one may note the detection of MACHOs (Alcock et al. 1993; Aubourg et al. 1993; Høg 1995; Høg et al. 1995), and the possible determination of

stellar angular diameters from the modulation depth of the interferometric fringes—a preliminary assessment indicates that the angular diameters of many hundreds of (nearby, giant) stars may be measurable in this way.

A 5-year mission has been considered, not only because of the resulting improvement in the achievable astrometric accuracy compared to a shorter mission, but also in order to establish an appropriate temporal baseline for the dynamical studies of asteroids, for the determination of the parameters (including orbital motion) of double and multiple stars, for the detection of possible planetary and brown dwarf companions, and for the photometric variability studies.

Throughout this paper, we shall assume that accuracies of the order of $10-20$ μas are achievable on both parallaxes and annual proper motions, although significantly better accuracies should be available for brighter objects, and the availability of improved photon-counting detectors would lead to a further increase in the astrometric performances and limiting magnitudes.
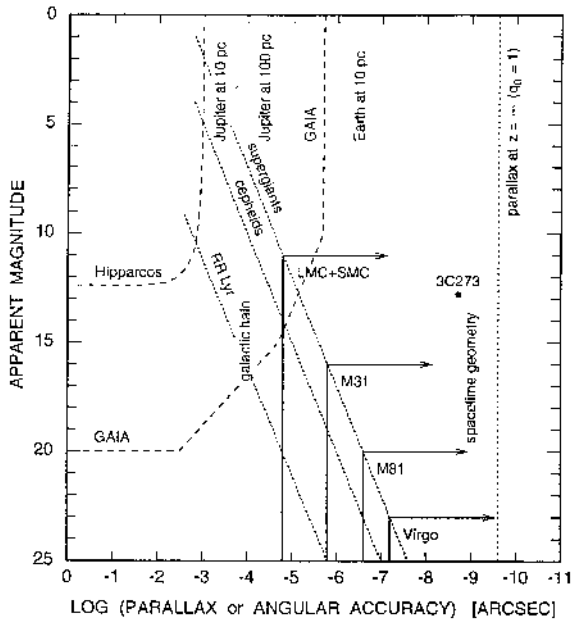

Fig. 2. Schematic overview of some objects and phenomena that can be studied by optical astrometry at various levels of accuracy and sensitivity. The region probed by Hipparcos (limited by 10^{-3} arcsec accuracy and magnitude 12.5) is indicated by the dashed curve at the upper left corner. The parameter region of interest for GAIA is also sketched. It includes direct (trigonometric) distances to all galactic supergiants and Cepheids visible from the sun, RR Lyrae stars and bright halo stars within 10 kpc from the sun, as well as the brightest stars in the Magellanic Clouds (LMC and SMC). Jupiter-like planets can be detected from the photocentric motions of stars out to a few hundred parsecs

2.1. *Physics and evolution of stars*

Stellar luminosities: luminosity estimates are based exclusively on determinations of stellar distances, themselves determined directly only from measurements of trigonometric parallaxes. With Hipparcos, direct distance measurements are limited to 100 pc with, say, 10% precision, a distance horizon insufficient to include rare but astrophysically important categories of stars such as O stars, Cepheids, and RR Lyrae variables. From ground-based observations, distance determinations and luminosity calibrations have been restricted to the main sequence, with indirect distance estimates, for example based on statistical calibrations, used to estimate stellar distances and luminosities for rarer spectral types. Parallaxes with a precision of 20 μas would reach to 5 kpc with 10% accuracy, or to 10 kpc with 20% accuracy. For the first time this would provide an extensive network of distance measurements throughout a significant fraction of our Galaxy, including the galactic centre, spiral arms, the halo, and the bulge.

Massive stars: although only a small fraction of stars in the Galaxy are more massive than 20 $M_{\odot}$, such stars, which spend most of their short lives as H-burning O-type stars, play an important role in galactic structure and evolution. Thus, accurate knowledge of the luminosity of these stars is important for comparing masses derived from stellar evolutionary models with those derived from stellar atmosphere models, for determining initial mass functions, and for studying stellar evolution in the high luminosity/high mass region of the Hertzsprung-Russell diagram. The absolute magnitudes of O stars are presently poorly determined (no O star is sufficiently close to the sun to have a trigonometric parallax accurately measured), the absolute visual magnitudes coming primarily from O stars in clusters and OB associations whose distances are themselves uncertain, but are typically around $1-2$ kpc. Typical apparent magnitudes are $V = 4-6$ mag.

Novae and nova-like variables: distance determinations to novae are required to interpret the energetics of the outburst, and to place these objects more securely within the context of evolutionary models. Distance estimates can be made through modelling of the shell expansion velocity, but such applications are restricted to particular periods after outburst, and also suffer from modelling uncertainties. Most galactic novae are brighter than $V = 12$ mag at maximum, although measurements to $V = 16$ mag or fainter would also allow the determination of distances to galactic novae observed over the last few decades. Related objects, such as dwarf novae, AM Her stars, symbiotic stars, and cataclysmic binaries could be studied, providing accurate luminosities needed to distinguish among alternate possible energy generation mechanisms. Many such nova-like variables would lie within the distance horizon and the magnitude limit (say, brighter than $V = 16$ mag) necessary to provide distances to better than 5%.

Planetary nebulae: planetary nebulae appear to indicate a very narrow mass range for the remnant star, and thus provide the possibility of being good distance indicators. However, because of their rarity, and therefore their typical distances, no satisfactory method yet exists for the their distance estimation. Parallax measurements of the central stars would lead to significant advance in the understanding of the formation and evolution of the shells, the status of the central stars, and the role of these objects as standard candles. Many tens of planetary nebulae would be measurable down to $V = 16$ mag.

Cepheids and RR Lyrae stars: in addition to the importance of these stars to models of stellar structure and evolution, Cepheids and RR Lyrae form the cornerstone of the extragalactic distance scale. Some 55 Cepheids and 26 RR Lyrae stars are known to lie within about 1000 pc, and are already contained with the Hipparcos observing programme, but most of these lie beyond about $300-400$ pc so that individual distance estimates from Hipparcos will be of limited value. In contrast, parallaxes at the 20 μas level would yield distance estimates to better than 2% for these objects. In turn, the details of their period-luminosity-colour relationship would also be significantly improved.

Stars in Open Clusters: only two open clusters (Hyades and Coma Ber) lie within the 100 pc distance horizon yielding distances from Hipparcos (from individual objects within the clusters) to better than 10% accuracy. Cluster studies are important for numerous reasons, mostly related to the fact that they represent a co-eval population of stars with well-defined initial chemical compositions. They can thus be used to follow the development of the formation of our Galaxy, and as a testbed for theories of stellar evolution. For example, they are amenable to studies of their dynamical behaviour, for the calibration of stellar luminosities and distances via properties such as the Wilson-Bappu effect and the mass-luminosity relation for binary stars, and for the calibration of the main sequence as a function of age and metallicity. Some 30 open clusters are considered to lie within about 500 pc, sufficient to provide individual distances to 1% accuracy with parallaxes measured at the 20 μas level.

Globular clusters: little or no information will be provided by Hipparcos on the internal dynamics, and luminosity calibration, of stars within globular clusters, due to the faint magnitude and high central density of stars in these clusters. However, Hipparcos will provide a reference system with respect to which proper motions of cluster stars can be derived from ground-based observations acquired over long periods of time. Ages of globular clusters have indicated a possible discrepancy with the age of the Universe derived from present estimates of H_0 and Ω_0. Many observational effects and theoretical complications make interpretation of globular clusters properties far from straightforward; but cluster age determinations

essentially require absolute magnitude calibrations of the main-sequence and, in particular, the turn-off point as a function of chemical composition. For absolute ages to be accurate to a billion years, essential for a resolution of the age conflict, the distance of the cluster must be determined with an accuracy of better than 3%. An observing programme reaching 15 mag and 20 μas accuracy on the parallaxes would include 20 or more globular clusters (such as 47 Tuc, ω Cen, M3, M5, and M15) lying between 5 and 10 kpc, and would yield *individual* distances accurate to better than $10-20$%. Some 10 or more of the brightest stars per cluster would be observable, resulting in mean cluster distances at least a factor of three better than these individual accuracies.

Metal-poor stars and primordial nucleosynthesis: recent determinations of beryllium and boron abundances in the metal-poor star HD 140283 have raised important questions about the origin of these elements: whether they originate from a high cosmic ray flux at the birth of the Galaxy, or from primordial nucleosynthesis (Gilmore et al. 1991). If the former possibility is ruled out, it would seem to indicate that the standard Big Bang model is wrong, and that newer inhomogeneous models would be required. Cosmic ray spallation, in contrast, makes a specific prediction of the B/Be ratio, although the Be abundances turn out to be very sensitive to whether the star is a subgiant or a dwarf. A clear parallax determination (ground-based parallaxes are generally quite inadequate) would clarify this question. While the specific instance of HD 140283 should be resolved by the Hipparcos measurements, this example illustrates the importance of individual parallax determinations for astrophysical studies, for example in allowing tests of both the inhomogeneity generated in the early universe at the quark-hadron phase transition, and the chemical and cosmic-ray evolution of the early Galaxy.

2.2. *Dynamics of stellar systems*

Visual and astrometric binaries: astrometric measurements at the accuracy achievable with GAIA will be able to resolve binaries with an apparent separation exceeding $2-5$ mas and a moderate magnitude difference. The astrometric and photometric characteristics of the components can be measured. For numerous systems with periods of a few years the absolute orbits can be determined and hence the individual component masses. Closer binaries, and systems with a faint companion, can in many cases still be detected, and their orbits determined, from the non-linear motion of the photocentre, as has been demonstrated in the case of the Hipparcos data by Bernstein & Bastian (1995). At the sensitivity level of GAIA some 25% of all stars may turn out to be non-single. This vast material on stellar duplicity will be essential for a correct interpretation of the astrometric and photometric parameters, as well as providing important constraints on theories of stellar formation.

Interacting binary systems: a rich variety of astrophysical problems related to interacting binary systems would become accessible with parallaxes in the $10-20$ μas range, and semi-major axes of the orbital systems of the order of tens of microarcseconds. That such systems could be detected and, under favourable circumstances, orbital elements solved, is again inferred from the work reported by Bernstein & Bastian (1995). The evolutionary history of interacting binary systems, and the origin of Type I supernovae, millisecond pulsars, low mass X-ray binaries, and globular cluster X-ray sources is intimately bound up with the behaviour of compact binaries with mass transfer and loss. Accurate knowledge of the stellar masses and orbital separation can be derived from astrometric measurements (yielding the orbital separation and the orbital inclination) combined with estimates of the mass function determined by radial velocity measurements. Many specific questions about accretion rates, precursors, mass distributions, and kinematic behaviour could be addressed with these data, including studies of the black hole candidates. Galactic black hole candidates have bright secondaries (9 mag in the case of Cyg X-1, and 12 mag or fainter in the case of V404 Cyg) and wide orbits (with orbital periods of about 6 days), which should yield definitive black hole masses by determining orbital separation and inclination. Evidently, further study is required to assess the true capabilities of GAIA for these systems as a function of magnitude and magnitude difference, separation, orbital period, and source structure.

Be star X-ray binaries: these are believed to consist of a recently formed neutron star and a Be star companion. The orbit has not yet circularized, and the eccentric motion produces periodic eruptions at periastron as the compact star passes through the mass outflow from the Be star. Measurement of the orbital parameters would yield information on the anisotropy of the supernova mass ejection. It is important to relate this to the kinematics of isolated pulsars, and to the physics of the explosion.

Dynamics of globular clusters: accurate proper motions of stars within globular clusters are required to yield information on the cluster's internal velocity dispersion, and thus constrain dynamical models of their formation and evolution. Within 47 Tuc, for example, proper motions of 20 μas/yr correspond to transverse velocities of 0.4 km/s. In addition, spectroscopic binaries have been detected in globular clusters with amplitudes of tens of km/s and periods of years, corresponding to separations of order 1 mas. Parallaxes and annual proper motions at the level of 50 μas or better would provide distances and orbital data necessary to clarify the formation and evolution of these binary systems, and their role in the formation of the millisecond pulsars now known to exist in the cores of globular clusters. Source crowding will severely limit the relevance of GAIA to these problems in the case of the 'baseline detection mode', while such applications will benefit considerably from the possible implementation of 'direct fringe detection' (see Sects. 3.5 and 5).

Galactic dynamics: the huge number, impressive accuracy, and faint limiting magnitude of the GAIA mission would totally revolutionise the dynamical studies of our Galaxy, which are now understood to be capable of providing considerable advances in our understanding of the structure and motions within the spiral arms, the disc and the outer halo. It is still unclear, for example, whether spiral arms are density wave enhancements in the background stellar distribution, or whether they are regions of enhanced star formation. If a density enhancement exists, it will affect the stellar motions in a characteristic way, and this could be tested on the relatively nearby Perseus arm (about 2 kpc distant). Stellar motions would be determined for stars near the arm, both foreground and background (distinguished by their parallaxes), with proper motion accuracy requirements of some 100 μas/yr, corresponding to about 1 km/s in space velocity. Similarly, the dynamical relationship between the disk and halo is still uncertain.

Dark matter within the disk: two recent programs have attempted to determine the surface density of the Galactic disk in the solar neighbourhood. One (Bahcall, Flynn & Gould 1992) finds evidence for dark matter in the disk, while the other (Kuijken & Gilmore 1989a, b) finds none. The issue is important because of the implications on the nature of dark matter—whether, if it exists, it admits matter in a baryonic or non-baryonic form. One source of uncertainty in the Bahcall et al. result, based on the distribution of K giants perpendicular to the galactic disk, is the error in the distances to individual stars. These are relatively bright objects (apparent magnitudes 10 and brighter); their distance scale could be recalibrated and substantially improved by direct parallax measurements of K giants with a range of metallicities. Parallaxes at the level of 50 μas would be required.

The mass of our Galaxy: the form of the rotation curve beyond the Sun is very sensitive to the existence and amount of dark matter near to it. No very reliable determination of the rotation curve has yet been derived. In an extension of the programmes being undertaken with the Hipparcos data, measurements of the distances and motions of disk stars are required at a range of galactic longitudes, resulting in the rotation curve at distances out to 15 kpc determined from stars with $V < 12$ mag. Parallaxes at the level of 20 μas for an accuracy of 20% in individual distances, and annual proper motions of about 200 μas/yr (or 10 km/s), would be required.

Proper motions of the Magellanic Clouds, and active galactic nuclei: different explanations for the dynamical behaviour of the LMC/SMC, in particular whether these systems are gravitationally bound to our own Galaxy, implies systematic proper motions of below about 1 mas/yr, very much at the limit of the Hipparcos capabilities. Large

numbers of stars measured in the LMC/SMC at the level of 50 μas or better, would clarify their dynamical relationship with our own Galaxy. Further out, the nuclei of active galaxies may be sufficiently point-like that their absolute proper motions may be measurable. At a redshift of 0.03, a transverse velocity of 1000 km/s corresponds to a proper motion of 3 μas/yr. Thus, transverse velocities of nearby galactic nuclei due, e.g., to galaxy cluster potentials, might be detectable. Further investigations will be required to determine whether the flux in truly unresolved points within such galactic or active galactic nuclei is suffient—and sufficiently stable with time—for such investigations to be feasible with GAIA.

The role of quasars: future astrometric missions, at levels of accuracy very much better than 1 μas, could determine the transverse motions of external galaxies and quasars routinely, and determine their kinematic properties independently of a dynamical model of the Universe. In the meantime, an astrometric programme reaching 15−16 mag would include a number of quasars (see, for example, Schilbach & Scholz 1995), and this in turn would allow a direct tie between the resulting reference system and an inertial reference frame, something which has not been possible (directly) in the case of Hipparcos. The considerable importance of this possibility is that the resulting proper motions of all the stars within the global observing programme would not be subject to arbitrary offsets in their proper motions, a fact critical for any dynamical interpretation of their motions.

2.3. Detection of planetary systems and brown dwarfs

The search for possible planetary systems beyond our own solar system has received considerable attention in recent years, with ground-based searches based on accurate radial velocity measurements being undertaken (e.g., Kürster et al. 1994), and a variety of other ground- and space-based possibilities based on imaging in the optical or in the infra-red, or dedicated astrometric techniques being considered.

The Hipparcos mission, due to its limited astrometric accuracy and mission duration, is unlikely to make any serious contribution to the possible detection of subsolar mass planetary companions around nearby stars, although the principles of the method have been convincingly demonstrated (Bernstein & Bastian 1995). The essential idea, that of detection of non-linear photocentric motions in the paths of nearby stars due to such 'planetary' companions, has been extensively studied and simulated, for example as part of NASA's TOPS (Towards Other Planetary Systems) and later as part of the POINTS programme. Preliminary studies in the context of GAIA have been reported by Casertano et al. (1995) and Casertano et al. (1995).

The probable size of the effect can be judged by considering the path of the Sun as seen from a distance of (say) 10 pc. The perturbation caused by Jupiter has an amplitude of 500 μas and a period of 5 years, while the effect of the Earth is a one-year period with 0.3 μas amplitude. With a mission length of 5 years and a target mission accuracy of 20 μas, GAIA should be able to provide annual normal points with an accuracy of 50 μas. This is sufficient to detect Jupiter-mass planets (at the 3σ level) out to 30 pc. This volume includes several thousand potential target stars, *all* of which can be monitored for possible companions. If the accuracy is instead 2 μas, which may be feasible for bright stars ($V < 10$ mag), then the detection horizon for Jupiter-mass planets is pushed beyond 100 pc and includes some 10^5 candidate stars. Screening all 50 million stars down to the survey limit of $V = 15-16$ mag for possible signatures of planetary and brown dwarf companions will provide a complete census of such bodies to well-defined detection limits.

2.4. General relativity

The reduction of the Hipparcos data has necessitated the inclusion of stellar aberration up to terms in $(v/c)^2$, and the general relativistic treatment of light bending due to the gravitational field of the Sun (and Earth). Light bending by the Sun amounts to 4 mas even at $90°$ (i.e., for light arriving perpendicular to the ecliptic). The astrometric residuals may be tested for any discrepancies with the prescriptions of general relativity; in principle this provides a constraint on the post-Newtonian light-bending term, γ, equal to unity in general relativity. Previous determinations of γ include the original observations of Dyson et al. (1920), and other optical determinations made at times of total solar eclipse, including the most recent published determinations of Jones (1976) based on the 1973 Mauritanian eclipse. All such previous determinations have been based on observations within a few $R_\odot$ of the solar limb. Other determinations based on VLBI observations and the Viking spacecraft (Shapiro) time-delay, also restricted to measurements made close to the solar limb, are also illustrated (see Soffel 1989). The best available measurements to date are those from the Viking observations (Reasenberg et al. 1979), and provide a precision of about 10^{-3} in γ.

The GAIA measurements would provide a precision of about 1 part in 10^6 or better in the determination of γ due to the Sun. Interestingly, this is close to the values predicted by those present theories which predict that the Universe started with a strong scalar component, and which is relaxing to the general relativistic value with time (e.g., Damour & Nordtvedt 1993). The importance of such theories is that they provide a possible route to the quantisation of gravity. For this reason, space experiments dedicated to the measurement of γ with a precision of about 1 part in 10^6 have been proposed, and are being considered for the ESA M4 round. We stress that GAIA would

provide this precision as a by-product of its astrometric and photometric campaign.

Light deflection has also been observed, with various degrees of precision, on distance scales of $10^9 - 10^{21}$ m, and on mass scales from $1 - 10^{13} M_\odot$, the upper ranges determined from the gravitational lensing of quasars (Dar 1992). Light-bending by the Earth is at the level of ~ 40 μas, and GAIA could therefore extend the domain of observations by two orders of magnitude in length-scale, and six orders of magnitude in mass. (The Pound-Rebka experiment verified the general relativistic prediction of a gravitational redshift for photons, an effect probing the time-time component of the metric tensor, while light deflection depends on both the time-space and space-space components). Light bending at the Jovian limb is predicted to amount to 17 mas.

At the level of accuracy expected from GAIA, even more subtle effects will start to become apparent, such as the quadrupole components of the gravitational fields of the Sun and the planets, and the 'frame-dragging' effects of their motions and rotations (see, e.g., Soffel 1989). Light modulation effects due to gravitational lensing by MACHOs (Alcock et al. 1993; Aubourg et al. 1993) might also be detectable (Høg 1995; Høg et al. 1995), and all of these possibilities will need to be studied in more detail.

The possible detection of gravity waves, as a result of the angular displacement of the apparent positions of stars has been considered by Fakir (1994). The speculative possibility that GAIA might achieve the first ever such detection, through the apparent motions of stars situated behind relatively nearby binary (giant) systems, has been briefly considered by Fakir (1995).

3. Technical considerations

3.1. Significance of global measurements

A future space astrometry mission aiming at the widest range of astrophysical problems must be able to perform *global* measurements. This means that positions, and changes in positions caused by proper motion and parallax, are determined in a reference system consistently defined over the whole sky. In practice this requires direct and accurate measurement of large angles (of the order of 1 rad). The need for wide-angle astrometry in order to construct a distortion-free reference system around the sky is fairly self-evident and need not be elaborated here. But wide-angle measurements are also crucial for the determination of *absolute parallaxes*, which may require some explanation.

The trigonometric parallax is the most important and direct way for determination of stellar distances. The traditional (ground-based) method to determine the parallax is to measure the annual oscillations of the star's apparent motion, as caused by the Earth's revolution around the Sun, against the background of distant stars. These measurements are confined to the field of view of a long-focus telescope, usually within $10-20$ arcmin. The measurable quantity is then the *relative* parallax, i.e., the difference between the parallax of the target star and the mean parallax of the background stars. To obtain the true or *absolute* parallax of the target star one must add a correction corresponding to the mean parallax of the background stars; this is often estimated simply from their mean magnitude and galactic coordinates. The size of the correction is typically $1-2$ mas, with a (model dependent) uncertainty of a substantial fraction of a milliarcsec (van Altena et al. 1988). This procedure is clearly inadequate when it comes to microarcsec astrometry.

The inestimability of absolute parallaxes in a small field depends on the circumstance that the parallax factor f, being proportional to the sine of the angle from the sun to the star, is nearly constant in the small field. The relative displacement of two stars with parallax p_1 and p_2 is consequently given by $f \cdot (p_1 - p_2)$. From a series of measurements made at different times of the year (with different f) it is possible to determine $p_1 - p_2$, but not the individual parallax values.

The situation is radically different if wide-angle measurements can be made with the same accuracy as in a small field. The relative displacement must now be written $f_1 p_1 - f_2 p_2$, where in general $f_1 \neq f_2$ for stars separated by a large angle (~ 1 rad). Measuring this displacement at different times of the year clearly allows to determine p_1 and p_2 individually, in other words, to obtain the absolute parallaxes.

This principle of absolute parallax determination by wide-angle measurements was implemented for the first time in the Hipparcos mission, and has been demonstrated to be a complete success: several independent tests on preliminary Hipparcos parallaxes have shown that their zero-point errors are much less than 1 mas, and possibly smaller than 0.1 mas. The same principle will apply to the GAIA mission, which therefore should be able to produce parallaxes that are absolute at the sub-μas level.

3.2. Outline of a mission concept

From a technical viewpoint, the general principles of the GAIA concept can be summarized in the following points:

a) it is a continuously scanning instrument, capable of measuring simultaneously the angular separations of several hundred star images as they pass across a field of view of about $1°$ diameter. Simultaneous multi-colour photometry of all astrometric targets is a necessary and integral part of the concept;

b) high angular resolution in the scanning direction is provided by optical interferometry on a baseline of ~ 3 m, the large field of view implying the use of Fizeau type interferometers, and the maximum baseline considered here being limited by the requirement for a non-deployable payload compatible with the Ariane 5 launcher envelope;

c) the wide-angle measuring capability is realized by means of two or more interferometers mechanically set at large angles to each other and scanning the same great circle on the sky. The precise 'basic angles' between the interferometer baselines are determined from the 360° closure condition on each great-circle scan, while short-term (< 3 hours) variations are passively controlled and monitored by internal metrology;

d) the whole sky is systematically scanned according to a pattern which permits, after only a few years of scanning, a complete separation of the astrometric parameters describing the motions and distances of single stars. A longer temporal baseline permits the determination of additional parameters relevant for instance to the detection of planetary companions.

Within this general outline a multitude of options exist which remain to be explored, optimized and weighed against each other. These include, for instance, the number and optical design of the Fizeau interferometers, choice of wavelength bands, detection systems, basic angles, metrology system, satellite layout, and orbit. A technically reasonable configuration which achieves the mission goals summarized in Table 1 is described in Lindegren & Perryman (1994a). Its main characteristics are as follows: three identical Fizeau interferometers (A, B, C) stacked on top of each other in a cylindrical body intended to fit into a dual-launch Ariane 5 envelope, pointed at angles of (for example) 54° (A−B), 78.5° (B−C), and 132.5° (A−C); each interferometer consisting of two 50 cm (or larger) aperture mirrors, with a baseline length of 2.5 m, and equipped with CCD detectors in the focal plane assembly. Further increase in the baseline length would result in a greatly increased complexity, since the payload structure would then have to be deployable.

While two interferometers are in principle sufficient to perform wide-angle astrometry, the addition of a third one provides full redundancy in the sense that any two of them can be used and still achieve the mission goals. Mindful of the criticality of alignment of an operational interferometer, the design concept involves a further level of redundancy: the outer part of the field of view is used in a non-interferometric mode (incoherent imaging, Sect. 3.6) which, even in the case of a total failure of the interferometric capability, would still provide an impressive overall astrometric accuracy (Table 3).

Parts of the proposed concept are derived directly from the Hipparcos mission and applied to the much more ambitious GAIA mission. We should underline here the three most important reasons why such a dramatic improvement in accuracy, compared with Hipparcos, can be considered achievable: a) increased optical baseline, from 0.3 m to 2.5 m; b) increased detector efficiency (going from a photon-counting image dissector tube to a CCD); c) multiplexed rather than sequential observations, resulting in increased observing time for target stars, and the possibility to observe all stars down to the instrument's sensitivity limits.

3.3. Scanning law and mission length

In contrast to point-and-stare missions as exemplified by the Hubble Space Telescope and the OSI and POINTS astrometry projects, the observation mode adopted for GAIA is based on continuous sky scanning. The rationale for this particular choice, which deeply affects nearly all technical, operational and scientific aspects of the mission, may be summarized as follows: a) it follows a proven concept, namely that of the Hipparcos mission; b) it is well suited for a global, survey-type mission with very many targets; c) astrophysical research primarily oriented towards the physics of stars and of the Galaxy as a whole is best served by a survey-type mission encompassing large and well-defined samples of a wide variety of objects; d) from a technical viewpoint, a continuously scanning satellite appears to be particularly efficient, and perhaps optimal, in terms of instrument stability and calibration, and observing time utilization.

The continuous scanning approach results in a variety of very important attributes, in particular resulting in the facts that the satellite provides a constant geometry with respect to the sun for the proposed scanning pattern; that critical instrument parameters such as the basic angle, scale value and geometrical field distortion are obtained from the closure conditions on each complete rotation of a few hours; that these calibrations are part of the normal observations and therefore do not require any overhead time; that no overhead is required for re-pointing the telescope; and that many objects are observed strictly simultaneously.

The interferometers are considered to scan the whole sky according to a pre-defined pattern similar to the 're-volving scanning law' used with Hipparcos. Thus, the axis of rotation (perpendicular to the viewing directions and interferometer baselines) is kept at a nominally fixed angle ξ from the sun, and describing a precessional motion about the solar direction at constant speed with respect to the stars. For Hipparcos the solar angle was $\xi = 43°$, but a slightly larger value (e.g., 55°) should be adopted for GAIA for improved uniformity of sky coverage. The spin period is about 3 hours, corresponding to a speed of 120 arcsec/s for the motion of star images across the field. The spin axis must be controlled to follow the nominal scanning law to within a few arcminutes in order to guarantee sufficient overlap between successive great-circle scans.

A baseline mission length of 5 years is proposed. Although the many advantages of the space environment permit very significant measurements to be made in a relatively short time, some of the more complex motions cannot be properly explored on a time scale of 2−3 years. This concerns, in particular, many binaries with periods

up to several years, and the detection of possible planetary and brown dwarf companions. As previously noted, the survey instrument is ideal for planetary detection, through its screening of a very large number of stars for their photocentric motions (Lindegren & Perryman 1994a; Casertano et al. 1995; Casertano et al. 1995), but measurements extending over a significant fraction of the orbital period are mandatory.

3.4. Size of interferometers, integration time and field of view requirements

The ultimate accuracy with which we can determine the direction to a point source of light is set by the dual nature of electromagnetic radiation, namely as waves (causing diffraction) and particles (causing a finite signal-to-noise ratio in the detection process). For an interferometer consisting of two (small) apertures separated by the baseline B, the fundamental limit for the directional accuracy perpendicular to the fringes is given by:

$$\sigma \geq \frac{\lambda}{2\pi B\sqrt{N}} \quad (1)$$

where λ is the wavelength and N the number of detected photons (Lindegren 1978). Given that $N \propto D^2$, where D is the diameter of each aperture, it follows that $\sigma \propto (BD)^{-1}$, and that the product BD should in general be maximized subject to the various constraints on size, mass and complexity. While a detailed trade-off must await further technical studies we have adopted $B = 2.45$ m and $D = 0.55$ m as representing a plausible interferometer compatible with the Ariane 5 launcher capabilities.

Equation (1) gives the limiting accuracy of a single angular measurement in the instrument frame. The positions, proper motions and trigonometric parallaxes of all the objects are eventually determined (in principle) by a least-squares combination of all such measurements collected over the whole mission. The accuracy of, say, the parallax of an individual star depends on the total weight of the measurements of the star, and a certain numerical factor determined by the geometry and temporal distribution of the measurements. For the parallax accuracy we may write, by a slight generalisation of Eq. (1):

$$\sigma_p = F\frac{\lambda_{\text{eff}}}{2\pi BR} \quad (2)$$

where F is a geometrical factor and R the total signal-to-noise ratio. Monte Carlo simulation of a revolving scanning law with $\xi = 55°$ gives $F \simeq 2$. Writing R instead of $\sqrt{N}$ allows the inclusion of certain other noise sources such as sky background.

A simple calculation based on Eq. (2) shows that, in the continuous scanning mode, where the total integration time per scan is limited, the required astrometric accuracy can only be reached with a sufficiently large field of view. For $B = 2.45$ m and $\lambda_{\text{eff}} = 500$ nm we have

$F\lambda_{\text{eff}}/(2\pi B) = 13$ mas. However, we wish to reach, at the end of the mission, a parallax error which is a thousand times smaller. This requires an accumulated signal-to-noise ratio of 10^3 per star, or at least 10^6 detected photons from the star. At 15 mag this translates into a total integration time of $\tau \sim 1000$ s, assuming a throughput of 10 m^{2}nm [the throughput is defined as $ATQ\Delta\lambda$, where A is the collecting (aperture) area, T the optical transmittance, Q the quantum efficiency, and $\Delta\lambda$ the wavelength bandwidth]. Now, for a continuously scanning instrument the total time spent inside the field of view by an average object is given by:

$$\langle\tau\rangle = \frac{\Omega}{4\pi}L \quad (3)$$

where Ω is the solid angle of the field of view and L the (effective) mission length. For $\langle\tau\rangle \sim 1000$ s and $L \sim 5$ years this gives $\Omega \sim 0.3$ deg^2 as a *minimum* field size.

Such a large field can only be realized in a Fizeau type interferometer. The use of Michelson interferometers appears to be ruled out by the adopted continuous scanning mode.

3.5. Spatial resolution of detector

Full utilization of the astrometric information of the diffraction images in principle requires that the spatial coordinates and, for a scanning instrument, the time, are recorded to adequate accuracy for each detected photon. Let us briefly consider the option to place such an (ideal) detector directly in the focal plane. The period of the Young's fringes produced by the interferometer is $\lambda_{\text{eff}}/B \simeq 40$ mas at 500 nm wavelength. The fringes must be recorded with a sufficiently high spatial resolution to retain most of the positional (phase) information. The sampling theorem requires only two samples per period, but considering also the modulation damping produced by the finite spatial and temporal sample width, we find that about four samples per fringe period is actually required. This corresponds to a pixel size of 10 mas. For a one-degree field this implies some 360 000 resolution elements (or pixels) across the detector. (In the perpendicular direction a much lower resolution can be used, because of the smaller aperture extent normal to the baseline.) This is clearly beyond what can realistically be achieved with current technology such as CCDs.

We do not exclude that detectors of sufficient resolution and geometrical linearity will become available in the coming years as a result of ongoing or potential development programmes. Such a possibility might be offered by new developments in superconducting detectors, which provide better efficiency, broader wavelength response (into the UV and infrared), lower noise, and energy resolution (Perryman et al. 1993; Lindegren & Perryman 1994b). However, the present goals can be achieved by means of conventional detectors, if a *modulating grid* is used to encode the phase information of

the fringes. The proposed modulation/detection scheme is outlined in Sect. 4.3.1.

3.6. *Coherent and incoherent imaging mode*

Operating the telescope as a Fizeau interferometer requires that the Airy disks formed by the two 0.55 m openings add coherently to form interference fringes. The condition is that the Airy disks fall at the same position in the focal plane (to within a fraction of the Airy radius), and that the optical pathlengths in the two interferometer arms are equal (to within a fraction of the coherence length $\lambda^2/\Delta\lambda$). For the nominal instrument these conditions can certainly be met in the central part of the field. Further away from the optical axis the fringe contrast may be seriously reduced, but the image still contains substantial astrometric information. This suggests that the outer part of the field can profitably be used in an *incoherent* imaging mode for additional astrometric and photometric measurements. The term 'incoherent' here simply means that the information is derived from the Airy envelope of the stellar images instead of the interference fringes. The theoretical loss of astrometric precision is $2B/D \simeq 9$, but in practice the loss is smaller because of other design factors.

For the incoherent imaging, CCD detectors can be used directly in the focal plane and operated in the Time Delayed Integration (TDI) mode. That is, the charge images are shifted along with the optical images and read out as they arrive at one side of the array. The principle of operation in this mode corresponds exactly to that of a proposed non-interferometric astrometric mission, Roemer (Lindegren et al. 1993a) or Roemer+ (Høg 1994), and the predicted performance is even slightly better than for Roemer+.

The addition of the incoherent imaging mode brings about some very significant advantages and a considerable strengthening of the overall mission performance:

a) compared with the use of a modulating grid, the CCD directly in the focal plane gives better light economy and much more efficient suppression of background radiation and nearby images, resulting in a fainter limiting magnitude ($V = 20$), albeit at a reduced accuracy;

b) the optimum wavelength bandwidth to be used in the coherent (interferometric) mode is largely determined by the pupil geometry. In contrast, the passbands for the incoherent mode can be chosen quite freely to serve a number of different purposes, including intermediate and narrow bands selected for astrophysical reasons (such as Strömgren and Hβ photometry to $V \sim 16$ mag) or especially designed for transformability to other photometric systems (Young 1994);

c) the incoherent imaging mode provides attitude information needed both in real time, for the operation of the detectors, and a posteriori, as required for proper interpretation of the coherent measurements in the data reductions;

d) positions determined by the incoherent imaging can be used as starting points for a solution based on the interferometric data. This brings the latter solution into the linear regime (phase errors $\ll 1$ rad), allowing considerable simplification of the processing and eliminating possible ambiguities and spurious solutions;

e) the CCD detectors surrounding the interferometric field can be used for coarse focussing and control of mirror alignments, since the star images from the two pupils must overlap completely and in all parts of the field when the mirrors are correctly adjusted. Since this function does not depend on the formation of fringes, it should be particularly useful for the first phase of fringe acquisition;

f) finally, the incoherent imaging mode provides a reliable and useful 'backup' facility in a scenario where one or more of the interferometers, for some technical reason, fails to produce fringes. In such a case the interferometric detectors do not produce any astrometrically useful data at all, whereas the incoherent mode still provides photometry at nearly the full accuracy and astrometry certainly at the sub-milliarcsec level.

4. The payload

4.1. *Fizeau interferometers*

As indicated in Sect. 3.4, a field of view size of the order of $1°$ is needed in order to accumulate sufficient integration time on each object. This implies that the interferometer is of the Fizeau type, i.e., optically equivalent to a single large telescope with two openings in the entrance pupil. The non-illuminated parts of the mirrors are removed, and we are basically left with two off-axis telescopes with a common focus.

In a first approximation we have considered the optical characteristics of the single 'underlying' 3 m telescope. An effective focal length of $10-15$ m is desirable for a convenient linear size of the field (a few tens of cm). Diffraction-limited imaging over a large field implies good correction for spherical aberration, coma and astigmatism, which suggests a three-mirror system. Very preliminary optimization of such a system led to a design whose main parameters are given elsewhere (Lindegren & Perryman 1994a; Loiseau & Shaklan 1995). The mirror diameters are calculated for an unvignetted field of $\pm0.8°$. The resulting Fizeau system is shown in Fig. 3.

The nominal performance of the three-mirror telescope is diffraction limited within a field radius of about $0.45°$. The outer part of the field, between $\simeq 0.5°$ and $0.8°$ radius, is still unvignetted and the geometrical images are still much smaller than the Airy disk of the 0.55 m openings; consequently this part of the field can be used for the incoherent imaging.
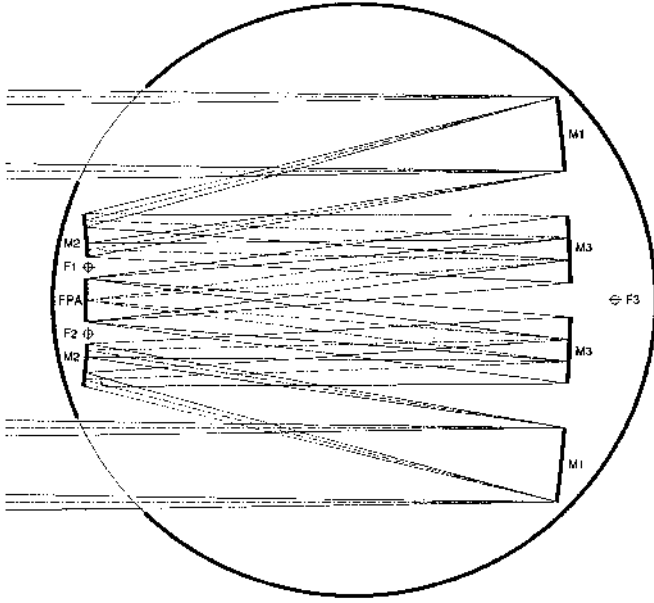
Fig. 3. Optical outline of one of the Fizeau interferometers. It is optically equivalent to a three-mirror telescope (M1, M2, M3) preceded by a double aperture. FPA = focal plane assembly. F1, F2, F3 indicate three fiducial points for controlling the mirrors and focal plane assembly (see text). This configuration results from a ray-tracing solution using the Code V package (Loiseau & Shaklan 1995)

The Fizeau interferometer fits into a circular envelope of 4.4 m diameter, with the focal plane assembly conveniently located near the envelope for passive cooling of the detectors. The layout permits internal baffling so that, for instance, the tertiary mirror is invisible from the outside. Three identical interferometers are stacked on top of each other in a cylindrical envelope, with baselines set at different angles (Sect. 3.2). These should be carefully chosen so that any two interferometers, or all three of them, can be used to perform wide-angle measurements (Makarov et al. 1995).

4.2. Alignment and stability requirements

For diffraction-limited performance the relative positions of all mirror elements need to be controlled, passively or actively, to within a fraction of a wavelength. In particular the superpositioning of the Airy disks from the two circular openings must be achieved over the whole one-degree field, which puts very stringent demands on the symmetry of the system. However, not all degrees of freedom need to be controlled to the same accuracy, and passive means should be used whenever possible. Active control must be based on error signals that can be derived from laser gauges, but also from the focal plane detectors. The metrology requirements for the optical alignment are

modest (some nm) and may be achieved with commercial gauge technique.

Much more stringent requirements are set by the stability requirements, and in particular variations of the angles between the interferometer baselines (the basic angles). While variations on time scales longer than a few hours (equal to the rotation period of the spinning satellite), as well as the absolute values of the basic angles, can be deduced from the closure conditions on each complete great-circle scan, all significant short-term variations must either be eliminated or monitored by the metrology system. On a 3-m baseline, 5 μas corresponds to a linear shift of one end by 70 pm. Thus, monitoring at the $10-50$ pm precision level is likely to be required, but only for *variations* occurring on time scales up to a few hours. The basic feasibility of differential laser gauging at the pm level has already been demonstrated in the POINTS and OSI projects (Noecker et al. 1993; Shao 1993).

The interferometer baseline orientation is defined by an arbitrary fixed point on the focal grid in combination with all the mirror elements. It can therefore be monitored with respect to three fiducial points, e.g., as indicated in Fig. 3 (F1, F2, F3), by measuring distances from these points to different parts of the mirrors and focal assembly. The basic angle can then in turn be monitored from the distances of the fiducial points of one interferometer with respect to the other.

4.3. Focal plane assembly

4.3.1. Coherent imaging mode (interferometer)

In the baseline concept the inner part of the focal surface, dedicated to the interferometric measurements, is covered by grids consisting of many narrow bands parallel to the interference fringes of the stellar images. The spatial frequency of the bands match the fringe frequency, thus producing a sinusoidal modulation of the light intensity behind the grid as the images (and fringes) move across the field, perpendicular to the bands. The grid may be designed to modulate either the amplitude of the image (by means of alternatingly opaque and transparent bands, or slits, as was used for Hipparcos), or the phase (by means of a phase hologram). In both cases the light modulation at the detector can be understood as interference of the two pupil images superposed in different spectral orders, as diffracted by the grid (Lindegren et al. 1994). The phase grid is the preferred design in view of its higher efficiency (Fig. 4).

Figure 5 shows a possible layout of the focal field. The inner part of the field is covered by a mosaic of 32 CCD detectors with corresponding grids, working at several different effective wavelengths. Each grid is divided into 'subfields' of 27×13.5 arcsec^2 to limit the background light and confusion from other stars.

An integrated 'light curve' of the modulation can be recorded by a CCD in which the electric charges generated by the photons are shifted back and forth at the (known) frequency of modulation. In principle $2n-1$ pixels are sufficient to store n phases of the light curve, where n may be in the range 4 to 8. At intervals of $1/n$ of the modulation period the charge image is shifted one pixel forward, and after $n-1$ such shifts it is quickly shifted back to its original position, immediately starting the integration of the next modulation cycle. The shift frequency must be accurately matched to the actual modulation frequency (depending on the grid period, the satellite rotation rate, and the image scale) to allow the light curve to be built up over many hundred modulation cycles before the whole CCD is quickly read out and the relevant samples saved. A field lens (Fig. 4) ensures that the illumination on the CCD is stationary during this process, by imaging the two circular openings of the interferometer at fixed points on the CCD.

The multiple shifting back and forth of the charges requires a high charge transfer efficiency, but should be quite feasible with current performance figures. A similar use of multiple hidden, fast image buffers on a CCD chip is proposed for the second generation of the Zurich Imaging Stokes Polarimeter (ZIMPOL II, Stenflo et al. 1992). Alternative detectors such as photon-counting avalanche photodiodes (APD) should also be considered in conjunction with a modulating grid. Although not yet available in arrays, commercially available discrete APD detectors have characteristics that are very well suited for the present application (size of sensitive area, high quantum efficiency over a very wide spectral range, low dark counts with moderate cooling, and high time resolution). The possible advantages of the more speculative superconducting tunnel junction detectors have already been noted.

Some additional design considerations are outlined in the following paragraphs.

a) Fringe/grid matching: for a narrow wavelength passband with effective wavelength λ_{eff} the diffraction image of the nominal interferometer will consist of a set of Young's fringes having an angular period of λ_{eff}/B radians, and with an outer envelope in the shape of an Airy disk corresponding to one of the pupils (radius to first minimum $= 1.22\lambda_{\text{eff}}/D$). The grid period must equal the fringe spacing in order to produce light modulation at the detector. Thus, the grid period is necessarily linked to the effective wavelength.

b) Optical bandwidth: the finite bandwidth ($\Delta\lambda$) reduces the fringe contrast as one moves away from the centre of the image. To get good modulation by the grid it is necessary that fringes are seen across the whole central Airy disk. This condition requires that $\Delta\lambda/\lambda_{\text{eff}} \lesssim D/B$. The optimum bandwidth is a compromise between the number of photons and the degree of modulation ob-
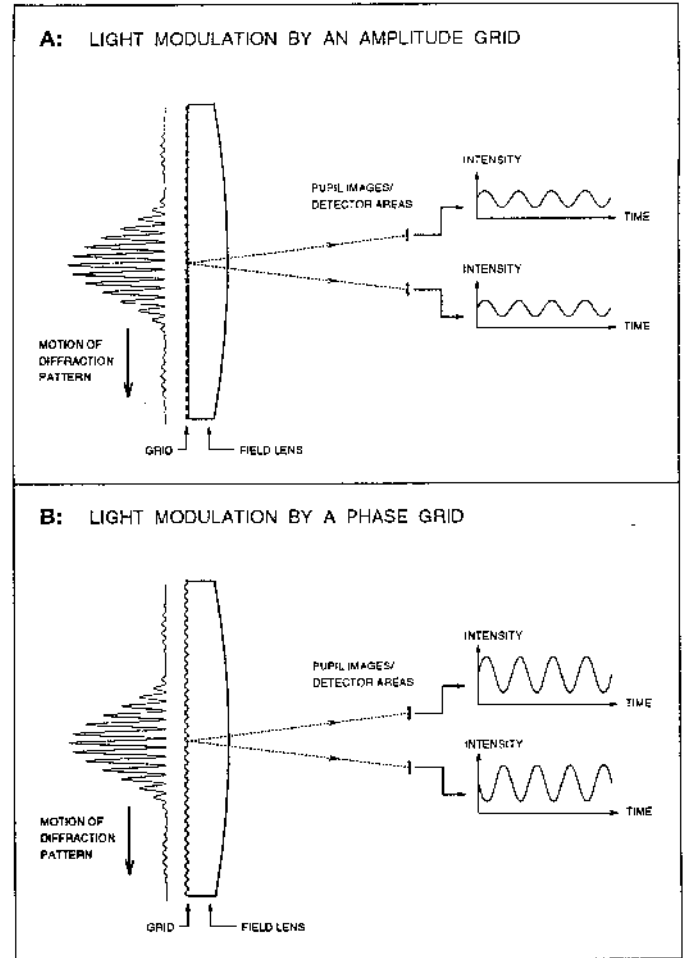
Fig. 4. Principle of light modulation by an amplitude grid **a)** and a phase grid **b)**. The phase grid transmits more light than the amplitude grid and gives deeper modulation, both factors improving the accuracy of the phase determination. However, for the phase grid the two images of the interferometric pupils are modulated in anti-phase, requiring separate detection of the two images

tained. Provisional calculations indicate that the standard error on the phase measurement is minimized for $\Delta\lambda/\lambda_{\text{eff}} \simeq 1.1D/B \simeq 0.25$ for the present pupil configuration.

c) Size of subfields: the light of any (unresolved) source falling on the grid will contribute to the detector output by superposing a sinusoidal component on the signal. Although it is possible to disentangle many such components in the data analysis (see below), it is of course desirable to obtain 'clean' signals for most observations, and only resort to the more complex analysis when necessary, e.g., for double stars and crowded fields. The instantaneous field of view covered by a single detector channel should therefore be small enough that, most of the time, there is no disturbing star within that area. At the limiting magnitude of about $V = 15.5$ mag we may consider any

Fig. 5. Schematic layout of grids and detectors in the focal plane. The inner part of the field, used in the interferometric (coherent) mode, is covered by a mosaic of phase grids, using different passband filters (A, B, ..., F) and a separate CCD behind each grid. The grids are divided in subfields with field lenses as shown in Fig. 5. The outer part of the field is used for incoherent imaging, using CCD detectors placed directly in the focal plane and operated in Time Delayed Integration mode. The detectors labelled w, u, v, b, y, I, Bw and Bn are preceded by colour filters for intermediate and narrow band photometry. Each CCD has two regions that are read out separately to increase the dynamic range. The pixels are long and narrow

star brighter than $V = 16.5$ mag as a potentially serious disturber. The mean density of such stars (Allen 1973) is about 3000 deg^{-2}. For the probability of disturbance to be less than 10%, say, the area must then be smaller than about 400 arcsec^2. The average sky background within this area is equivalent to one star of magnitude 16: this contributes an acceptable level of photon noise but no modulation. In the design of Fig. 5 the subfield area is 365 arcsec^2.

d) Image synthesis: in contrast with the standard aperture synthesis problem, the 'image' to be analyzed by GAIA is usually not a continuous brightness distribution, but consists of a relatively small number of point sources. Each point source must however be modelled by several photometric and geometric parameters corresponding to

its spectrum, proper motion and parallax. The resulting multi-dimensional synthesis problem requires a reasonable observational coverage not only in the uv plane, but also in the time domain. The 'revolving scanning law' (Sect. 3.3) provides sufficient temporal distribution and diversity of scanning directions. The uv coverage is further improved by covering the whole useful spectral range by overlapping passbands of $\Delta\lambda/\lambda_{\rm eff} \simeq 0.25$, each band corresponding to a different spatial frequency in the given scanning direction.

4.3.2. Incoherent imaging mode

The outer part of the unvignetted 1.6° field includes a number of CCD detectors (20 as sketched in Fig. 5) placed directly in the focal plane and operated in the Time Delayed Integration (TDI) mode. The pixel width, 6 μm = 108 mas, is matched to the size of the Airy disk of the 0.55 m pupils ($1.22\lambda/D = 160$ to 360 mas) but completely damps out any interference fringes (period $\lambda/B = 30$ to 65 mas). Some of the CCDs are equipped with intermediate and narrow-band colour filters, e.g., corresponding to a modified $uvby\beta + I$ system. These provide accurate and astrophysically important photometric information on all stars observed by the interferometer and for selected fainter stars. The remaining CCDs have no filters, and are used for complementary astrometric measurements and for attitude determination and coarse mirror control.

Table 2. Astrometric performance of GAIA in its interferometric (coherent imaging) mode, using a modulating grid (baseline) and with a hypothetical high-resolution detector (direct fringe detection). Stars of spectral type G0 have been assumed, and a mission length of 5 years. A dash means that the star is too faint to be measured against the background within the 27×13.5 arcsec^2 subfield assumed for the baseline option. Units: mas = milliarcsec for parallaxes (p), mas/yr for proper motions (p.m.)

V [mag]	Baseline option		Direct fringe detection	
	p [mas]	p.m. [mas/yr]	p [mas]	p.m. [mas/yr]
10	0.002	0.001	<0.001	<0.001
12	0.003	0.002	0.001	<0.001
14	0.006	0.004	0.002	0.001
15	0.011	0.006	0.003	0.002
16	0.020	0.012	0.005	0.003
18	–	–	0.012	0.007
20	–	–	0.030	0.018

5. Accuracy estimates

For the interferometric (coherent) mode, the light modulation recorded by the detectors can be evaluated by

integrating the intensity over the relevant detector area, and with respect to λ weighted by the stellar photon flux, instrument and filter transmittances, and detector quantum efficiency. Such calculations have been carried out using realistic assumptions on all instrumental characteristics. The accuracy of the phase determination, and hence on the angular coordinate of the star in the scanning direction, is then computed for a maximum likelihood estimator. The resulting errors are finally transformed into global astrometric errors by assumptions similar to those used for the Hipparcos accuracy assessment. Results of a preliminary accuracy assessment are given in the first three columns of Table 2 ('baseline option').

Table 3 gives the photon statistical limits on the astrometric and photometric precisions achievable in the incoherent imaging mode. The actual accuracy will be lower depending on how well the geometric and photometric properties of the CCDs can be calibrated, and how stable they are on time scales up to several weeks (the time period needed to achieve a complete mapping of irregularities). However, photometric accuracies better than 0.01 mag should be obtained for all survey stars ($V < 15.5$ mag), as well as sub-mas astrometric accuracy for many more fainter stars.

The columns in Table 2 headed 'direct fringe detection' refer to a hypothetical detector with sufficient spatial resolution to replace the modulating grid, as discussed in Sect. 3.5. In this option we assume a detector with similar characteristics as a CCD, except that the pixel width in the scanning direction is much smaller than the fringe period, say $< 0.5\ \mu$m. As indicated by the table, this would dramatically enhance the performance both in terms of accuracy and limiting magnitude, while at the same time allowing observations in crowded regions, such as in clusters or at low galactic latitudes.

While the use of a modulating grid is considered feasible with existing technology, and therefore retained as the baseline option, we wish to emphasize that it does not fully utilize the potential of the interferometric method. Serious consideration should therefore be given to alternative detection methods which may become feasible in the near future. The inclusion of the 'direct fringe detection' option in Table 2 may serve as an illustration of the improvement that could result from using a nearly optimal detector.

6. Other mission characteristics

The principal spacecraft characteristics needed to accommodate the identified goals are summarised as follows:

Orbit: the satellite could be operated in geostationary orbit, or at the Earth-Moon triangulation libration point L5; the advantages of the latter (Companys et al. 1993) include a reduced particle radiation background, and absence of eclipses; but operation at L5 would demand at least two ground stations, and X-band for telemetry transmission.

Data Rate: a telemetry data rate of several hundred kbits/s is foreseen.

Mass: first estimates indicate a payload mass of below 600 kg, a bus mass of 700−800 kg, a bi-propellant fuel mass of between 600 kg (L5) and 1100 kg (geostationary orbit); giving a total launch mass of between 2100−2700 kg.

Power: the payload and total spacecraft power requirements are estimated at around 400 W and 1100 W respectively.

Deployable elements: these would include only a second antenna to ensure omnidirectional RF coverage, and (possibly) the solar panels (arranged so as to minimise the dynamical perturbations due to solar radiation pressure).

Thermal control: each interferometer would be contained within an actively-controlled thermal environment to ensure geometric stability. The CCDs at the focal surface are specifically arranged to lie close to the outer surface of the spacecraft (Fig. 3) to allow passive cooling to deep space.

7. Comparison with other missions

The success of Hipparcos, and the resulting resurgence of interest in the astrophysical capabilities of astrometry, has led to a variety of recent proposals for space missions dedicated to astrometry. These missions are divided into two categories: those that aim for Hipparcos-type accuracy on a comparable number of stars (i.e., some one hundred thousand objects); and those that aim for a significantly higher astrometric accuracy.

In the former category are three proposed missions from scientists within the former Soviet Union (Lomonosov, Regatta-Astron, and AIST). These have been under consideration for several years, are still in the study phase, and all face uncertain financial support. The objectives of all three missions was to achieve Hipparcos-type (milliarcsec) accuracy on up to a few hundred thousand stars, taking the Hipparcos goals as roughly indicative of the state of the art. Their scientific importance was based on maintaining the reference frame established by the Hipparcos mission throughout the coming decades— the optical reference frame defined by Hipparcos will degrade with time as a result of the uncertainties on the individual proper motions, such that a mission in 10−20 years from now, with individual accuracies in the range 1−2 mas will still have a useful and important, if not fundamentally new, scientific value.

Very recently, a Japanese Astrometric Satellite has been mentioned (Yoshizawa, private communication). This is at an early study phase only, no details have been made available, but it is considered to lie within the same type of mission concept as those described in the previous paragraph.

Table 3. Predicted mean errors in astrometry and photometry for GAIA in its incoherent imaging mode (using the outer part of the field of view). Only photon and readout noise has been taken into account. Stars of spectral type G0 have been assumed, and a mission length of 5 years. Filter and CCD characteristics are given at the bottom. The wide-band channel 'W' is primarily intended for astrometric observation of the faintest stars and is not equipped with a filter. For faint stars, a dash means a signal-to-noise ratio ≤ 2.0 on a single CCD crossing. Units: mas = milliarcsec for parallaxes (p), mas/yr for proper motions (p.m.), and milli-magnitudes for the photometry (calculations kindly provided by E. Høg)

| | Astrometry | | | Photometry [milli-mag] | | | | | | | | |
V [mag]	p [mas]	p.m. [mas/yr]	W	w	u	v	b	Bn	Bw	y	I
8	0.002	0.002	0.1	0.5	0.3	0.1	0.1	0.3	0.1	0.1	0.1
10	0.003	0.002	0.1	1.2	0.7	0.3	0.2	0.7	0.2	0.2	0.1
12	0.008	0.005	0.1	3.0	1.8	0.8	0.6	1.7	0.6	0.6	0.3
14	0.020	0.011	0.1	7.8	4.5	2.0	1.4	4.2	1.4	1.4	0.6
16	0.05	0.03	0.3	22.5	12.1	5.0	3.7	11.1	3.7	3.7	1.6
18	0.14	0.08	0.7	–	44.1	14.0	10.1	39.6	10.1	10.1	4.3
20	0.5	0.3	2.2	–	–	–	38.5	–	38.5	38.8	14.2
Central wavelength [nm]			–	320	350	411	467	486	486	547	800
Filter FWHM [nm]			–	20	30	25	25	3	25	25	140
Peak transmission			–	0.30	0.40	0.60	0.70	0.70	0.70	0.70	0.96
QE of CCD			–	0.64	0.68	0.74	0.78	0.78	0.78	0.76	0.62

Significantly higher astrometric accuracy is the goal of the JPL's OSI and SONATA concepts (Shao 1993; Buscher et al. 1995), and the proposed POINTS mission (Reasenberg et al. 1994a). These programmes aim to reach accuracies at the level of a few μas or better, per observation, with some or all of these *pointed* (rather than scanning) instruments being able to benefit, in a similar way to GAIA, from repeated measurements throughout the mission. Thus the limiting magnitude of POINTS is between $14-18$ mag depending on adopted mission concept, OSI's accuracy target is 5 microarcsec on 20 mag stars, while SONATA is targetting 0.5 microarcsec for 14 mag objects after 4 hours. A recently-proposed low-cost alternative to POINTS is NEWCOMB (Reasenberg et al. 1994b), which currently aims to reach an accuracy of about 100 microarcsec in a single measurement of about 4 min.

However, being pointed rather than scanning instruments, these proposed missions are able to contain only a relatively small target observing list, of say several hundred, or several thousand, pre-selected programme stars. Their high intended accuracy makes these programmes highly interesting for specific scientific goals (such as the detection of sub-solar mass planetary companions around selected nearby stars); however, the relatively few objects which can be observed, and their (generally) relatively bright limiting magnitude, means that these missions do not significantly overlap with, and therefore do not duplicate, the majority of the scientific goals attainable by GAIA.

For completeness, we note that there is a continuing effort to develop ground-based optical interferometers, especially in the U.S.A., building on the success of the Mk III optical interferometer (see, e.g., Hummel et al. 1994). These instruments should ultimately provide very high relative astrometric accuracy over small fields (for example, on bright double or multiple systems) at the 10 microarcsec level (Shao & Colavita 1992; Colavita & Shao 1993; Shao 1995); they are not seriously being considered as tools for large-angle (positional) astrometry, which is now considered to be mandatory for the large-scale determination of parallaxes and proper motions.

8. Conclusion

Building on the Hipparcos experiences, a mission with dramatically enhanced scientific goals has been proposed, and has been recommended for inclusion in ESA's long-term scientific programme Horizon 2000+. Many tens of millions of objects, including solar system and extragalactic objects, would be observed. Astrometric accuracies of the order of 10 μas (better for brighter stars) are achievable. This would yield distances and luminosities to objects within 10 kpc, and motions individually significant *throughout* our Galaxy. Information on double and multiple systems (including stellar masses), an immense data base of multi-colour, multi-epoch photometry, valuable information on the space-time metric, and angular diameters of hundreds of stars, would become available.

Initial system studies indicate that such a mission could be comfortably accommodated within the

(dual-launch) Ariane 5 envelope, and within the financial envelope of the ESA cornerstone missions. The satellite is proposed for 5 years of operations, and could be positioned in either a geostationary, or L5 (Earth-Moon Lagrangian point) location. It would be a first in space interferometry, would continue recent progress in an important scientific discipline, and seems to involve an acceptable balance between overall feasibility and required technological development.

In a sense, the accuracies and scientific goals identified here may be considered as a 'minimum' objective. If considered appropriate, technological developments could permit enhanced scientific performance: in particular we may identify the following as areas in which improvements might occur: i) improved (more efficient, energy sensitive) detectors, including detectors sensitive in the near infra-red; ii) enlarged optical elements (if, for example, light-weighting technology can be combined with further optimisation of the optical layout and associated baffling requirements); iii) larger interferometric baselines (if, for example, a large stable *deployable* structure could be foreseen).

Note added in proof: the Proceedings of the RGO–ESA Symposium on 'Future Possibilities for Astrometry in Space, ESA SP–379, which took place after submission of this manuscript, includes many further scientific and technical details of the topics considered in this paper.

References

Alcock C., et al., 1993, Nat 365, 621

Aubourg E., et al., 1993, Nat 365, 623

Allen C.W., 1973, Astrophysical Quantities. Athlone Press (3rd edition)

Bahcall J.N., Flynn C., Gould A., 1992, ApJ 389, 234

Bernstein H.-H., Bastian U., 1995, Future Possibilities for Astrometry in Space, ESA SP–379, 55

Battrick B., 1994, Horizon 2000 Plus: European Space Science in the 21st Century, ESA SP–1180

Bonnet R.M., 1995, ESA Bull. 81, 6–17

Buscher D., et al., 1995, Appl. Opt. 34, 1081

Casertano S., Lattanzi M., Perryman M.A.C., 1995, Future Possibilities for Astrometry in Space, ESA SP–379, 47

Casertano S., Lattanzi M., Perryman M.A.C., Spagna A., 1995, Boulder Workshop on 'Detection and Study of Extra-Solar Terrestrial Planets', May 1995 (in press)

Colavita M.M., Shao M., 1992, Bull. AAS 182, 5116

Companys V., et al., 1993, Proc. 44th IAF Cong., Graz

Dainty C.J., et al., 1994, Report of the ESA Lunar Interferometry Study Team. ESA Publication (in press)

Damour T., Nordtvedt K., 1993, Phys. Rev. Lett. 70, 2217

Dar A., 1992, Nuc. Phys. B (Suppl.) 28A, 321

Dyson F.W., Eddington A.S., Davidson J., 1920, Phil. Trans. Roy. Soc. 220A, 291

Fakir R., 1994, Astrophys. J. 426, 74

Fakir R., 1995, Future Possibilities for Astrometry in Space, ESA SP–379, 113

Favata F., Perryman M.A.C., 1995, Future Possibilities for Astrometry in Space, ESA SP–379, 153

Gilmore G., Edvardsson B., Nissen P.E., 1991, ApJ 378, 17

Høg E., 1994, IAU Symp. No. 166, Kluwer, p. 317

Høg E., 1995, Future Possibilities for Astrometry in Space, ESA SP–379, 125

Høg E., Novikov I.D., Polnarev A.G., 1995, A&A 294, 287

Hummel C.A., et al., 1994, AJ 108, 326

Jones B.F., 1976, AJ 81, 455

Kovalevsky J., Turon C., 1991, Adv. Space Res. 11, 5

Kovalevsky J., et al., 1995, A&A 304, 34

Kuijken K, Gilmore G., 1989a, MNRAS 239, 571

Kuijken K, Gilmore G., 1989b, MNRAS 239, 605

Kürster, et al., 1994, ESO Newslett. 76, 51

Lindegren L., 1978. In: Prochazka F.V. and Tucker R.H. (eds.), Modern Astrometry, IAU Coll. No. 48, p. 197

Lindegren L., Kovalevsky J., 1995, A&A 304, 189

Lindegren L., Perryman M.A.C., 1994a, internal report submitted to the ESA Horizon 2000+ Survey Committee

Lindegren L., Perryman M.A.C., 1994b, IAU Symp. 166, p. 337

Lindegren L., Perryman M.A.C., Loiseau S., 1995, SPIE Symp. Spaceborne Interferometry II, Orlando, 2477, 91

Lindegren L., Bastian U., Gilmore G., et al., 1993a, Roemer: Proposal for the Third Medium Size ESA Mission (M3), Lund Observatory

Lindegren L., Perryman M.A.C., Bastian U., et al., 1993b, GAIA — Global Astrometric Interferometer for Astrophysics, proposal for a Cornerstone Mission concept submitted to ESA in October 1993

Lindegren L., Perryman M.A.C., Bastian U., et al., 1994, Amplitude and Intensity Spatial Interferometry II, SPIE, Vol. 2200, p. 599

Lindegren L., et al., 1995, A&A 304, 44

Loiseau S., Shaklan S., 1995, SPIE Symp., Spaceborne Interferometry II, Orlando, 2478, 269

Makarov V.V., Høg E., Lindegren L., 1995, Exper. Astron. 6, 211

Noecker M.C., Phillips J.D., Babcock R.W., Reasenberg R.D., 1993. In: Reasenberg R.D. (ed.), Spaceborne Interferometry, SPIE, Vol. 1947, p. 174

Perryman M.A.C., Foden C.L., Peacock A., 1993, Nucl. Instr. Methods A 325, 319

Perryman M.A.C., et al., 1995, A&A 304, 69

Reasenberg R.D., et al., 1979, ApJ 234, L219

Reasenberg R.D., Babcock R.W., Murison M.A., et al., 1994a, Amplitude and Intensity Spatial Interferometry II, SPIE, Vol. 2200, p. 2

Reasenberg R.D., Babcock R.W., Phillips J.D., Johnston K.J., Simon R.S., 1994b, Amplitude and Intensity Spatial Interferometry II, SPIE, Vol. 2200, p. 18

Schilbach E., Scholz R.D., 1995, Future Possibilities for Astrometry in Space, ESA SP–379, 127

Shao M., 1993. In: Reasenberg R.D. (ed.), Spaceborne Interferometry, SPIE, Vol. 1947, p. 89

Shao M., 1995, Ap&SS 223, 119

Shao M., Colavita M.M., 1992, A&A 262, 353

Soffel M., 1989, Relativity in Astrometry. Springer-Verlag

Stenflo J.O., Keller C.U., Povel H.P., 1992, LEST Foundation Technical Report No. 54, Institute of Theoretical Astrophysics, Oslo

van Altena W.F., Lee J.T., Hanson R.B., Lutz T.E., 1988. In: Philip A.G.D. (ed.), Calibration of Stellar Ages. Davis Press, Schenectady, p. 175

Young A.T., 1994, A&A 288, 683